

Институт проблем управления
им. В.А. Трапезникова РАН

УПРАВЛЕНИЕ БОЛЬШИМИ СИСТЕМАМИ

Выпуск 111
Сентябрь 2024

**СБОРНИК
ТРУДОВ**

ISSN 1819-2467

Регистрационный номер Эл №ФС77-44158 от 09 марта 2011 г.

Москва – 2024

РОССИЙСКАЯ АКАДЕМИЯ НАУК

Институт проблем управления

им. В.А. Трапезникова

УПРАВЛЕНИЕ БОЛЬШИМИ СИСТЕМАМИ

СБОРНИК ТРУДОВ

Выпуск 111

Москва – 2024

КООРДИНАЦИОННЫЙ СОВЕТ

Академики: Васильев С.Н., Желтов С.Ю., Каляев И.А., Куржанский А.Б., Попков Ю.С., Черноусько Ф.Л.; члены-корреспонденты РАН: Галяев А.А., Савватеев А.В.; д-ра техн. наук: Кузнецов О.П. (ИПУ РАН), Павлов Б.В. (ИПУ РАН), Рутковский В.Ю. (ИПУ РАН).

РЕДАКЦИОННАЯ КОЛЛЕГИЯ

Главный редактор: академик Новиков Д.А. **Зам. главного редактора:** д-р физ.-мат. наук Губко М.В. **Отв. секретарь:** канд. физ.-мат. наук Горбунова А.В. **Редактор:** канд. техн. наук Квинто Я.И.

Д-ра техн. наук: проф. Алескеров Ф.Т. (ГУ ВШЭ), проф. Алчинов А.И. (ИПУ РАН), проф. Андриевский Б.Р. (ИТМО), проф. Афанасьев В.Н. (МИЭМ, ГУ ВШЭ), проф. Бахтадзе Н.Н. (ИПУ РАН), проф. Бурков В.Н. (ИПУ РАН), проф. Вишневский В.М. (ИПУ РАН), д-р физ.-мат. наук проф. Ерешко Ф.И. (ВЦ РАН), д-ра техн. наук: проф. Калянов Г.Н. (ИПУ РАН), проф. Каравай М.Ф. (ИПУ РАН); д-р экон. наук, проф. Клочков В.В. (ИПУ РАН); д-р техн. наук доцент Коргин Н.А. (ИПУ РАН); д-ра физ.-мат. наук: проф. Кушнер А.Г., проф. Лазарев А.А. (ИПУ РАН); д-ра техн. наук: проф. Лебедев В.Г. (ИПУ РАН), проф. Мандель А.С. (ИПУ РАН); д-р биол. наук проф. Михальский А.И. (ИПУ РАН); д-р экон. наук, проф. Нижегородцев Р.М. (ИПУ РАН); д-ра физ.-мат. наук: проф. Райгородский А.М. (МФТИ), проф. Рапопорт Л.Б. (ИПУ РАН); д-ра техн. наук: проф. Самуйлов К.Е. (РУДН), проф. Сидельников Ю.В. (МАИ), проф. Совлуков А.С. (ИПУ РАН); д-ра физ.-мат. наук: проф. Соловьев С.Ю. (МГУ), проф. Угольницкий Г.А. (ЮФУ); д-ра техн. наук: проф. Уткин В.А. (ИПУ РАН), проф. Хоботов Е.Н. (ИПУ РАН); д-ра физ.-мат. наук: проф. Чхартишвили А.Г. (ИПУ РАН), проф. Щербаков П.С. (ИПУ РАН).

РЕГИОНАЛЬНЫЕ РЕДАКЦИОННЫЕ СОВЕТЫ

Арзамас – д-р физ.-мат. наук проф. Пакшин П.В. (АПИ НГТУ); **Волгоград** – д-ра физ.-мат. наук: проф. Воронин А.А., проф. Лосев А.Г. (ВолГУ); **Воронеж** – д-р техн. наук, проф. Баркалов С.А., д-р физ.-мат. наук, проф. Головинский П.А. (ВГУ), д-р техн. наук, проф. Подвальный С.Л. (ВГТУ); **Иркутск** – д-р техн. наук проф. Зоркальцев В.И. (ЛИН СО РАН), д-р физ.-мат. наук, проф. Лакеев А.В. (ИДСТУ СО РАН); **Казань** – д-р физ.-мат. наук, проф. Маликов А.И., д-р техн. наук, проф. Сиразетдинов Р.Т. (КГТУ-КАИ); **Липецк** – д-ра техн. наук: проф. Погодаев А.К., проф. Сараев П.В. (ЛГТУ); **Самара** – д-ра экон. наук: проф. Богатырев В.Д., проф. Гераськин М.И.; **Петрозаводск** – д-р физ.-мат. наук, проф. Мазалов В.В., д-р техн. наук, доц. Печников А.А. (ИПМИ КарНЦ РАН); **Санкт-Петербург** – д-р физ.-мат. наук: проф. Петросян Л.А. (СПбГУ), д-р техн. наук проф. Фуртат И.Б. (ИПМ РАН).

Адрес редакции: 117997, г. Москва, ул. Профсоюзная, д. 65.

Адрес в интернете: ubs.mtas.ru.

СОДЕРЖАНИЕ

Системный анализ

Искаков М.Б., Искаков А.Б.

Теорема существования равновесия в безопасных стратегиях по Рени 6

Седых И.А., Макаров К.Н.

Идентификация квадратичных комплекснозначных динамических окрестностных моделей на кластеризованных данных и без кластеризации 66

Егоркин А.А.

Определение центральности графа алгоритмом PageRank с учетом весов связей 81

Информационные технологии в управлении

Рей А.С., Широкий А.А.

Комплексная оценка информационных рисков. II: Алгоритм идентификации структуры дерева комплексной оценки интегрального информационного риска 97

Сетевые модели в управлении

Крыгин А.А., Тарасова С.М.

Снижение размерности задачи нахождения критических узлов сети..... 118

Управление в социально-экономических системах

Авдеева З.К., Коврига С.В.

Когнитивные карты для анализа и моделирования в системе стратегического планирования государства 147

Владова А.Ю.

Формирование групповой и индивидуальной траекторий успеваемости по данным e-learning-платформ..... 179

Клочков В.В., Егошин С.Ф.

Организационно-экономический механизм внедрения технологий автоматизации управления воздушными судами в рамках авиатранспортных систем 197

***Управление техническими системами
и технологическими процессами***

**Вандиловская П.А., Крыгин А.А., Лукинова О.В.,
Рощин А.А.**

Метод поиска разрезов графа для задачи управления инженерной инфраструктурой..... 226

Тырсин А.Н., Кашеев С.Е.

Моделирование взаимовлияния скважин для анализа эффективности систем заводнения на малых выборках данных 247

Управление подвижными объектами и навигация

Мартынова Л.А., Павлов А.А.

Метод выбора энергоэкономичного пути обхода сложного препятствия автономным необитаемым подводным аппаратом 266

***Надежность и диагностика средств
и систем управления***

**Баженов С.Г., Вересников Г.С., Голев А.В.,
Гончаренко В.И., Ерофеев Е.В., Лазурин Г.А.,
Скрябин А.В., Тимофеева А.Д., Феденюк В.А.**

Разработка методики и средств проведения ресурсных испытаний для построения системы ранней диагностики технического состояния электромеханического рулевого привода беспилотного воздушного судна..... 286

***Программы и системы моделирования объектов,
средств и систем управления***

Зверкина Г.А., Кошелев А.А.

О моделировании поведения одного марковского процесса на основе метода моделирования случайных величин с помощью интенсивностей 306

ТЕОРЕМА СУЩЕСТВОВАНИЯ РАВНОВЕСИЯ В БЕЗОПАСНЫХ СТРАТЕГИЯХ ПО РЕНИ

Искаков М. Б.¹, Искаков А. Б.²
(ФГБУН Институт проблем управления
им. В.А. Трапезникова РАН, Москва)

Статья является продолжением цикла статей 2018, 2022, 2023 годов посвященного теоретическому обоснованию равновесия в безопасных стратегиях (РБС), как концепции решения игровых задач. Представлен метод конструирования теорем существования РБС из известных теорем существования равновесия Нэша (РН). При этом исходная теорема существования РН приводится к стандартной формулировке, которая, как условие, вставляется в текст мета-теоремы существования РБС. Согласно данному методу из теоремы Рени (1999) существования равновесий Нэша получены и доказаны две альтернативных теоремы существования РБС. Общая схема вывода теорем существования выглядит следующим образом. В разделе 2 кратко изложены теоремы, опубликованные в предыдущих работах автора. В разделе 3 приводятся две оригинальные теоремы из работы Рени. В разделе 4 дается подробная интерпретация условий теорем Рени в сравнении с условиями теоремы Дебре. В разделе 5 дается подробный анализ теоремы Рени. На ряде примеров условие теоремы интерпретируется как условие отсутствия точек перескока или точек, гарантирующих наилучший ответ. В разделе 6 формально, методом мета-теоремы, строятся два критерия существования для РБС, использующие исходные теоремы существования РН. В разделах 7 и 8 сформулированы и доказаны две теоремы, специально доработанные для решения прикладных задач (пространственная конкуренция Хотеллинга, конкуренция за ренту Таллока, олигополия Бертрана – Эджворта). Все рассмотренные теоремы сведены в итоговую таблицу.

Ключевые слова: равновесие в безопасных стратегиях, равновесие Нэша, теоремы существования решения игровых задач, теорема Рени, задача пространственной конкуренции Хотеллинга, задача конкуренции за ренту Таллока – Скапердаса.

¹ Михаил Борисович Искаков, к.т.н. (mih_isakov@mail.ru).

² Алексей Борисович Искаков, к.ф.м.н. (isk_alex@mail.ru).

1. Введение

Концепция РБС была предложена в статье [1, 2]. Поскольку модель равновесия была более широкой, чем РН, то первоначальное направление исследований сводилось к поиску задач, для которых оно работает. Сначала проверялось наличие РБС для отдельных задач.

Если рассматривать РН как базовую классическую модель решения в теории игр, то для очень многих задач оно не просто существует, но имеется большое множество РН, причем возникает проблема выбора из этого множества предпочтительных решений. Противоположный случай менее распространен, но имеется ряд классических задач, для которых РН не существует. Такой ряд наиболее известных и важных задач был рассмотрен в классической статье [9], и именно он был взят за основу для тестирования модели РБС.

В работе Дасгупты и Маскина рассматриваются задачи: пространственной конкуренции Хотеллинга [5, 13], дуополия Бертрана – Эджворта [6, 11], модель рынка страхования Ротшильда – Стиглица – Вильсона [15, 20]. Кроме них была также рассмотрена задача конкуренции за ренту Таллока – Скапердаса [17, 18, 19]. Оказалось, что весь рассмотренный ряд наиболее известных задач без РН имеет решение РБС, причем с хорошими свойствами, сравнительно с альтернативными подходами к их решению: равновесием Нэша в смешанных стратегиях, *ad hoc* решениями для частного случая конкретной задачи, стратегий наказания при переходе к динамической модели повторяющихся игр, коллективно оптимальных моделей решения (оптимальность по Парето, кооперативные, коалиционные игры и т.д.). Все РБС решения этого ряда задач представлены в итоговой статье [12].

Сравнение с альтернативами проведено в [4]. В качестве РБС-решений были найдены принципиально новые решения с убедительными интерпретациями; другие методы не давали таких решений. Сравнительно с решениями в смешанных стратегиях, достоинством РБС является простота и возможность получить решение в явном виде. Сравнительно с народными теоремами – единственность равновесий или их небольшое количество (одно-два). Сравнительно с *ad hoc* методами – универсальная

применимость к многим разным задачам. Сравнительно с методами решения в угрозах и контругрозах (которые кажутся наиболее близкими по идеологии к РБС) получались совершенно другие решения, которые проявили различие в подходе к возникающим игровым угрозам. Подход РБС стремится найти устойчивые профили не на основе стратегий наказания, а через введение дополнительных правил индивидуального поведения игроков, следующих принципу осторожности.

Поскольку для представительной выборки задач без РН были найдены РБС решения, то естественно возник вопрос, можно ли получить какие-то общие условия существования РБС, очертить круг задач, к которым применимы данные решения, т.е. построить теорию существования РБС. Данная статья является непосредственным продолжением статей [3, 4]. Статьи написаны по плану, сформулированному совместно с Клодом д'Апремоном и Алексеем Исаковым во время работы над статьей [12]. Автор выражает глубокую благодарность своим соавторам.

План теоретического обоснования равновесия в безопасных стратегиях как концепции решения игровых задач предполагал преодоление ряда проблем. В настоящий момент разрешены следующие из них.

1. В качестве теоретической основы метода сформулирован вывод определения равновесия в безопасных стратегиях (РБС) как развитие концепции теоретико-игрового равновесия по Нэшу (РН).

2. На этой основе разработан метод конструирования теорем существования РБС из известных теорем существования РН. При этом исходная теорема существования РН приводится к стандартной формулировке, которая вставляется в текст мета-теоремы существования РБС как условие.

3. Метод конструирования теорем существования опробован на трех исходных теоремах: Дебре (1952) [10], Рени (1999) [14], Бика (2009) [8].

4. Полученные теоремы существования РБС опробованы и доработаны на хрестоматийных задачах Хотеллинга [13, 5], Таллока – Скапердаса [17, 18, 19], Бертрана – Эджворта [6, 11], не имеющих РН при некоторых значениях параметров. При этом

если условия стандартной формулировки не выполняются для тестовых задач, слишком жесткие, то разрабатывается смягченный вариант теоремы, под условия которого попадают тестовые задачи.

5. Таким образом, в рамках общего метода сформулированы три рабочих теоремы существования РБС решений, применимые на тестовых эталонных задачах. Разрешение первого пункта плана опубликовано в [4], второго – в [3]. В первой из статей дается система определений РБС, выведенная из определения РН, а также обсуждается место концепции РБС среди других моделей ограниченной рациональности. Во второй статье на основе этой аналогии строится метод конструирования теорем существования РБС из известных теорем существования РН. Теорема существования РБС по теории существования социального равновесия Дебре с тестированием на указанных задачах опубликована в совместной работе [12]. В данной статье предложенный метод получения теорем существования РБС тестируется на теории существования РН Рени.

2. Необходимые исходные понятия, утверждения и теоремы

Приведем введенные определения и доказанные результаты из предыдущих статей [3, 4], используемые в этой. Система определений равновесия в безопасных стратегиях. Рассматривается некооперативная игра n участников $G = (i \in N, s_i \in S_i, u_i \in \mathbb{R})$ в нормальной форме.

Определение 2.1. Угрозой игрока j игроку i в профиле стратегий s называется такая пара профилей $\{s, (s'_j, s_{-j})\}$, что $u_j(s'_j, s_{-j}) > u_j(s)$ и $u_i(s'_j, s_{-j}) < u_i(s)$. Профиль стратегий s называется **содержащим угрозу**, а s' – **угрожающим** игроку i со стороны игрока j .

Определение 2.2. Стратегия s_i игрока i называется **безопасной стратегией** для этого игрока при заданных стратегиях s_{-i} всех других игроков, если профиль s не содержит угроз игроку i . Профиль стратегий s называется **безопасным профилем**, если все его стратегии безопасны.

Определение 2.3. **Безопасным отклонением** игрока i в профиле s называется такая стратегия s'_i , что $u_i(s'_i, s_{-i}) > u_i(s)$ и $u_i(s'_i, s'_j, s_{-ij}) \geq u_i(s)$ для любой угрозы $\{(s'_i, s_{-i}), (s'_i, s'_j, s_{-ij})\}$ любого игрока $j \neq i$ игроку i .

Определение 2.4. Безопасный профиль стратегий называется **равновесием в безопасных стратегиях (РБС)**, если ни один игрок не может увеличить свой выигрыш безопасным отклонением.

Первое важное свойство РБС вытекает непосредственно из его определения.

Утверждение 2.1. Любое равновесие Нэша является равновесием в безопасных стратегиях.

Доказательство. В равновесии Нэша нет угроз, поэтому он является безопасным профилем. Никакой игрок в равновесии Нэша не может увеличить свой выигрыш никаким отклонением. Оба условия РБС выполнены. $\square$

Чтобы охарактеризовать содержащиеся в игре угрозы количественно, можно поставить в соответствие исходной игре такую, в которой функции выигрышей учитывают ожидание наихудшей имеющейся в профиле угрозы. Обозначим как $Q_i(s_{-i})$ множество безопасных стратегий игрока i при заданных стратегиях s_{-i} других игроков. Заметим, что $Q_i(s_{-i})$ может оказаться пустым, если все стратегии игрока i небезопасны при s_{-i} . Обозначим как $Q^{(i)}$ множество безопасных профилей игрока в игре.

Определение 2.5. **Безопасный выигрыш** игрока i в профиле стратегий s равен его выигрышу при реализации наихудшей угрозы в небезопасном профиле и совпадает с его выигрышем в безопасном профиле, т.е.

$$v_i(s) = \begin{cases} \inf_{i \neq j, s'_i: u_j(s'_j, s_{-j}) > u_j(s)} u_i(s'_j, s_{-j}), & s_i \notin Q_i(s_{-i}), \\ u_i(s), & s_i \in Q_i(s_{-i}). \end{cases}$$

Соответствующей игре G **игрой угроз** называется игра $\tilde{G} = (S_i, v_i)_{i=1}^N$.

Введённая функция $v_i(s)$ позволяет охарактеризовать безопасность профиля и безопасность отклонений простым способом. Будем рассматривать только такие отклонения $s \xrightarrow{i} s'$, для

которых $u_i(s') \geq u_i(s)$ (будем называть их выгодными отклонениями).

Утверждение 2.2. Профиль s безопасен для игрока $i \Leftrightarrow u_i(s) = v_i(s)$.

Профиль s небезопасен для игрока $i \Leftrightarrow u_i(s) > v_i(s)$.

Выгодное отклонение $s \xrightarrow{i} s'$ безопасно для игрока $i \Leftrightarrow v_i(s') \geq u_i(s)$.

Выгодное отклонение $s \xrightarrow{i} s'$ небезопасно для игрока $i \Leftrightarrow v_i(s') < u_i(s)$.

Утверждение 2.3. Если безопасный профиль s^* является строгим равновесием Нэша в игре угроз $\tilde{G}$, то он является РБС в исходной игре. Если s^* является РБС в исходной игре G , то s^* является он является равновесием Нэша в игре угроз $\tilde{G}$.

Слабое равновесие в безопасных стратегиях.

Определение 2.6. Безопасное отклонение s'_i игрока i в профиле s называется **тривиальным**, если существует такая угроза $\{(s'_i, s_{-i}), (s'_i, s'_j, s_{-ij})\}$ игрока $j \neq i$ игроку i , что $u_i(s'_i, s'_j, s_{-ij}) = u_i(s)$.

Определение 2.7. Безопасный профиль стратегий называется **слабым равновесием в безопасных стратегиях**, если ни один игрок не может сделать не тривиальное безопасное отклонение.

Утверждение 2.3 ставит в соответствие РБС и равновесие Нэша. Теорема утверждает более сильный результат, как необходимое и достаточное условие.

Теорема 2.1. Профиль s^* является равновесием в безопасных стратегиях тогда и только тогда, когда выполняются три условия:

- 1) s^* – безопасный профиль;
- 2) s^* – равновесие Нэша в игре угроз $\tilde{G}$;
- 3) $\forall i, \forall s'_i: v_i(s'_i, s_{-i}^*) = v_i(s^*) \Rightarrow u_i(s'_i, s_{-i}^*) = u_i(s^*)$.

Доказательство приведено в [3] как «теорема 1». □

Следствие. Профиль в игре G является слабым равновесием в безопасных стратегиях тогда и только тогда, когда он безопасен в игре G и является равновесием Нэша в игре угроз $\tilde{G}$.

Метатеорема. Теоремы существования РБС строятся, как конструктор, из известных теорем существования РН, приведенных к стандартному виду. В литературе формулировки этих теорем различны. Требуется эквивалентно переформулировать их так, чтобы они различались только определенным условием, которое будет переменной частью в метатеореме.

Рассмотрим произвольную игру в нормальной форме: $G = (N, S_i, u_i(s_1, \dots, s_n), i \in N, s_i \in S_i)$. Обозначим как « $(\#)$ » некоторое ограничивающее условие, однозначно определяющее подмножество множества всех возможных игр G . Это условие и будет переменной частью в метатеореме.

Пример 2.1. Определим функцию наилучшего ответа: $BR_i(s_{-i}) = \arg \max_{s_i \in S_i} u_i(s_i, s_{-i})$. Пусть условие задается так:

$(\#1) := \langle \forall i, \forall s_{-i}, \exists BR_i(s_{-i}) \rangle$, функции $BR_i(s_{-i})$ непрерывны по s_{-i} . Это условие является достаточным условием для существования РН в игре.

Метод конструирования теорем существования РБС основан на двух идеях. Первой составляющей является понятие сильной угрозы.

Определение 2.8. Угроза игроку i в профиле s является **сильной**, если существует безопасная стратегия $s' = (s'_i, s_{-i})$ такая, что $u_i(s') = v_i(s') > v_i(s)$. Для игрока i в игре G выполняется **условие сильных угроз**, если для него содержащиеся в любом опасном профиле угрозы являются сильными. Игра G называется игрой **с сильными угрозами**, если для всех игроков в ней выполняется условие сильных угроз.

Вторая составляющая идеи метатеоремы существования РБС состоит в том, что если для игры выполняется требование сильных угроз и имеется некоторая известная теорема существования равновесия Нэша, то можно потребовать выполнения условий этой теоремы только на безопасном множестве (или даже на некотором предпочтительном подмножестве этого множества, содержащего в себе наилучшую безопасную альтернативу – достигаемое максимальное значение безопасного выигрыша), и этого будет достаточно для существования РБС.

Пусть имеется некоторое верное утверждение (исходная теорема, приведенная к стандартному виду): «Если для игры выполняется условие (#), то в игре существует равновесие Нэша». Пусть это условие (#) выполняется на множествах безопасных стратегий игроков. Такое предположение надо сформулировать строго. Введем константу $C_{\min} \leq u_i(s), \forall i, s$, которая обозначает некоторое штрафное значение выигрыша.

Определение 2.9. Пусть дана игра $G(S_i, u_i)_{i=1}^N$ с множествами безопасности $Q^{(i)} \subseteq S$. Игра $\bar{G}_{Q^{(i)}}(S_i, \bar{u}_i), \bar{u}_i(s) = \begin{cases} u_i(s), s \in Q^{(i)}, \\ C_{\min}, s \notin Q^{(i)} \end{cases}$ называется **соответствующей ей обрезанной игрой**. Условие (#) существования равновесия Нэша **выполняется для игры G на безопасных множествах $Q^{(i)} \subseteq S$** , если оно выполняется для соответствующей обрезанной игры.

Таким образом, если для игры выполняются два условия: *условие теоремы существования равновесия Нэша на безопасном множестве* и *условие сильных угроз*, то можно ожидать, что в данной игре имеется РБС. Условие исходной теоремы (#) обеспечивает наличие равновесия в нужном множестве, а условие сильных угроз гарантирует его устойчивость в смысле РБС для всей игры. Метатеорема:

Теорема 2.2. Пусть верно утверждение: «Если для игры выполняется условие (#), то в игре существует равновесие Нэша». Если для игры $G(S_i, u_i)_{i=1}^N$ выполняется условие сильных угроз, а на её безопасных множествах $Q^{(i)} \subseteq S$ выполняется условие (#) существования равновесия Нэша, тогда в игре G существует равновесие в безопасных стратегиях.

Доказательство приведено в [3] как «теорема 2». □

Для рассмотренного выше примера 2.1 полученная при помощи метатеоремы 2.2 теорема существования РБС будет формулироваться так: «Если для игры $G(S_i, u_i)_{i=1}^N$ выполняется условие сильных угроз, а на её безопасных множествах $Q^{(i)} \subseteq S$ выполняется условие $\forall i \forall s_{-i} \exists! BR_i(s_{-i})$, функции $BR_i(s_{-i})$ непрерывны по s_{-i} , тогда в игре G существует равновесие в безопасных стратегиях».

Локальный вариант теоремы. Утверждение теоремы носит общий характер и содержательно слабо, так как требуемые в ней условия достаточно сильны. Ослабление условия *сильных угроз* заключается в том, что его выполнение требуется только по отношению к некоторому множеству $B = \times_{i=1}^N B_i$, где множества B_i предполагаются компактными выпуклыми подмножествами S_i .

Определение 2.10. Для игрока i выполняется **условие сильных угроз в B** , если для каждого $s_{-i} \in B_{-i}$ существует непустое подмножество $\tilde{Q}_i(s_{-i}) \subseteq Q_i(s_{-i}) \cap B_i$ такое, что для каждой стратегии $s_i \notin \tilde{Q}^{(i)}$ существует стратегия $s'_i \in \tilde{Q}^{(i)}$ такая, что $u_i(s'_i, s_{-i}) = v_i(s'_i, s_{-i}) > v_i(s_i, s_{-i})$. Игра G называется игрой **с сильными угрозами по отношению к B** , если для всех игроков выполняется условие сильных угроз в B .

Теорема 2.3. Пусть верно утверждение: «Если для игры выполняется условие (#), то в игре существует равновесие Нэша». Если игра $G(S_i, u_i)_{i=1}^N$ является игрой с сильными угрозами по отношению к B , а на её безопасных множествах $Q^{(i)} \subset S$ выполняется условие (#) существования равновесия Нэша, тогда в игре G существует равновесие в безопасных стратегиях.

Доказательство приведено в [3] как «теорема 3». $\square$

3. Теорема Рени (1999)

Статья Рени [14] предлагает результат существования равновесий по Нэшу в чистых стратегиях для большого класса разрывных игр. Покажем подробно на ее примере, как работает метатеорема. Результаты Рени основаны на условии, называемом *гарантией лучшего ответа*. Игра гарантирует лучший ответ, если для каждой неравновесной стратегии s^* и каждого предела вектора выигрыша u^* , возникающего при приближении стратегий игроков к s^* , у некоторого игрока i имеется стратегия, приносящая выигрыш строго выше u_i , даже если другие игроки слегка отклонятся от s^* . Главный результат заключается в том, что игры с компактными выпуклыми пространствами стратегий и с выигрышами, которые являются квазивогнутыми по собственным

стратегиям, обладают равновесием по Нэшу в чистых стратегиях, если при этом они гарантируют лучший ответ.

Гарантия лучшего ответа объединяет и обобщает два условия: *взаимную верхнюю полунепрерывность* и *гарантию выигрыша*. Игра взаимно верхнеполунепрерывна, если всякий раз, когда выигрыш какого-либо игрока прыгает вниз, выигрыш какого-то другого игрока прыгает вверх. Игра гарантирует выигрыш, если для каждого вектора стратегий s у каждого игрока есть стратегия, которая фактически гарантирует ему выигрыш, который он получает в s , даже если другие игроки сыграют немного иначе, чем в s . Оба условия выполняются во многих экономических играх, и их часто довольно просто проверить.

Формально эти понятия задаются следующим образом. Имеется N игроков. Каждый игрок $i = 1, 2, \dots, N$ имеет множество чистых стратегий S_i – непустое компактное подмножество топологического векторного пространства, а также ограниченную функцию $u_i: S \rightarrow \mathbb{R}$, где $S = \times_{i=1}^N S_i$. При этих условиях игра $G = (S_i, u_i)_{i=1}^N$ называется *компактной игрой*.

Вектор выигрышей игроков обозначается как $u: S \rightarrow \mathbb{R}^N$ и определяется как $u(s) = (u_1(s), \dots, u_N(s))$ для каждого $s \in S$. Графиком вектора выигрышей является подмножество множества $S \times \mathbb{R}^N$, определяемое как $\{(s, u) \in S \times \mathbb{R}^N \mid u = u(s)\}$. Наконец, если каждое S_i выпукло и для каждого i и каждого $s_{-i} \in S_{-i}$, $u_i(\cdot, s_{-i})$ квазивогнута на S_i , то определяется, что $G = (S_i, u_i)_{i=1}^N$ *квазивогнута*.

Определение 3.1. (Рени-1). Игрок i **может гарантировать выигрыш** $\alpha \in \mathbb{R}$ при $s \in S$, если существует $s'_i \in S_i$, такое что $u_i(s'_i, s'_{-i}) \geq \alpha$ для всех s'_{-i} в некоторой открытой окрестности s_{-i} .

Таким образом, выигрыш может быть гарантирован игроком i в профиле s , если игрок i имеет стратегию, которая гарантирует по крайней мере этот выигрыш, даже если другие игроки немного отклоняются от s .

Определим множество *графика вектора выигрышей игры* как $\Gamma(G) = \{(s, u(s)), s \in S\}$. Дальнейшее определение использует понятие *замыкания графика вектора выигрышей игры*. В контексте метрического пространства под этим понимается

следующее: пара $(s^*, u^*) \in \overline{\Gamma(G)}$ находится в замыкании графика вектора выигрышей, если предельные значения вектора выигрышей u^* являются пределом последовательности векторов выигрышей $(s_i^k, u_i(s_i^k)) \xrightarrow{k \rightarrow \infty} (s^*, u^*)$, некоторой последовательности стратегий, сходящейся к s^* . При этом может быть $u(s^*) \neq u^*$.

Определение 3.2. (Рени-2). Игра $G = (S_i, u_i)_{i=1}^N$ **гарантирует лучший ответ** (better-reply secure – BRS), если всякий раз, когда (s^*, u^*) находится в замыкании графика её вектора выигрышей и s^* не является равновесием, некоторый игрок i может гарантировать выигрыш строго больше u_i^* в профиле s^* .

Таким образом, игра гарантирует лучший ответ, если для любого неравновесного профиля стратегий s^* и любого вектора выигрышей u^* , возникающего как предел при последовательности стратегий, приближающихся к s^* , некоторый игрок i имеет стратегию, дающую выигрыш строго больше u_i^* , даже если другие игроки немного отклоняются от s^* .

Теперь можно сформулировать основную теорему Рени.

Теорема 3.1 (Рени, т. 3.1 [14]). Если игра $G = (S_i, u_i)_{i=1}^N$ компактна, квазивогнута и гарантирует лучший ответ, то она обладает равновесием Нэша в чистых стратегиях.

Соответствующее условие (#2):= «игра $G = (S_i, u_i)_{i=1}^N$ компактна, квазивогнута и гарантирует лучший ответ».

В то время как гарантия лучшего ответа проверяется непосредственно, иногда проще проверить другие приводящие к ней условия. Рени предложил также два довольно полезных условия, которые в совокупности подразумевают гарантию лучшего ответа.

Определение 3.3. (Рени-3). Игра $G = (S_i, u_i)_{i=1}^N$ **гарантирует выигрыш** (payoff secure – PS), если для каждого $s \in S$ и каждого $\varepsilon > 0$ каждый игрок i может гарантировать выигрыш $u_i(s) - \varepsilon$ в (профиле стратегий) s .

Гарантия выигрыша требует, чтобы для каждого профиля стратегий s каждый игрок имел (такую) стратегию, которая фактически гарантирует ему выигрыш, который он получает в точке s , даже если остальные игроки слегка отклонятся от s .

Определение 3.4. (Рени-4). Игра $G = (S_i, u_i)_{i=1}^N$ **взаимно верхнеполунепрерывна (в.в.п.н.)** (reciprocally upper semicontinuous – RUSC), если всякий раз, когда (s, u) находится в замыкании графа её вектора выигрышей и $u_i(s) \leq u_i$ для каждого игрока i , тогда $u_i(s) = u_i$ для каждого игрока i .

Взаимная в.п.н. была впервые предложена под названием свойства *дополняющих разрывностей*. Она означает, что выигрыш некоторого игрока должен прыгать вверх всякий раз, когда выигрыш некоторого другого игрока прыгает вниз.

Утверждение 3.1. (Рени). Если игра $G = (S_i, u_i)_{i=1}^N$ взаимно верхнеполунепрерывна и гарантирует выигрыш, то она гарантирует лучший ответ.

Следствие 3.1 (Рени, сл. 3.3 [14]). Если игра $G = (S_i, u_i)_{i=1}^N$ компактна, квазивогнута, взаимно верхнеполунепрерывна и гарантирует выигрыш, то в ней имеется равновесие Нэша в чистых стратегиях.

Соответствующее условие (#3):= «игра $G = (S_i, u_i)_{i=1}^N$ компактна, квазивогнута, взаимно верхнеполунепрерывна и гарантирует выигрыш»

Результаты Рени были развиты в статье [5], которые также использованы при разработке теоремы существования РБС, так что здесь их также следует привести.

Определение 3.5. (Баг, Джофр). Игра $G = (S_i, u_i)_{i=1}^N$ **слабо взаимно верхнеполунепрерывна (с.в.в.п.н.)** (weakly reciprocally upper semicontinuous – WRUSC), если для любого $(s^*, u^*) \in \overline{\Gamma(G)} \setminus \Gamma(G)$ существует игрок i и $s'_i \in S_i$, такие что $u_i(s'_i, s_{-i}^*) > u_i^*$.

Утверждение 3.2. (Баг, Джофр, утв. 1 [7]). Если игра $G = (S_i, u_i)_{i=1}^N$ слабо взаимно верхнеполунепрерывна и гарантирует выигрыш, то она гарантирует лучший ответ.

Соответствующее условие (#4):= «игра $G = (S_i, u_i)_{i=1}^N$ компактна, квазивогнута, слабо взаимно верхнеполунепрерывна и гарантирует выигрыш»

Таким образом, для существования равновесия Нэша в игре достаточно компактности, квазивогнутости, свойств гарантии выигрыша и с.в.в.п.н.

4. Анализ и интерпретация условий теоремы Рени. Сравнение с теоремой Дебре

Чтобы продемонстрировать соотношения между разными теоремами существования, рассмотрим несколько примеров. Хотя теорема Рени более мощная, но следствие Рени оперирует более понятными и легко проверяемыми условиями, поэтому именно следствие, как правило, становится основой для дальнейших исследований. То, что применимость теоремы Рени существенно шире, чем её следствия, иллюстрирует следующий пример.

Пример 4.1. Игра $G = (S_i, u_i)_{i=1}^N$, где $N = 2$, $S_i = [0, 1]$,

$$u_i(s_i, s_{-i}) = \begin{cases} 1 - \left(s_i - \frac{1+s_{-i}}{2}\right)^2, & s_i \in \left[s_{-i}, \frac{3+s_{-i}}{4}\right], \\ 0, & s_i \notin \left[s_{-i}, \frac{3+s_{-i}}{4}\right]. \end{cases}$$

Игра имеет единственное равновесие Нэша $(s_1^*, s_2^*) = (1, 1)$. Она компактна, квазивогнута, верхнеполунепрерывна. Игра гарантирует лучший ответ, так как для $\forall (s_i, s_{-i}) \neq (1, 1)$ либо один, либо другой игрок могут отклониться в точку максимума своей целевой функции $s_i = \frac{1+s_{-i}}{2}$ так, что малое отклонение игрока $(-i)$ не выведет исход игры в область, где $u_i(s_i, s_{-i}) = 0$, т.е. условие определения Рени-2 выполняется. Но игра не гарантирует выигрыш, так как для точек $s_i < 1$, $s_{-i} = 1$ игрок i может отклониться в единственную точку с ненулевым выигрышем $(s_i^*, s_{-i}^*) = (1, 1)$, а любое малое отклонение игрока $(-i)$ выведет исход игры в область, где $u_i(s_i, s_{-i}) = 0$, т.е. условие определения Рени-3 гарантированного выигрыша не выполняется. То есть для этой игры выполняются условия теоремы Рени, где ограничение накладывается на хотя бы одного из двух игроков, но ни следствие Рени, ни утверждение Бага – Жофра с ограничением на обоих игроков не могут быть применены.

В задаче Хотеллинга возникает именно такая ситуация, что заставило сформулировать более сложные варианты теорем в последующих разделах 7 и 8.

Сравнение теорем Дебре и Рени.

Теорема Дебре (1952) формулируется следующим образом [10].

Полигедрон – множество в $\mathbb{R}^n$ гомеоморфное геометрическому полигедрону (т.е. объединению конечного числа выпуклых оболочек в $\mathbb{R}^n$). Он очевидно замкнут.

Пусть при заданном s_{-i} (т.е. действиях всех остальных) выбор i -го агента ограничен *непустым компактным* множеством $A_i(s_{-i}) \subset S_i$. Агент i выбирает s_i из $A_i(s_{-i})$ так, чтобы максимизировать $u_i(s_{-i}, s_i)$, которая предполагается непрерывной по s_i в $A_i(s_{-i})$. Множества $A_i(s_{-i})$ интерпретируются как задающие совокупность социально приемлемых выборов. Это делает интуитивным следующее определение *социального равновесия по Дебре*.

Определение (Дебре). s^* является **точкой социального равновесия**, если для всех $i = 1, \dots, n$: $s_i^* \in A_i(s_{-i}^*)$ и $u_i(s^*) = \max_{s_i \in A_i(s_{-i}^*)} u_i(s_{-i}^*, s_i)$.

График функции $A_i(s_{-i})$ определяется как подмножество $S_{-i} \times S_i$ следующего вида: $\Gamma_i = \{(s_{-i}, s_i) \mid s_i \in A_i(s_{-i})\}$. Для любого s_{-i} , множество $A_i(s_{-i})$ всегда предполагается непустым.

Теорема 4.1. (Дебре, т.1 [10]). Пусть для всех $i = 1, \dots, n$ множества S_i – стягиваемые полиэдры, $A_i(s_{-i})$ – многозначные функции из S_{-i} в S_i с замкнутыми графиками Γ_i , u_i – непрерывные функции из Γ_i в дополненную ось действительных чисел такие, что $\varphi_i(s_{-i}) = \max_{s_i \in A_i(s_{-i})} u_i(s_{-i}, s_i)$ непрерывна. Если для каждого i и s_{-i} множество $M_{s_{-i}} = \{s_i \in A_i(s_{-i}) \mid u_i(s_{-i}, s_i) = \varphi_i(s_{-i})\}$ стягиваемо, то существует точка социального равновесия.

Если взять в качестве множества социально приемлемых выборов все множество стратегий $A_i(s_{-i}) = S_i$, то из теоремы Дебре получается теорема существования РН.

Теорема 4.2. Пусть для всех $i = 1, \dots, n$ множества S_i – стягиваемые полиэдры, графики $\Gamma_i^S = \{(s_{-i}, s_i) \mid s_i \in S_i\}$ замкнуты, u_i – непрерывные функции из Γ_i в дополненную ось дей-

ствительных чисел такие, что $\varphi_i(s_{-i}) = \max_{s_i \in S_i} u_i(s_{-i}, s_i)$ непрерывна. Если для каждого i и s_{-i} множество $M_{s_{-i}} = \{s_i \in S_i \mid u_i(s_{-i}, s_i) = \varphi_i(s_{-i})\}$ стягиваемо, то существует точка РН.

Если взять в качестве множества социально приемлемых выборов множество безопасных стратегий стратегий $A_i(s_{-i}) = Q_i(s_{-i})$, то из теоремы Дебре получается теорема существования РБС. Соответствующее условие (#5): «в игре $G = (S_i, u_i)_{i=1}^N$ для всех $i = 1, \dots, n$ множества S_i – стягиваемые полиэдры, $Q_i(s_{-i})$ – многозначные функции из S_{-i} в S_i с замкнутыми графиками Γ_i , u_i – непрерывные функции из Γ_i в дополненную ось действительных чисел такие, что $\varphi_i(s_{-i}) = \max_{s_i \in Q_i(s_{-i})} u_i(s_{-i}, s_i)$ непрерывна. Если для каждого i и s_{-i} множество $M_{s_{-i}} = \{s_i \in Q_i(s_{-i}) \mid u_i(s_{-i}, s_i) = \varphi_i(s_{-i})\}$ стягиваемо»

Теорема 4.3. Пусть для всех $i = 1, \dots, n$ множества S_i – стягиваемые полиэдры, $Q_i(s_{-i})$ – многозначные функции из S_{-i} в S_i с замкнутыми графиками Γ_i , u_i – непрерывные функции из Γ_i в дополненную ось действительных чисел такие, что $\varphi_i(s_{-i}) = \max_{s_i \in Q_i(s_{-i})} u_i(s_{-i}, s_i)$ непрерывна. Если для каждого i и s_{-i} множество $M_{s_{-i}} = \{s_i \in Q_i(s_{-i}) \mid u_i(s_{-i}, s_i) = \varphi_i(s_{-i})\}$ стягиваемо, то существует точка РБС.

В качестве частного, более простого, вытекающего из теоремы Дебре утверждения можно взять пример 2.1: если функции наилучшего ответа игроков непрерывны, то они пересекутся и в этой точке будет существовать равновесие Нэша или точка социального равновесия Дебре.

Доказанные Дебре и Рени результаты пересекаются, но не совпадают и не вкладываются друг в друга. Приведём два примера игр, для которых действует одна теорема и не действует другая.

Пример 4.2. Игра $G = (S_i, u_i)_{i=1}^N$, где $N = 2$, $S_i = [0, 1]$,

$$u_i(s_i, s_{-i}) = \begin{cases} 1, & s_i = s_{-i}, \\ 0, & s_i \neq s_{-i}. \end{cases}$$

Здесь выполняются условия теоремы Дебре, функции наилучшего ответа игроков непрерывны. Условие гарантированного выигрыша Рени не выполняется, поскольку максимумы целевых функций достигаются на разрывах функций выигрыша, в выколотых вверх точках.

Пример 4.3. Игра $G = (S_i, u_i)_{i=1}^N$, где $N = 2$, $S_i = [0, 1]$,

$$u_1(s_1, s_2) = \begin{cases} 1 - \left(\frac{3}{4} - s_1\right)^2, & s_2 \leq \frac{1}{3}, \\ 1 - \left(\frac{1}{4} - s_1\right)^2, & s_2 > \frac{1}{3}; \end{cases} \quad u_2(s_1, s_2) = 1 - \left(\frac{1}{2} - s_2\right)^2.$$

Здесь условия теоремы Дебре не выполняются, непрерывности наилучшего ответа нет. Все условия теоремы Рене выполнены, в том числе и свойство гарантированного выигрыша.

Тем не менее следует прояснить, **насколько подходы Дебре и Рени пересекаются**, степень их общности. Для этого поясним на примере и в последующем теоретическом обобщении, как принцип непрерывности функции наилучшего ответа проявляется в рамках подхода Рени.

Пример 4.4. Рассмотрим две игры $G = (S_i, u_i)_{i=1}^N$, где $N = 2$, $S_i = [0, 1]$. Вариант А:

$$u_1(s_1, s_2) = \begin{cases} 1 - \left(\frac{3}{4} - s_1\right)^2, & s_2 \leq \frac{1}{2}, \\ 1 - \left(\frac{1}{4} - s_1\right)^2, & s_2 > \frac{1}{2}; \end{cases}$$

$$u_2(s_1, s_2) = \begin{cases} 1 - \left(\frac{1}{4} - s_2\right)^2, & s_2 \leq \frac{1}{2}, \\ 1 - \left(\frac{3}{4} - s_2\right)^2, & s_1 > \frac{1}{2}. \end{cases}$$

Вариант Б:

$$u_1(s_1, s_2) = \begin{cases} 1 - \left(\frac{3}{4} - s_1\right)^2, & s_2 \leq \frac{1}{2}, \\ 1 - \left(\frac{1}{4} - s_1\right)^2, & s_2 > \frac{1}{2}; \end{cases} \quad u_2(s_1, s_2) = 1 - (s_1 - s_2)^2.$$

В этих играх нет равновесных по Нэшу точек, нет непрерывности функций наилучшего ответа (см. рис. 1), не выполняется свойство лучшего гарантированного ответа. Если рассмотреть более подробно, то условие лучшего гарантированного ответа не

выполняется в обеих играх только на одной точке: $(s_1, s_2) = (0,5; 0,5)$. Именно в этой точке в обоих случаях происходит «перескок одной через другую» функций наилучшего ответа. То есть если сравнить примеры 4.3 и 4.4, получается, что условие лучшего гарантированного ответа Рени «ловит» не все разрывы данных функций, а только те, в которых функции пере-скакивают друг через друга, создавая предпосылку для отсутствия решения игры по Нэшу. Вывод из этого наблюдения надо сформулировать строго.

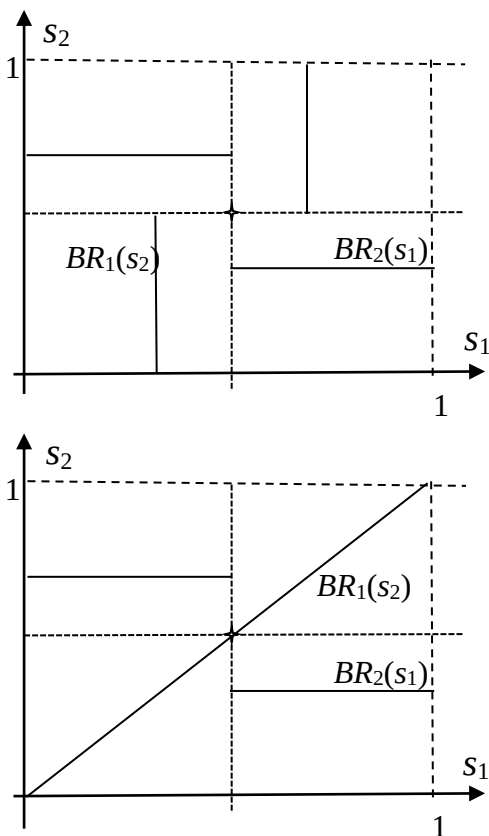


Рис. 1. Примеры 4.3, 4.4, функции наилучшего ответа и точки перескока

5. Точки перескока

Точки перескока. Интересно было бы узнать ответ на вопрос, насколько общими являются условия утверждений из п. 3. Чтобы получить его, рассмотрим более подробно условие гарантированного лучшего ответа Рени в примерах 4.3 и 4.4, обращая внимание прежде всего на точки, где оно нарушается. Если записать отрицание определения Рени-2 гарантирования лучшего ответа, привязанное к профилю s^* , то получится следующее.

Определение 4.1. Профиль s^* , не являющийся равновесием Нэша, называется **негарантирующим лучший ответ** в игре $G(S_i, u_i)_{i=1}^N$, если $\exists (s^*, u^*) \in \overline{G(G)}: \forall i, \forall \alpha > u_i^*, \forall s'_i \in S_i, \forall$ окрестности $U(s_{-i}^*) \subset S_{-i}$ (т.е. открытого множества, содержащего s_{-i}^*), $\exists s'_{-i} \in U(s'_{-i}, s_{-i}^*)$ такое, что $u_i(s'_i, s'_{-i}) < \alpha$.

В примере 4.3 второй игрок не может гарантировать лучший ответ во всех профилях s , где $s_2 = \frac{1}{2}$, при этом его стратегии являются лучшими ответами. Первый игрок не может гарантировать лучший ответ на множестве профилей: $(s_2 \leq \frac{1}{3}, s_1 = \frac{3}{4}) \cup (s_2 > \frac{1}{3}, s_1 = \frac{1}{4}) \cup (\frac{1}{4} \leq s_1 < \frac{3}{4}, s_2 = \frac{1}{3})$. При этом на первых двух подмножествах стратегия первого игрока является наилучшим ответом. В профилях третьего подмножества в точке разрыва своей функции лучшего ответа, по стратегии партнёра, стратегия первого игрока лежит между предельными значениями этой функции справа и слева. Единственная точка $(s_1, s_2) = (0,25; 0,5)$, где ни один игрок не может гарантировать лучший ответ, является равновесием Нэша.

В примере 4.4, рассуждая аналогично, можно увидеть, что единственной точкой, где оба игрока не могут гарантировать лучший ответ, есть точка $(s_1, s_2) = (0,5; 0,5)$. При этом в случае А стратегии обоих игроков лежат между предельными значениями своих функций лучшего ответа, в точке разрыва. А в случае Б это верно только для одного первого игрока, а второй игрок выбирает стратегию наилучшего ответа.

Следует заметить, что нарушения условия гарантирования лучшего ответа не всегда связано с точкой перескока. Это может

получиться, например, из-за того, что лучший ответ является выколотой вверх точкой разрыва функции выигрыша, что демонстрирует пример 4.2.

Для игр двух участников с одномерными стратегиями и кусочно-непрерывными функциями лучшего ответа возможны только два случая перескока, показанные на примере: перескок одной функции через непрерывную другую и встречный перескок двух функций (см. рис. 2). Только при наличии таких точек, в которых нарушается условие лучшего гарантированного ответа Рени, в игре возможно отсутствие равновесия Нэша. Соответствующие точки однозначно определяются либо как точка разрыва функции одного игрока и значение функции другого, либо как точки разрыва функции обоих игроков.

Подводя итог, можно сказать, что теорема Рени развивает идею теоремы Дебре, хотя на первый взгляд это не очевидно. Она является «в целом» более тонким и мощным необходимым условием существования равновесия Нэша, хотя далеко не достаточным. Слова «в целом» означают, что формально условия теоремы Рени не вкладываются полностью в условия теоремы Дебре, что было показано выше на примерах. Для того чтобы поймать одну и ту же содержательную ситуацию, используются различные формальные инструменты. Поэтому и строго вложения полученных результатов нет.

Такое далёкое отступление от РБС к теме сравнения теорем Рени и Дебре и перескоков функций наилучшего ответа друг через друга потребовалось, чтобы прояснить далеко не очевидный содержательный смысл сложных конструкций теоремы Рени и некоторые их тонкости, которые окажутся полезными на следующем этапе конструирования теоремы существования РБС на основе теоремы Рени и приведения всех требуемых этой теоремой условий. Следует заметить, что теорема Рени требует аккуратной работы с предельными точками, а это требует серьёзной доработки условий, сформулированных в теоремах 2.1, 2.2 и 2.3.

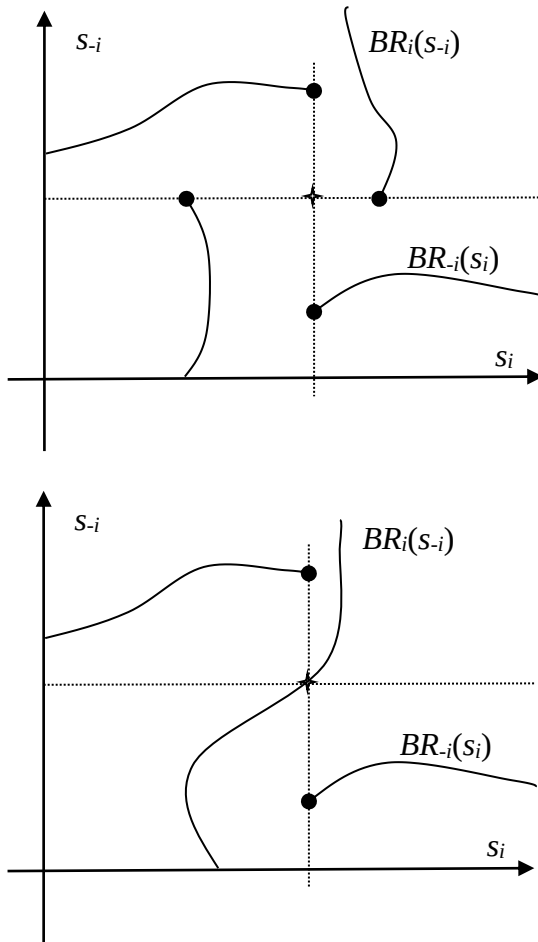


Рис. 2. Функции лучшего ответа и точки перескока для двух игроков

6. Теоремы существования РБС по Рени

Воспользуемся теоремой 2.3, подставив в неё в качестве исходных утверждения, сформулированные Рени. В качестве базовых формулировок можно взять собственно основную теорему

Рени и следствие (из утверждения) Рени. Поскольку второй вариант имеет условия, которые, хотя и более слабы по существу, но проверяются легче, то на практике, в публикациях развивающих подход Рени, как правило используется именно следствие Рени. Оба варианта требуют ввести уточняющие формулировки соответствующих определений.

Конкретизируем условие (#2) из теорем 2.2, 2.3 для случая теоремы Рени. Игра должна быть *компактна, квазивогнута и гарантировать лучший ответ* на безопасных множествах. Компактная игра, естественно, будет компактной и на безопасных множествах.

Определение 6.1. Игра $G(S_i, u_i)_{i=1}^N$ **квазивогнута на безопасных стратегиях**, если для любого игрока i при любом окружении s_{-i} функции $u_i(s)$ квазивогнуты по s_i при $s_i \in Q_i(s_{-i})$.

Это определение неявно предполагает, что множества безопасных стратегий $Q_i(s_{-i}) \subseteq S_i$ являются выпуклыми, так как если $Q_i(s_{-i})$ не выпуклы, то функции $u_i(s)$ не могут быть определены как квазивогнутые.

Чтобы применить условие *гарантирования лучшего ответа* к обрезанной функции в теореме 2.3 (определение 2.9), удобно разделить его на два случая: безопасных и небезопасных для игрока профилей. Для этого сначала следует ввести определение понятия, исходно данного для игры в целом, привязанное к игроку, его конкретному окружению (ситуации игры для него), либо даже к отдельному профилю.

Определение 6.2. Игрок i **имеет гарантированный безопасный ответ при окружении** s_{-i} , если $\exists s'_i \in S_i, \exists$ открытая окрестность $U_{s_{-i}} \subseteq S_{-i}, \forall s'_{-i} \in U_{s_{-i}}: s' = (s'_i, s'_{-i}) \in Q^{(i)}$. Игрок i **имеет гарантированный безопасный ответ в игре** G , если он его имеет при любых окружениях.

Это определение требует подробного комментария. На первый взгляд, оно искусственно переусложнено, но это необходимо. В теореме Рени требуется очень тонкая работа со всеми недостижимыми предельными точками $(s^*, u^*) \in \overline{\Gamma(G)} \setminus \Gamma(G)$. При этом небольшие различия в обозначениях множеств безопасных стратегий при введении РБС, при взятии их замыканий обретают различное значение. Напомним, что множество $Q^{(j)} \subseteq S$ было

определено при введении *обрезанной игры* (определение 3.2) как множество профилей безопасных для игрока j в игре, т.е. $Q^{(j)} = \{s: s_j \in Q_j(s_{-j}) \subseteq S_j\}$. Хотя сами обозначаемые множества профилей идентичны, но замыкание множества $\overline{Q^{(j)}} \subseteq S$ и совокупность замыканий множеств $\overline{Q_j(s_{-j})} \subseteq S_j$ образуют различные множества. Таким образом, « s^* – безопасные профили» является более узким множеством, чем « $s_j^* \in \overline{Q_j(s_{-j})} \subseteq S_j$ », которое содержит предельные точки множества при его замыкании в измерении S_j , а это множество, в свою очередь, является более узким, чем множество $s^* \in \overline{Q^{(j)}} \subseteq S$, где замыкание берётся уже на множестве S и где содержатся также предельные точки и по измерениям S_{-j} . Теорема Рени требует рассмотрения именно последнего, самого широкого множества предельных точек $(s^*, u^*) \in \overline{\Gamma(G)}$.

Поскольку для случая небезопасных профилей дополнительные ограничения накладываются на множество безопасных для игрока профилей, куда происходит безопасный переход игрока, то имеет смысл соединить в отдельном термине требования условия сильных угроз и требования, вытекающие из подхода Рени. Определение формулируется для всех частных вариантов: для профиля, окружения, игрока, игры в целом.

Определение 6.3. Небезопасный для игрока i профиль s **доминируется безопасными стратегиями**, если угроза игроку i в нём является сильной и у него имеется гарантированный безопасный ответ. Для игрока i **безопасные стратегии доминируют при окружении s_{-i} (в игре G)**, если любая небезопасная для него стратегия при окружении s_{-i} (в игре G) доминируется безопасными стратегиями. Игра G называется **доминирующей в безопасных стратегиях**, если безопасные стратегии доминируют для всех игроков.

Для безопасных профилей условие выглядит следующим образом.

Определение 6.4. Безопасная стратегия s_i игрока i в профиле s , не являющимся РБС, **имеет гарантированный лучший ответ в безопасных стратегиях**, если $\forall (s, u) \in \overline{\Gamma(G(s, \bar{u}(s)))} \exists \alpha > u_i$

$\exists s'_i \in Q_i(s_{-i}), \exists$ открытая окрестность $U_{s_{-i}} \subseteq S_{-i}: \forall s'_{-i} \in U_{s_{-i}}: s' = (s'_i, s'_{-i}) \in Q^{(i)}$ (безопасен для i) и $u_i(s') \geq \alpha > u_i$.

Теперь, наконец, можно собрать то условие, которое требуется теоремой Рени, – *гарантию лучшего ответа*, помня о том, что в каждом профиле требуется его выполнение хотя бы для одного игрока.

Определение 6.5. Игра $G(S_i, u_i)_{i=1}^N$ **гарантирует лучший ответ в безопасных стратегиях**, если $\forall s \in S, \exists$ игрок i такой что: он имеет гарантированный безопасный ответ, если его стратегия s_i не безопасна, и гарантированный лучший ответ в безопасных стратегиях, если его стратегия s_i безопасна.

Итак, сформулированный «в лоб» вариант теоремы Рени, применённый для задачи поиска РБС, будет иметь следующий вид.

Теорема 6.1. *Предположим, игра $G(S_i, u_i)_{i=1}^N$ компактна, квазивогнута на безопасных стратегиях, все множества безопасных стратегий в ней являются выпуклыми компактными множествами, гарантирует лучший ответ в безопасных стратегиях, с сильными угрозами, тогда в игре G существует РБС.*

Доказательство. Условия «компактна, квазивогнута на безопасных стратегиях, имеет лучший ответ в безопасных стратегиях» являются ослаблением условий Рени на безопасные стратегии. Условие «гарантированного безопасного ответа» обеспечивает соблюдение условий теоремы Рени для не безопасных профилей. Условие «выпуклости безопасных множеств» необходимо для корректности условия квазивогнутости на безопасных стратегиях. А условие «с сильными угрозами» обеспечивает, что осторожные игроки не будут выходить за пределы безопасных стратегий. Общая схема доказательства следующая: введение обрезанной функции; применение к ней теоремы Рени; доказательство, что равновесие Нэша обрезанной игры будет РБС исходной игры.

Рассмотрим обрезанную игру $\bar{G} = (S_i, \bar{u}_i)_{i=1}^N$, где

$$\bar{u}_i(s) = \begin{cases} u_i(s), & s \in Q^{(i)}, \\ C_{\min}, & s \notin Q^{(i)}. \end{cases}$$

Проверкой определений 6.1, 6.2, 6.4 показываем, что обреченная игра $\bar{G} = (S_i, \bar{u}_i)_{i=1}^N$ компактна, квазивогнута, имеет лучший гарантированный ответ. Рассмотрим определение 6.1. Квазивогнутость игры означает квазивогнутость функции выигрыша $u_i(s_i, s_{-i})$ по s_i при любом окружении s_{-i} . То есть квазивогнутость на выпуклом множестве $Q_i(s_{-i})$ означает, что для любой константы C множество $M_C^{u_i}(s_{-i}) = \{s_i \in Q_i(s_{-i}): u_i(s_i, s_{-i}) \geq C\} \subseteq S_i$ выпукло. Это же является условием квазивогнутости игры $\bar{G}$.

Определение 6.2 для игры G эквивалентно выполнению требований определения Репу-2 лучшего гарантированного ответа для игры $\bar{G}$, для небезопасных профилей игры G : если $s_i \notin Q_i(s_{-i})$, то игрок i может в игре $\bar{G}$ гарантировать себе выигрыш $u_i(s_i', s_{-i}) \geq C_{min} = \bar{u}_i(s_i, s_{-i})$ выбором $s_i': (s_i', s_{-i}) \in Q^{(i)} \setminus \bar{Q}^{(i)}$.

Определение 6.4 для игры G эквивалентно выполнению требований определения Репу-2 лучшего гарантированного ответа для игры $\bar{G}$, для безопасных профилей игры G : если $s_i \in Q_i(s_{-i})$, то игрок i может в игре $\bar{G}$ гарантировать себе выигрыш $u_i(s_i', s_{-i}) \geq u_i(s_i, s_{-i}) = \bar{u}_i(s_i, s_{-i})$ выбором $s_i': (s_i', s_{-i}) \in Q^{(i)} \setminus \bar{Q}^{(i)}$.

Выполняются условия теоремы Рени, $\Rightarrow$ в игре $\bar{G} = (S_i, \bar{u}_i)_{i=1}^N$ имеется равновесие Нэша s^* . Согласно условию сильных угроз, это есть безопасный профиль игры G .

Применим условие сильных угроз. Пусть s^* равновесие Нэша в игре $\bar{G}$. Для любого i рассмотрим отклонение $s_i': v_i(s_i', s_{-i}^*) > u_i(s^*)$, где v_i – безопасный выигрыш. Безопасным для i профиль (s_i', s_{-i}^*) не может быть, так как s^* – равновесие Нэша в игре $\bar{G}$. То есть $v_i(s_i', s_{-i}^*) > u_i(s_i', s_{-i}^*)$. По определению игры с сильными угрозами, $\exists s_i'': u_i(s_i'', s_{-i}^*) > v_i(s_i', s_{-i}^*)$, где s_i'' – безопасная для i стратегия. $\Rightarrow u_i(s_i'', s_{-i}^*) > v_i(s_i', s_{-i}^*) > u_i(s^*)$, чего не может быть, так как s^* – равновесие Нэша в игре $\bar{G}$. Значит отклонение s_i' невозможно, для $\forall s_i': v_i(s_i', s_{-i}^*) \leq u_i(s^*)$, для безопасных отклонений, и $v_i(s_i', s_{-i}^*) < u_i(s^*)$, для не безопасных отклонений. То есть, согласно теореме 2.1, s^* – РБС игры $G(S_i, u_i)_{i=1}^N$. $\square$

Аналогично строим условие (#3) из теоремы 2.2 для **следствия Рени**. Игра должна быть *компактна, квазивогнута, гарантировать выигрыш* и быть *взаимно верхнеполунепрерывной* на безопасных множествах. Выполнение двух последних свойств следствия Рени является более узким частным случаем выполнения условия *гарантирования лучшего ответа* теоремы Рени и теоремы 6.1.

Аналогично формулируем требуемое понятие *гарантированного выигрыша* сначала для профиля.

Определение 6.6. Безопасная стратегия s_i игрока i в профиле s **имеет гарантированный выигрыш в безопасных стратегиях**, если $\forall \varepsilon > 0, \exists s'_i \in Q_i(s_{-i}), \exists$ открытая окрестность $U_{s_{-i}} \subseteq S_{-i}: \forall s'_{-i} \in U_{s_{-i}}: s' = (s'_i, s'_{-i}) \in Q^{(i)}$ (безопасен для i) и $u_i(s') \geq u_i(s) - \varepsilon$.

Теперь можно сформулировать понятие для игры, помня, что в этом случае оно должно соблюдаться для всех игроков.

Определение 6.7. Игра $G(S_i, u_i)_{i=1}^N$ **гарантирует выигрыш в безопасных стратегиях**, если для $\forall s \in S, \forall$ игрока i : имеется гарантированный безопасный ответ, если его стратегия s_i не безопасна, и гарантированный выигрыш в безопасных стратегиях, если его стратегия s_i безопасна.

Взаимная верхняя полунепрерывность на содержательном уровне формулируется как следующее свойство игры в точках разрыва целевых функций игроков: если в такой точке выигрыш какого-либо игрока скачком падает вниз, то должен найтись хотя бы один игрок, у которого в этой же точке выигрыш скачком изменяется вверх [5, 8, 9].

Определение 6.8. Игра $G(S_i, u_i)_{i=1}^N$ **взаимно-верхнеполунепрерывна в безопасных стратегиях (в.в.п.н.б.с.)** в профиле s , если для соответствующей этой игре обрезанной игры $\bar{G}(S_i, \bar{u}_i)_{i=1}^N$, для $\forall (s, u) \in \overline{\Gamma(\bar{u})}$ выполняется условие: $u_i(s) \leq \bar{u}_i$, $\forall i \Rightarrow u_i(s) = \bar{u}_i, \forall i$. Игра $G(S_i, u_i)_{i=1}^N$ **в.в.п.н.б.с.**, если она в.в.п.н.б.с. любом профиле $s \in S$.

А формально, если при просто верхней непрерывности значения всех функций в точках разрыва определяются по верхнему предельному значению, то для взаимной верхней полунепрерывности достаточно, чтобы это свойство выполнялось хотя бы для

одного игрока (из тех, у кого в этой точке имеется разрыв целевой функции). Возвращаясь к играм в безопасных стратегиях, можно заметить, что, во-первых, безопасные выигрыши игроков как раз часто разрывны на границах их множеств безопасности и эти множества безопасности часто незамкнуты, не содержат своих границ. А во-вторых, почти во всех рассмотренных играх с непрерывными множествами стратегий свойства взаимной верхней полунепрерывности и просто верхней полунепрерывности либо выполняются, либо не выполняются одновременно. Поэтому было сочтено возможным ослабить ограничение взаимной верхней полунепрерывности до простой верхней полунепрерывности, при которой каждого игрока можно рассматривать отдельно, и это намного удобнее, так как множества безопасности у всех игроков разные, границы их не совпадают и формулировать для них взаимность показалось слишком сложным и громоздким. Но здесь сначала получим формулировку теоремы «в лоб», по следствию Рени, с точным соблюдением всех условий.

Теорема 6.2. *Предположим, игра $G(S_i, u_i)_{i=1}^N$ компактна, квазивогнута на безопасных стратегиях, все множества безопасных стратегий в ней являются выпуклыми компактными множествами, гарантирует выигрыш на безопасных стратегиях, взаимно-верхнеполунепрерывна в безопасных стратегиях, с сильными угрозами, тогда в игре G существует РБС.*

Доказательство. Рассмотрим игру $\bar{G}$. Определение 6.7 $\Rightarrow \bar{G}$ гарантирует выигрыш. Определение 6.8 $\Rightarrow$ верхняя полунепрерывность в безопасных стратегиях $\bar{G}$. То есть $\bar{G}$ компактна, квазивогнута, гарантирует выигрыш, взаимно-верхнеполунепрерывна $\Rightarrow$ {утверждение Рени и следствие из него} $\Rightarrow \bar{G}$ компактна, квазивогнута, имеет лучший гарантированный ответ $\Rightarrow \bar{G}$ имеет равновесие по Нэшу. Далее – как в предыдущей теореме. $\square$

Чтобы получить **локальный вариант теорем** в множестве $B = \times_{i=1}^N B_i, B_i \subseteq S_i$, достаточно использовать определение 2.10 сильных угроз по отношению к B и переписать определения 6.1–6.8 «в множестве B »: на множествах $s_i \in Q_i(s_{-i}) \cup B_i$ вместо $s_i \in Q_i(s_{-i})$ для стратегий окружения $s_{-i} \in B_{-i}$. Тогда теоремы 6.1 и 6.2 предстанут в следующем виде.

Теорема 6.1'. *Предположим, игра $G(S_i, u_i)_{i=1}^N$ компактна, квазивогнута на безопасных стратегиях в V , все множества безопасных стратегий в ней являются выпуклыми компактными множествами в V , имеет лучший ответ в безопасных стратегиях в V , имеет лучший ответ в безопасных стратегиях в V , с сильными угрозами по отношению к V , тогда для игры G существует РБС в V .*

Теорема 6.2'. *Предположим, игра $G(S_i, u_i)_{i=1}^N$ компактна, квазивогнута на безопасных стратегиях в V , все множества безопасных стратегий в ней являются выпуклыми компактными множествами в V , гарантирует выигрыши на безопасных стратегиях в V , верхнеполунепрерывна в безопасных стратегиях в V , с сильными угрозами по отношению к V , тогда для игры G существует РБС в V .*

Доказательство теорем 6.1' и 6.2'. Введём игру $\bar{G}_V =, = (B_i, \bar{u}_i)_{i=1}^N$, определённую на множествах $s_i \in B_i$. В этой игре согласно теоремам 6.1 и 6.2 существует равновесие Нэша s^* , совпадающее с РН в игре $\bar{G} = (S_i, \bar{u}_i)_{i=1}^N$. Это равновесие s^* игры $\bar{G}_V(B_i, \bar{u}_i)_{i=1}^N$ при условии сильных угроз по отношению к V является РБС исходной игры $G(S_i, u_i)_{i=1}^N$. $\square$

Итак, получены теоремы 6.1 и 6.2, как формальное применение теоремы 2.2 к теореме Рени и следствию Рени. Но требования условий этих теорем оказываются слишком жёсткими для многих прикладных задач. Хотелось бы иметь такие утверждения, которые позволили бы применять их к задачам, рассматриваемым ниже. С другой стороны, желательно, чтобы эти условия были просты и легко проверяемы, формулировались бы в терминах простых свойств функций выигрыша или безопасного выигрыша. Также желательно, чтобы было достаточно выполнения этих условий только на локальном узком подмножестве безопасных стратегий, не привлекая значения целевых функций за их пределами. То есть чтобы достаточно было бы для всей игры выполнения локального условия сильных угроз, которые сами по себе гарантируют, что игра не выйдет из положения устойчивости, за пределы желаемого хорошего множества, а за этими пределами игра может вести себя сколь угодно плохо. Таким образом,

в дальнейших более сложных формальных конструкциях реализуется план коррекции теорем 6.1 и 6.2 в соответствии с этими пожеланиями, при этом доказательство полученных результатов существенно усложнилось.

7. Большая теорема. Первый вариант

Этот раздел написан в соавторстве с А.Б. Исаковым.

Строим теорему на основе теоремы 6.2 и следствия Рени. Сначала сформулируем модифицированный вариант условий Рени на *безопасных множествах* из этих утверждений: *квазивогнутость, гарантирование выигрыша, верхнюю полунепрерывность*. Для начала в качестве критерия существования равновесия Нэша в «безопасной области» воспользуемся стандартными условиями непрерывности и квазивогнутости функций выигрыша. В компактной и квазивогнутой игре с непрерывными функциями выигрыша существует равновесие Нэша (см. например, Дебре (1952) [10]).

Ещё раз сформулируем соответствующие понятия на множествах безопасности игроков, уделяя особенное внимание границам этих множеств, причём как той части граничного множества, которая входит в ограничиваемое им множество, так и той, что не входит. В качестве данных множеств рассматриваются множества безопасных стратегий и являющиеся их совокупностью множества безопасных профилей игроков. Но при этом совокупность границ множеств безопасных стратегий хотя и вложена в границу множества безопасных профилей, но часто с ней не совпадает. Кроме того, на этих множествах определены функции выигрыша и безопасного выигрыша игроков, в которых, особенно на границах множеств безопасности, встречаются точки недостижимых пределов. Для работы с ними будут рассматриваться (аналогично конструкциям у Рени (1999) [14]) графики этих функций, взятые как множества, и предельные точки этих множеств, в особенности не принадлежащие этим множествам. Теперь зададим эти конструкции формально.

Пусть $Q_i(s_{-i}) \subseteq S_i$ обозначает множество безопасных стратегий игрока i при заданных стратегиях других игроков s_{-i} ,

а $Q^{(i)} \subseteq S$ – множество всех профилей, в которых стратегии игрока i безопасны $\{s: s_i \in Q_i(s_{-i})\}$. Будут рассматриваться отдельно случаи внутренних безопасных стратегий $s \in Q^{(i)} \setminus \partial Q^{(i)}$ и граничных безопасных стратегий $s \in Q^{(i)} \cap \partial Q^{(i)}$. При рассмотрении условия квазивогнутости также потребуется понятие выпуклой оболочки множества безопасных стратегий $s_i \in \text{co}(Q_i(s_{-i}))$.

График функции безопасного выигрыша определяется как подмножество $S \times \mathbb{R}$ следующего вида: $\Gamma(v_i) = \{(s, v_i): v_i = v_i(s)\}$. При этом существенным будет рассмотрение точек, принадлежащих замыканию графика функции, $(s, v_i^*) \in \overline{\Gamma(v_i)}$, и точек недостижимых пределов функции безопасного выигрыша $(s, v_i^*) \in \overline{\Gamma(v_i)} \setminus \Gamma(v_i)$.

Теперь, переходя к рассмотрению условия Рени, прежде всего остановимся на условии непрерывности. Естественно ожидать, что функция выигрыша $u_i(s)$ и особенно безопасного выигрыша $v_i(s)$ игрока i может быть разрывной на границах множества его безопасных стратегий $\partial Q^{(i)}$. Поэтому можно наложить несколько усиленное требование непрерывности на $u_i(s) = v_i(s)$ только во внутренних безопасных стратегиях, когда $s \in Q^{(i)} \setminus \partial Q^{(i)}$, а случай границы безопасных стратегий, где $s \in Q^{(i)} \cap \partial Q^{(i)}$, рассмотреть особо.

Определение 7.1. Игра $G(S_i, u_i)_{i=1}^N$ является **непрерывной во внутренних безопасных стратегиях (continuous at interior secure strategies)**, если функция выигрыша $u_i(s)$ каждого игрока i непрерывна при $s \in Q^{(i)} \setminus \partial Q^{(i)}$.

Условие гарантированности выигрыша, предложенное Рени, будет наложено только на границе безопасных стратегий, $s \in Q^{(i)} \cap \partial Q^{(i)}$. Разумеется, непрерывность игры во внутренних безопасных стратегиях также означает, что игра гарантирует там выигрыш в том же смысле, что и определение 6.6 для границы безопасных стратегий, поскольку $u_i(s) = v_i(s)$ при $s \in Q^{(i)} \setminus \partial Q^{(i)}$ (в соответствии с утверждением 2.2).

Определение 7.2. Игра $G(S_i, u_i)_{i=1}^N$ **гарантирует выигрыш на границе безопасных стратегий (payoff secure at boundary secure strategies)**, если для каждого игрока i для каждого $s \in$

$Q^{(i)} \cap \partial Q^{(i)}$ и для каждого $\varepsilon > 0$ существует s'_i такое, что $v_i(s'_i, s'_{-i}) \geq u_i(s) - \varepsilon$ для всех s'_{-i} в некоторой окрестности s_{-i} .

Рассмотрим теперь квазивогнутость. Определение квазивогнутости 6.1 (к теореме 6.1) неявно предполагало выпуклость множеств безопасности. Теперь следует дать более точное и более сложное определение этого понятия при предположении, что этой выпуклости нет. Следуя стандартным определениям выпуклого анализа, определим квазивогнутую оболочку функции выигрыша $u_i(s_i, s_{-i})$ игрока i на множестве его безопасных стратегий $s_i \in Q_i(s_{-i})$ при заданных стратегиях других игроков s_{-i} в форме следующей функции, определённой при $s_i \in co(Q_i(s_{-i}))$: $\hat{u}_i(s_i, s_{-i}) = \inf\{h(s_i): h: co(Q_i(s_{-i})) \rightarrow \mathbb{R}, \text{ таких, что } h \text{ квазивогнута и на } Q_i(s_{-i}) \ u_i(\cdot, s_{-i}) \leq h(\cdot)\}$, т.е. квазивогнутой оболочки обрезанной функции. Тогда квазивогнутость игры в безопасных стратегиях можно определить следующим образом.

Определение 7.3. Игра $G(S_i, u_i)_{i=1}^N$ **квазивогнута в безопасных стратегиях (quasiconcave in secure strategies)**, если функция выигрыша каждого игрока i совпадает с ее квазивогнутой оболочкой на множестве безопасных стратегий этого игрока, т.е. $\forall i, \forall s: s_i \in Q_i(s_{-i})$ выполняется $u_i(s) = \hat{u}_i(s)$.

Итак, выполнение условий данных определений даёт выполнение следующих из условий Рени: квазивогнутость, гарантирование выигрыша на безопасных множествах, (верхняя полу-) непрерывность на внутренних безопасных множествах.

Теперь определим формально условие, аналогичное условию сильных угроз, согласно которому игроки с небезопасными стратегиями всегда могут отклониться в безопасные стратегии и увеличить при этом свой выигрыш, защищённый от угроз. В этом случае будем говорить, что игроки имеют *лучшую безопасную альтернативу (a better secure alternative)*.

Определение 7.4. При заданных стратегиях других игроков s_{-i} игрок i имеет **лучшую безопасную альтернативу (a better secure alternative) (BSA)** в непустом множестве $\tilde{Q}_i(s_{-i}) \subseteq Q_i(s_{-i})$, если для каждого $((s_i, s_{-i}), v_i^*) \in \overline{\Gamma(v_i)}, s_i \notin \tilde{Q}_i(s_{-i})$ и для каждого $((s_i, s_{-i}), v_i^*) \in \overline{\Gamma(v_i)} \setminus \Gamma(v_i), (s_i, s_{-i}) \in \partial Q^{(i)}$ существует $s'_i \in \tilde{Q}_i(s_{-i})$ такое, что $u_i(s'_i, s_{-i}) > v_i^*$.

Заметим, что в условие BSA включены также предельные значения функции $v_i(s)$ в точках разрывности, которые могут возникать на границе безопасных стратегий при $s \in Q^{(i)} \cap \partial Q^{(i)}$ и $s \in \partial Q^{(i)} \setminus Q^{(i)}$.

Определение 7.5. Игра $G(S_i, u_i)_{i=1}^N$ **предоставляет лучшую безопасную альтернативу (или является BSA-игрой) (provides better secure alternative or is a BSA-game)**, если для всех $s \in S$ каждый игрок i имеет лучшую безопасную альтернативу в некотором непустом множестве $\tilde{Q}_i(s_{-i}) \subseteq Q_i(s_{-i})$. Кроме того, если для всех $s \notin \times_{i=1}^N \tilde{Q}_i(s_{-i})$ существует игрок j такой, что $s_j \notin co(\tilde{Q}_j(s_{-j}))$, то игра $G(S_i, u_i)_{i=1}^N$ **предоставляет лучшую безопасную альтернативу в выпуклом множестве**.

Если попытаться нестрого-словесно охватить в целом требования, налагаемые всеми определениями, то можно сказать следующее. *Квазивогнутость* на невыпуклых безопасных множествах должна выполняться в смысле, обобщённом на этот случай. *Лучшая безопасная альтернатива* даёт игрокам возможность отклониться в безопасные стратегии и при этом увеличить свой выигрыш, защищённый от угроз. При этом в точке отклонения безопасный выигрыш игрока должен быть непрерывным по стратегиям окружения других игроков, чтобы выполнялось требование *гарантированного выигрыша*. Это особенно важно для граничных точек множества безопасных профилей, откуда должна быть обеспечена возможность отклонения туда, где есть непрерывность безопасного выигрыша по стратегиям окружения, даже если точка отклонения находится на границе множества безопасных стратегий или профилей. То есть необходимо обеспечить игроку в точке его отклонения как *безопасность* от односторонних угроз других игроков по отдельности, так и *гарантированность* от малых отклонений других игроков во всей их совокупности. Теперь можно сформулировать и доказать теорему существования.

Теорема 7.1. Если игра $G(S_i, u_i)_{i=1}^N$ компактна, квазивогнута в безопасных стратегиях, непрерывна во внутренних безопасных стратегиях, гарантирует выигрыши на границе безопасных стратегий и предоставляет лучшую безопасную альтернативу в выпуклом множестве, то в игре G существует РБС.

Доказательство.

Общий план доказательства, состоящего из шести частей.

(1) В специальной лемме доказывается представление для квазивогнутой оболочки функции, которое является основным рабочим инструментом, используемым для доказательства основного результата. (2) Вводится модифицированная игра, функцией выигрыша в которой принимается квазивогнутая оболочка безопасного выигрыша исходной игры. Для этой функции доказываются условия Баг, Джофр [7], из выполнения которых следует существование равновесия Нэша в этой игре. А именно, доказываются свойства гарантирования выигрыша (3) и слабой взаимной верхней полунепрерывности (4). Далее доказывается, что существующее равновесие Нэша модифицированной игры является безопасным профилем исходной игры (5). После чего доказывается, что этот профиль является РБС в исходной игре (6). В конце каждой части и всего доказательства приводится краткий конспект хода рассуждений без формальных выкладок, выделенный *курсивом*.

(1). Прежде всего докажем представление для квазивыпуклой оболочки функции (по аналогии с представлением, полученным в Бик, [8]), которое нам понадобится при доказательстве.

Лемма 7.1. *Квазивогнутую оболочку функции f , определяемую в виде:*

(1) $s \in Y \subseteq \mathbb{R}^d: \tilde{f}(s) = \inf\{h(s): h: Y \rightarrow \mathbb{R}, \text{ таких, что } h \text{ квазивогнута и на } Y \quad f(\cdot) \leq h(\cdot)\},$

можно представить в следующей форме:

(2) $\tilde{f}(s) = \sup\{\min\{f(s^{(1)}), \dots, f(s^{(d+1)})\}\},$

где d – это размерность множества $Y \subseteq \mathbb{R}^d$ и supremum берётся по всем семействам точек $\{s^{(1)}, \dots, s^{(d+1)}\}$ из Y таким, что $s \in \text{co}\{s^{(1)}, \dots, s^{(d+1)}\}$.

Доказательство Леммы.

В соответствии с (Бик [8], Утверждение 2.2) квазивогнутую оболочку (1) можно представить в форме

$$(3) \quad \hat{f}(s) = \sup_{\substack{\forall k \in \mathbb{N}, \{s^{(1)}, \dots, s^{(k)}\} \in Y^k: \\ s \in \text{co}\{s^{(1)}, \dots, s^{(k)}\}}} \{\min\{f(s^{(1)}), \dots, f(s^{(k)})\}\}.$$

Докажем, что (2) и (3) совпадают, т.е. $\tilde{f} = \hat{f}$. Очевидно, что $\hat{f} \geq \tilde{f}$, поскольку supremum в (3) берётся по более широкому множеству, чем в (2). С другой стороны, для всех $\forall k \in \mathbb{N}$ и для всех таких семейств точек $\{s^{(1)}, \dots, s^{(k)}\} \in Y^k$, что $s \in co\{s^{(1)}, \dots, s^{(k)}\}$, в соответствии с теоремой Каратеодори (1911) о выпуклой оболочке, из множества точек $\{s^{(1)}, \dots, s^{(k)}\}$ всегда можно выбрать $d + 1$ таких точек $\{\tilde{s}^{(l)}\} \in \{s^{(1)}, \dots, s^{(k)}\}$, $l = 1, \dots, d + 1$, (быть может повторно), что $s \in co\{\tilde{s}^{(1)}, \dots, \tilde{s}^{(d+1)}\}$. В этом случае, очевидно, выполняется

$$\min\{f(\tilde{s}^{(1)}), \dots, f(\tilde{s}^{(d+1)})\} \geq \min\{f(s^{(1)}), \dots, f(s^{(k)})\}.$$

Переходя в этом неравенстве к supremum по всем $k \in \mathbb{N}$ и по всем соответствующим семействам точек $\{s^{(1)}, \dots, s^{(k)}\}$, получаем $\tilde{f}(s) \geq \hat{f}(s)$. Таким образом, равенство $\tilde{f} = \hat{f}$ доказано, и представление квазивогнутой оболочки в виде (2) выполняется. $\square$

(2). Определим модифицированную игру $\hat{\hat{G}}(S_i, \hat{v}_i)_{i=1}^N$, где $\hat{v}_i(s_i, s_{-i})$ – квазивогнутая оболочка функции $v_i(s_i, s_{-i})$ на множестве $s_i \in S_i$. По построению игра $\hat{\hat{G}}$ квазивогнутая. Она также компактная, поскольку исходная игра G компактная. Для дальнейшего доказательства нам понадобится представление для квазивогнутой оболочки $\hat{v}_i(s)$ полученное в Лемме 7.1.:

$$(4) \quad \hat{v}_i(s_i, s_{-i}) = \sup_{\substack{\{s_i^{(1)}, \dots, s_i^{(d+1)}\} \in (S_i)^{(d+1)}: \\ s_i \in co\{s_i^{(1)}, \dots, s_i^{(d+1)}\}}} \left\{ \min \left\{ v_i(s_i^{(1)}, s_{-i}), \dots, v_i(s_i^{(d+1)}, s_{-i}) \right\} \right\},$$

$$\forall s_i \in S_i \subseteq \mathbb{R}^d,$$

где d – размерность множества $S_i \subseteq \mathbb{R}^d$ и супремум берётся по всем семействам точек $\{s_i^{(1)}, \dots, s_i^{(d+1)}\}$ из S_i таким, что $s_i \in co\{s_i^{(1)}, \dots, s_i^{(d+1)}\}$.

В соответствии с утверждением 2.2, $\forall s_i \in Q_i(s_{-i}): u_i(s_i, s_{-i}) = v_i(s_i, s_{-i})$. Введём определённую на $\{s: s_i \in co(Q_i(s_{-i}))\}$ функцию $\hat{u}_i(s)$ квазивогнутой оболочки

обрезанной функции. Используя приведённое выше представление леммы 7.1 получаем для этой функции следующее неравенство:

$$\begin{aligned}
 (5) \quad \hat{u}_i(s_i, s_{-i}) &= \\
 &= \sup_{\substack{\{s_i^{(1)}, \dots, s_i^{(d+1)}\} \in (Q_i(s_{-i}))^{(d+1)}: \\ s_i \in co\{s_i^{(1)}, \dots, s_i^{(d+1)}\}}} \left\{ \min \left\{ v_i \left(s_i^{(1)}, s_{-i} \right), \dots, v_i \left(s_i^{(d+1)}, s_{-i} \right) \right\} \right\} \leq \\
 &\leq \sup_{\substack{\{s_i^{(1)}, \dots, s_i^{(d+1)}\} \in (S_i)^{(d+1)}: \\ s_i \in co\{s_i^{(1)}, \dots, s_i^{(d+1)}\}}} \left\{ \min \left\{ v_i \left(s_i^{(1)}, s_{-i} \right), \dots, v_i \left(s_i^{(d+1)}, s_{-i} \right) \right\} \right\} = \\
 &= \hat{v}_i(s_i, s_{-i}).
 \end{aligned}$$

То есть квазивогнутая оболочка обрезанной функции ограничена сверху квазивогнутой оболочкой функции безопасного выигрыша.

Мы докажем, что в модифицированной игре $\hat{G}$ существует равновесие Нэша, используя теорему существования, предложенную Рени (1999) [14] и улучшенную Баг, Джофр (2006) [7]. Напомним еще раз определения из этих работ.

Определение 7.6 (Рени-3). Игра $G = (S_i, u_i)_{i=1}^N$ **гарантирует выигрыш** (payoff secure – PS), если для любого игрока i , для любого $s \in S$ и любого $\varepsilon > 0$, существует $s'_i \in S_i$, такое что $u_i(s'_i, s'_{-i}) \geq u_i(s) - \varepsilon$ для всех s'_{-i} в некоторой открытой окрестности s_{-i} .

Определение 7.7 (Баг, Джофр [7]). Игра $G = (S_i, u_i)_{i=1}^N$ **слабо взаимно верхнеполунепрерывна (с.в.в.п.н.)** (weakly reciprocally upper semicontinuous – WRUSC), если для любого $(s^*, u^*) \in \bar{\Gamma}(u) \setminus \Gamma(u)$ существует игрок i и $s'_i \in S_i$, такие что $u_i(s'_i, s_{-i}^*) > u_i^*$.

Баг, Джофр [7] доказали, что в каждой компактной, квазивогнутой, (PS) и (WRUSC) игре существует равновесие Нэша. Мы докажем, что модифицированная игра $\hat{G}$ удовлетворяет этим условиям, а затем покажем, что каждое равновесие Нэша в $\hat{G}$ является РБС в исходной игре G .

(3). Докажем, что $\hat{G}$ гарантирует выигрыш (определение 7.6). Выберем любого игрока i , любой профиль $s = (s_i, s_{-i}) \in S$ и любое

$\varepsilon > 0$. Из (4) следует, что существует $s'_i \in S_i$ такое, что $v_i(s'_i, s_{-i}) > \hat{v}_i(s) - \varepsilon/2$. Либо $s'_i \in Q_i(s_{-i})$ и мы имеем $u_i(s'_i, s_{-i}) = v_i(s'_i, s_{-i})$, либо $s'_i \notin Q_i(s_{-i})$, и поскольку $((s'_i, s_{-i}), v_i(s'_i, s_{-i})) \in \overline{\Gamma(v_i)}$, то из BSA условия (определения 7.4 и 7.5) для игрока i следует, что существует $s''_i \in Q_i(s_{-i})$ такое, что $u_i(s''_i, s_{-i}) > v_i(s'_i, s_{-i})$. Таким образом, в обоих случаях существует $s''_i \in Q_i(s_{-i})$ такое, что $u_i(s''_i, s_{-i}) = \hat{v}_i(s) - \varepsilon/2$. Игра G по условию теоремы непрерывна во внутренних безопасных стратегиях (определение 7.1) и гарантирует выигрыш на границе безопасных стратегий (определение 7.2). Отсюда следует, что условие гарантированности выигрыша (в смысле определения 6.6) для u_i удовлетворяется во всех безопасных стратегиях, и в частности для $s''_i \in Q_i(s_{-i})$. Используя это условие для $\varepsilon/2$, мы получаем, что существует $s'''_i \in Q_i(s_{-i})$ такое, что $v_i(s'''_i, s'_{-i}) = u_i(s''_i, s_{-i}) - \varepsilon/2 > \hat{v}_i(s) - \varepsilon$ для всех s'_{-i} в некоторой окрестности s_{-i} . Из представления квазивогнутой оболочки (4) следует, что $\hat{v}_i(s'''_i, s'_{-i}) \geq v_i(s'''_i, s'_{-i})$, и мы получаем условие гарантированности выигрыша для функции $\hat{v}(s)$, т.е. по определению Рени $\hat{G}$ гарантирует выигрыш (определение 7.6 (Рени-3)).

Таким образом, для получения требуемой определением Рени-3 гарантированного выигрыша оценки требуется ряд переходов: от квазивогнутой оболочки безопасного выигрыша $\hat{v}_i(s)$ к безопасному выигрышу $v_i(s)$, к выигрышу в исходной игре $u_i(s)$ и обратно, к $v_i(s)$ и $\hat{v}_i(s)$. Оценка гарантированного выигрыша получается следующим образом: $\hat{v}_i(s'''_i, s'_{-i}) \geq \{\text{представление (4) квазивогнутой оболочки по лемме 7.1}\} \geq v_i(s'''_i, s'_{-i}) \geq \{\text{условия теоремы непрерывности во внутренних безопасных стратегиях и гарантирования выигрыша на границе безопасных стратегий, определения 7.1, 7.2}\} \geq u_i(s''_i, s_{-i}) - \varepsilon/2 > \{\text{определения 7.4, 7.5 лучшей безопасной альтернативы}\} > v_i(s'_i, s_{-i}) - \varepsilon/2 \geq \{\text{представление (4) квазивогнутой оболочки по лемме 7.1}\} \geq \hat{v}_i(s) - \varepsilon$.

(4). Докажем, что $\hat{G}$ является WRUSC. Рассмотрим произвольный вектор $(s^*, \hat{v}^*) \in \overline{\Gamma(\hat{v})} \setminus \Gamma(\hat{v})$. Существует последователь-

ность профилей $\{s^n\} \rightarrow s^*: \{\hat{v}(s^n)\} \rightarrow \hat{v}^*$. Для произвольного i используем выражение (4) для $\hat{v}_i(s^n)$ и выберем последовательность семейств точек $\{s_i^{n(1)}, \dots, s_i^{n(d+1)}\}$ такую, что

$$\forall n: s_i^n \in \text{co}\{s_i^{n(1)}, \dots, s_i^{n(d+1)}\}, \min\{v_i(s_i^{n(1)}, s_{-i}^n), \dots, v_i(s_i^{n(d+1)}, s_{-i}^n)\} \xrightarrow{n \rightarrow \infty} \hat{v}_i^*.$$

Поскольку игра $\tilde{G}$ компактна, то из этой последовательности можно выбрать такую подпоследовательность $\{\tilde{s}_i^{n(l)}, \dots, \tilde{s}_i^{n(d+1)}\}$, что все её члены и соответствующие величины функций сходятся:

$$\{\tilde{s}_i^{n(l)}\} \rightarrow s_i^{(l)}, \{v_i(\tilde{s}_i^{n(l)}, \tilde{s}_{-i}^n)\} \rightarrow v_i^{(l)} \geq \hat{v}_i^*, l = 1, \dots, d+1.$$

Поскольку для всех n мы имеем $\tilde{s}_i^n \in \text{co}\{\tilde{s}_i^{n(1)}, \dots, \tilde{s}_i^{n(d+1)}\}$ и $\{\tilde{s}_i^n\} \rightarrow s_i^*$, то в пределе мы получаем $s_i^* \in \text{co}\{s_i^{(1)}, \dots, s_i^{(d+1)}\}$.

Возможны два случая. Либо существуют такие i и l , что точка $\left((s_i^{(l)}, s_{-i}^*), v_i^{(l)}\right) \in \overline{\Gamma(v_l)}$ удовлетворяет BSA условию для игрока i (определение 7.4). В таком случае в соответствии с этим условием существует такая $s'_i \in \tilde{Q}_i(s_{-i}^*) \subseteq Q_i(s_{-i}^*)$, что $u_i(s'_i, s_{-i}^*) > v_i^{(l)} \geq \hat{v}_i^*$. Учитывая (5) и условие квазивогнутости $\hat{u}_i(s'_i, s_{-i}^*) = u_i(s'_i, s_{-i}^*)$ для $s'_i \in \tilde{Q}_i(s_{-i}^*) \subseteq Q_i(s_{-i}^*)$, получаем $\hat{v}_i(s'_i, s_{-i}^*) \geq \hat{u}_i(s'_i, s_{-i}^*) = u_i(s'_i, s_{-i}^*) > \hat{v}_i^*$, т.е. условие WRUSC выполняется для точки $(s^*, \hat{v}^*)$ (определение 7.7 Баг, Джофр [7]).

Во втором случае для любых i и l точка $\left((s_i^{(l)}, s_{-i}^*), v_i^{(l)}\right) \in \overline{\Gamma(v_l)}$ не удовлетворяет BSA условию для игрока i (определение 7.4). Отрицание условий определения означает, что $\forall i, \forall l: s_i^{(l)} \in Q_i(s_{-i}^*)$, т.е. $v_i(s_i^{(l)}, s_{-i}^*) = u_i(s_i^{(l)}, s_{-i}^*)$. Кроме того, по BSA-условию, это означает, что либо $\left((s_i^{(l)}, s_{-i}^*), v_i^{(l)}\right) \in \Gamma(v_i)$; либо $(s_i^{(l)}, s_{-i}^*) \notin \partial Q^{(i)}$, т.е. $v_i = u_i$ в некоторой окрестности $(s_i^{(l)}, s_{-i}^*)$, и из непрерывности u_i во внутренних безопасных стратегиях также следует непрерывность v_i в $(s_i^{(l)}, s_{-i}^*)$.

В обоих случаях имеем: $\forall l: \hat{v}_i^* \leq v_i^{(l)} = \lim_{n \rightarrow \infty} v_i(\tilde{s}_i^{n(l)}, \tilde{s}_{-i}^n) = v_i(s_i^{(l)}, s_{-i}^*) = u_i(s_i^{(l)}, s_{-i}^*)$.

Поскольку $s_i^* \in co\{s_i^{n(1)}, \dots, s_i^{n(d+1)}\}$, отсюда, согласно квазивогнутости $u_i(s)$ на безопасных стратегиях, получаем:

$$\begin{aligned} \hat{u}_i(s^*) &\geq \min\{u_i(s_i^{(1)}, s_{-i}^*), \dots, u_i(s_i^{(d+1)}, s_{-i}^*)\} = \\ &= \min\{v_i(s_i^{(1)}, s_{-i}^*), \dots, v_i(s_i^{(d+1)}, s_{-i}^*)\} \geq \hat{v}_i^*. \end{aligned}$$

Учитывая, что согласно (5), $\hat{u}_i(s_i, s_{-i}) \leq \hat{v}_i(s_i, s_{-i})$, имеем $\hat{v}_i(s^*) \geq \hat{v}_i^*$ для любых i . Но поскольку с самого начала мы предполагали $(s^*, \hat{v}^*) \notin \Gamma(\hat{v})$, то существует такое j , что $\hat{v}_j(s^*) > \hat{v}_j^*$. Это означает, что условие WRUSC (определение 7.7 Баг, Джоффр [7]) удовлетворяется для игрока j и $s_j' = s_j^*$.

Таким образом, чтобы получить оценку слабой взаимной верхней полунепрерывности для произвольной точки недостижимого предела $(s^*, \hat{v}^*)$ из замыкания графика игры, требуется следующий ряд шагов: для каждого игрока i построить последовательность окружающих профиль s^* семейств точек $\{s_i^{n(1)}, \dots, s_i^{n(d+1)}\}$, сходящихся к ней: $\{s^n\} \rightarrow s^*: \{\hat{v}(s^n)\} \rightarrow \hat{v}^*$; выделить из неё предельное семейство (набор) $\left((s_i^{(l)}, s_{-i}^*), v_i^{(l)}\right)$ окружающий s^* и ограничивающих сверху значение недостижимого предела $v_i^{(l)} \geq \hat{v}_i^*$; рассмотреть два случая выполнения хотя бы для одной из точек предельного семейства и хотя бы одного игрока / (невыполнения для всех точек и игроков) условий BSA; в первом случае оценкой будет значение совпадающих функций выигрыша в точке отклонения в безопасную область $v_i = u_i = \hat{u}_i = \hat{v}_i(s_i', s_{-i}^*)$; во втором случае оценка $\hat{v}_i(s_i', s_{-i}^*)$ получается через минимальное значение из предельных точек окружающего семейства $v_i(s_i^{(l)}, s_{-i}^*) = v_i^{(l)}$, лежащих в безопасном множестве; строгое выполнение неравенства $\hat{v}_j(s^*) > \hat{v}_j^*$ хотя бы для одного игрока следует из того, что что профиль $(s^*, \hat{v}^*)$ является точкой недостижимого предела.

(5). Покажем, что равновесие игры $\hat{G}$ является безопасным профилем. Мы доказали, что игра $\hat{G}$ квазивогнута, гарантирует выигрыш (PS) и является WRUSC. По теореме Баг, Джофр [7] в ней существует равновесие Нэша s^* .

Предположим, что s^* является не безопасным профилем. Тогда $s^* \notin \times_{i=1}^N \tilde{Q}_i(s_{-i}^*)$ и в соответствии с BSA условием существует игрок j со стратегией $s_j^* \notin co(\tilde{Q}_j(s_{-j}^*))$ в профиле s^* . Используя представление (4) для $\hat{v}_j(s^*)$, выберем такую последовательность семейств точек $\{s_j^{n(1)}, \dots, s_j^{n(d+1)}\}$, чтобы

$$\begin{aligned} \forall n: s_j^* &\in \\ &\in co\{s_j^{n(1)}, \dots, s_j^{n(d+1)}\}, \min\{v_j(s_j^{n(1)}, s_{-j}^*), \dots, v_j(s_j^{n(d+1)}, s_{-j}^*)\} \\ &\xrightarrow{n \rightarrow \infty} \hat{v}_j(s^*). \end{aligned}$$

Существует некоторое положительное расстояние $\rho_0 = dist\left(\{s_j^*\}, co\left(\tilde{Q}_j(s_{-j}^*)\right)\right) > 0$ между точкой $\{s_j^*\}$ и замкнутым множеством $co\left(\tilde{Q}_j(s_{-j}^*)\right) \subseteq S_j$. Для каждого n выберем такую точку $\{s_j^{n(l)}\} \in \{s_j^{n(1)}, \dots, s_j^{n(d+1)}\}$, чтобы расстояние $dist\left(\{s_j^{n(l)}\}, co\left(\tilde{Q}_j(s_{-j}^*)\right)\right) \geq \rho_0 > 0$. Её можно выбрать благодаря выпуклости множества $co\left(\tilde{Q}_j(s_{-j}^*)\right)$. Выберем из последовательности $\{s_j^{n(l)}\}$ сходящуюся подпоследовательность $\{\tilde{s}_j^{n(l)}\} \rightarrow s_j'$ так, чтобы значения соответствующих величин v_j также сходились: $\{v_j(\tilde{s}_j^{n(l)}, s_{-j}^*)\} \rightarrow v_j' \geq \hat{v}_j(s^*)$. Для каждого n расстояние $dist\left(\{\tilde{s}_j^{n(l)}\}, co\left(\tilde{Q}_j(s_{-j}^*)\right)\right) \geq \rho_0 > 0$ и потому $s_j' \notin \tilde{Q}_j(s_{-j}^*)$. Из условия BSA (определение 7.4) для точки $((s_j', s_{-j}^*), v_j') \in \overline{\Gamma(v_j)}$ существует такая стратегия $s_j'' \in \tilde{Q}_j(s_{-j}^*) \subseteq Q_j(s_{-j}^*)$, что $u_j(s_j'', s_{-j}^*) > v_j' \geq \hat{v}_j(s^*)$. Поскольку (s_j'', s_{-j}^*) является безопасным профилем для игрока j , то $u_j(s_j'', s_{-j}^*) = v_j(s_j'', s_{-j}^*) \leq \hat{v}_j(s_j'', s_{-j}^*)$, т.е. мы получаем $\hat{v}_j(s_j'', s_{-j}^*) > \hat{v}_j(s^*)$ в противоречии с условием, что s^* является равновесием в $\hat{G}$. Поэтому наше

предположение было неверным и s^* является безопасным профилем.

Таким образом эта часть доказательства проводится от противного: принимается допущение, что существующее равновесие Нэша s^* игры $\hat{G}$ не является безопасным профилем, и из него выводится противоречие. Согласно условию лучшей безопасной альтернативы в выпуклом множестве, это равновесие должно хотя бы для одного игрока находиться вне выпуклой оболочки его множества безопасных стратегий, на некотором расстоянии ρ_0 от него. Дальнейшее рассуждение строится при помощи тех же приёмов, что и в предыдущих частях доказательства: через построение последовательности окружающих s^* семейств точек и выделение предельного семейства из этой последовательности осуществляется переход от функции квазивогнутой оболочки безопасного выигрыша к безопасному выигрышу; потом из безопасного множества выбирается лучшая безопасная альтернатива, в которой все функции выигрыша, безопасного выигрыша и их квазивогнутые оболочки равны $v_i = u_i = \hat{u}_i = \hat{v}_i$ и выигрыш в которой превосходит выигрыш в точке равновесия $\hat{v}_i(s^*)$. Полученное противоречие опровергает сделанное предположение.

(6). Докажем существование РБС в исходной игре. Покажем, что $v(s^*) = \hat{v}(s^*)$. Выберем произвольное i и воспользуемся опять выражением (4) для $\hat{v}_i(s^*)$. Возможны два случая.

В первом случае из любой последовательности семейств точек $\{s_i^{n(1)}, \dots, s_i^{n(d+1)}\} \in (S_i)^{d+1}$, удовлетворяющих условиям

$$(6) \quad \forall n: s_i^* \in \text{co}\{s_i^{n(1)}, \dots, s_i^{n(d+1)}\}, \min\{v_i(s_i^{n(1)}, s_{-i}^*), \dots, v_i(s_i^{n(d+1)}, s_{-i}^*)\} \xrightarrow{n \rightarrow \infty} \hat{v}_i(s^*)$$

можно выбрать такую подпоследовательность

$$(7) \quad \{\tilde{s}_i^{n(1)}, \dots, \tilde{s}_i^{n(d+1)}\}: \{\tilde{s}_i^{n(l)}\} \rightarrow s_i^{(l)}, \{v_i(\tilde{s}_i^{n(l)}, s_{-i}^*)\} \rightarrow v_i^{(l)} \geq \hat{v}_i(s^*);$$

$$l = 1, \dots, d + 1$$

что для всех l BSA условия не выполняются для $\left((s_i^{(l)}, s_{-i}^*), v_i^{(l)}\right) \in \overline{\Gamma(v_l)}$ (определение 7.4). Это означает, что

$\forall l: s_i^{(l)} \in \tilde{Q}_i(s_{-i}^*) \subseteq Q_i(s_{-i}^*)$, и также, что либо $\left((s_i^{(l)}, s_{-i}^*), v_i^{(l)} \right) \in \Gamma(v_i)$, либо $\left(s_i^{(l)}, s_{-i}^* \right) \notin Q^{(i)}$, т.е. $u_i = v_i$ в некоторой окрестности $\left(s_i^{(l)}, s_{-i}^* \right)$, и из непрерывности u_i во внутренних безопасных стратегиях также следует непрерывность $v_i(s)$ в $\left(s_i^{(l)}, s_{-i}^* \right)$. В обоих случаях мы имеем $v_i\left(s_i^{(l)}, s_{-i}^*\right) = v_i^{(l)}$, и следовательно,

$\forall l: u_i(s_i^{(l)}, s_{-i}^*) = v_i(s_i^{(l)}, s_{-i}^*) = v_i^{(l)}$ и $\hat{u}_i(s^*) \geq \min\{v_i^{(1)}, \dots, v_i^{(d+1)}\} \geq \hat{v}_i(s^*)$. Но из (5) мы имеем $\hat{u}_i(s^*) \leq \hat{v}_i(s^*)$, и потому $\hat{u}_i(s^*) = \hat{v}_i(s^*)$. Из квазивогнутости игры G в s^* и безопасности профиля s^* мы получаем $v_i(s^*) = u_i(s^*) = \hat{u}_i(s^*) = \hat{v}_i(s^*)$, и требуемое равенство $v_i(s^*) = \hat{v}_i(s^*)$ доказано.

Во втором случае существует последовательность семейств точек $\{s_i^{n(1)}, \dots, s_i^{n(d+1)}\} \in (S_i)^{d+1}$, удовлетворяющих (6) и таких, что в любой подпоследовательности (7) существует такое l , что BSA условие удовлетворяется для точки $\left((s_i^{(l)}, s_{-i}^*), v_i^{(l)} \right) \in \overline{\Gamma(v_i)}$. В соответствии с определением 7.4 мы получаем

$$\exists s_i'' \in Q_i(s_{-i}^*): u_i(s_i'', s_{-i}^*) > v_i^{(l)} \geq \hat{v}_i(s^*), \hat{v}_i(s_i'', s_{-i}^*) \geq u_i(s_i'', s_{-i}^*)$$

в противоречие с тем, что s^* является равновесием Нэша в $\hat{G}$. Таким образом, равенство $v_i(s^*) = \hat{v}_i(s^*)$ доказано для каждого i .

Поскольку $\forall i, \forall l: \hat{v}_i(s^*) \geq v_i(s^*)$, отсюда следует, что s^* – равновесие Нэша не только в игре $\hat{G}$, но также равновесие Нэша в игре угроз $\tilde{G} = (S_i, v_i)_{i=1}^N$. Рассмотрим отклонение s'_i произвольного игрока i . Если $s'_i \in Q_i(s_{-i}^*)$, то $u_i(s'_i, s_{-i}^*) = v_i(s'_i, s_{-i}^*) \leq v_i(s^*) = u_i(s^*)$ и отклонение не выгодно. Если $s'_i \notin Q_i(s_{-i}^*)$, то $\left((s'_i, s_{-i}^*), v_i(s'_i, s_{-i}^*) \right) \in \overline{\Gamma(v_i)}$, и в соответствии с BSA условием существует такое $s_i'' \in Q_i(s_{-i}^*)$, что $u_i(s_i'', s_{-i}^*) > v_i(s'_i, s_{-i}^*)$. Поскольку $u_i(s_i'', s_{-i}^*) = v_i(s_i'', s_{-i}^*) \leq v_i(s^*) = u_i(s^*)$, мы получаем $v_i(s'_i, s_{-i}^*) < u_i(s^*)$, и в соответствии с утверждением 2.2 отклонение s'_i не безопасно. Таким образом, в безопасном профиле s^* ни один игрок не имеет выгодных безопасных отклонений, т.е. s^* является РБС в исходной игре G .

Таким образом, для доказательства $v(s^*) = \hat{v}(s^*)$ снова были рассмотрены всевозможные последовательности окружающих равновесный профиль наборов точек (s, v) , из них выделяются предельные наборы, ограничивающие сверху значение $\hat{v}_i(s^*)$; далее применяется условие BSA, если оно не выполняется для всех последовательностей и предельных точек, то s^* находится во внутренних безопасных стратегиях, если выполняется хотя бы в одном случае, то из BSA-условия получается противоречие. Из того, что $v(s^*) = \hat{v}(s^*) = u(s^*) = \hat{u}(s^*)$ во внутренних безопасных стратегиях из элементарных свойств РБС следует, что s^* – РБС.

Итак, результат во всех частях доказательства достигается переходами между функциями $v, \hat{v}, u, \hat{u}$ и обратно, которые осуществлялись посредством формулы представления леммы 7.1, отклонения в безопасную область по условиям определения лучшей безопасной альтернативы и через построение последовательности окружающих семейств профилей и выделения из них предельного набора. $\square$

Заметим, что **условие квазивогнутости** в безопасных стратегиях **может быть ослаблено**. Доказательство практически не изменяется, если потребовать выполнения условия квазивогнутости функций выигрыша не на множествах безопасных стратегий, а только на их подмножествах $\tilde{Q}_i(s_{-i}) \subseteq Q_i(s_{-i})$ из BSA условия. Назовем их BSA-множествами.

Определение 7.8. BSA-Игра $G(S_i, u_i)_{i=1}^N$ называется **квазивогнутой на BSA-множествах**, если функция выигрыша каждого игрока i совпадает со своей квазивогнутой оболочкой на BSA-множествах, т.е. $\forall i, \forall s: s_i \in \tilde{Q}_i(s_{-i})$ выполняется $u_i(s) = \hat{u}'_i(s)$, где $\hat{u}'_i(s_i, s_{-i}) = \inf\{h(s_i): co(\tilde{Q}_i(s_{-i})) \rightarrow \mathbb{R}: \text{квазивогнутая, } u_i(s_i, s_{-i}) \leq h(s_i), s_i \in \tilde{Q}_i(s_{-i})\}$.

Отсюда получаем следующее Следствие.

Следствие 7.1. Если игра $G(S_i, u_i)_{i=1}^N$ компактна, непрерывна во внутренних безопасных стратегиях, гарантирует выигрыш в границе безопасных стратегий и предоставляет лучшую безопасную альтернативу в выпуклом множестве и квазивогнута на BSA-множествах, то в игре G существует РБС.

Доказательство. Будем проводить его по схеме доказательства теоремы 7.1, отмечая необходимые изменения, заключающиеся в замене множества безопасных стратегий на его BSA-подмножество из определения 7.4: $Q_i(s_{-i}) \rightarrow \tilde{Q}_i(s_{-i})$; а также квазивогнутой оболочки функции обрезанной на безопасном множестве на соответствующую функцию на BSA-множестве из определения 7.6: $\hat{u}_i(s) \rightarrow \hat{u}'_i(s)$.

(1). Основным инструментом доказательства является представление (2) квазивогнутой оболочки функции (1) из леммы 7.1 без изменений.

(2). Для модифицированной игры $\hat{G}(S_i, \hat{v}_i)_{i=1}^N$, определяемой равенством (4), при помощи представления леммы 7.1 вместо оценки (5) получаем оценку

$$\begin{aligned} (8) \quad \hat{u}'_i(s_i, s_{-i}) &= \sup_{\substack{\{s_i^{(1)}, \dots, s_i^{(d+1)}\} \in (\tilde{Q}_i(s_{-i}))^{(d+1)}: \\ s_i \in co\{s_i^{(1)}, \dots, s_i^{(d+1)}\}}} \left\{ \min \left\{ v_i \left(s_i^{(1)}, s_{-i} \right), \dots, v_i \left(s_i^{(d+1)}, s_{-i} \right) \right\} \right\} \leq \\ &\leq \sup_{\substack{\{s_i^{(1)}, \dots, s_i^{(d+1)}\} \in (S_i)^{(d+1)}: \\ s_i \in co\{s_i^{(1)}, \dots, s_i^{(d+1)}\}}} \left\{ \min \left\{ v_i \left(s_i^{(1)}, s_{-i} \right), \dots, v_i \left(s_i^{(d+1)}, s_{-i} \right) \right\} \right\} = \\ &= \hat{v}_i(s_i, s_{-i}) \end{aligned}$$

То есть квазивогнутая оболочка обрезанной на BSA-множестве функции ограничена сверху квазивогнутой оболочкой функции безопасного выигрыша.

(3). Доказательство, что $\hat{G}$ гарантирует выигрыш – идентично соответствующему доказательству теоремы, учитывая, что используемые там отклонения s'_i, s''_i, s'''_i определяются по BSA-условию, а это означает, что эти отклонения принадлежат не только безопасному множеству, но и его BSA-подмножеству: $s'_i, s''_i, s'''_i \in \tilde{Q}_i(s_{-i})$.

(4). Доказательство, что $\hat{G}$ является WRUSC. Как в теореме, для произвольного вектора $(s^*, \hat{v}^*) \in \overline{\Gamma(\hat{v})} \setminus \Gamma(\hat{v})$ строится последовательность окружающих профиль s^* семейств точек $\{s_i^{n(1)}, \dots, s_i^{n(d+1)}\}$, сходящихся к ней, и выделяется набор точек

$\left((s_i^{(l)}, s_{-i}^*), v_i^{(l)} \right), i = 1, \dots, N, l = 1, \dots, d$. Если в наборе существует точка, удовлетворяющая BSA-условию, то по нему $\exists s'_i \in \tilde{Q}_i(s_{-i}^*) : u_i(s'_i, s_{-i}^*) > v_i^{(l)} \geq \hat{v}_i^*$. В соответствии с (8) и условием квазивогнутости следствия $\hat{u}'_i(s'_i, s_{-i}^*) = u_i(s'_i, s_{-i}^*)$, $s'_i \in \tilde{Q}_i(s_{-i}^*)$, получается условие WRUSC для точки $(s^*, \hat{v}^*)$: $\hat{v}_i(s'_i, s_{-i}^*) \geq \hat{u}_i(s'_i, s_{-i}^*) = \hat{u}'_i(s'_i, s_{-i}^*) > \hat{v}_i^*$. Если ни одна точка набора не удовлетворяет BSA-условию, то соответствующая часть доказательства идентична теореме.

(5). *Доказательство, что равновесие Нэша игры $\hat{G}$ является безопасным профилем.* В доказательстве теоремы при предположении, что профиль s^* (равновесие Нэша) не является безопасным, получено противоречие. При этом доказательство проведено для профиля $s^* \notin \times_{i=1}^N \tilde{Q}_i(s_{-i}^*)$, т.е. к данным профилям принадлежат не только небезопасные, но также и любые профили, не принадлежащие BSA-подмножеству, что и требуется для доказательства следствия.

(6). *Доказательство существования РБС в исходной игре.* Так как в предыдущем пункте уже доказано, что профиль s^* принадлежит BSA-подмножеству, то это значит, что и на этом более узком множестве по условию (следствия) квазивогнутости на BSA-подмножествах выполняется $u_i(s^*) = \hat{u}'_i(s^*)$, которое требуется для доказательства равенства $v_i(s^*) = \hat{v}_i(s^*)$. Дальнейшее доказательство идентично доказательству теоремы. $\square$

Ослабим условия существования ещё на один шаг и покажем, что для существования РБС будет достаточно, чтобы условия теоремы 7.1 выполнялись на более узком множестве стратегий. Рассмотрим компактную область профилей игры в некотором прямоугольном множестве $s \in B = \times_{i=1}^N B_i$, где $B_i \subseteq S_i$ – непустое замкнутое выпуклое множество стратегий игрока i . Переопределим условия теоремы 7.1, если их брать только по отношению к множеству B .

Определение 7.1'. Игра $G(S_i, u_i)_{i=1}^N$ является **непрерывной во внутренних безопасных стратегиях в B** (continuous at inte-

rior secure strategies in B), если функция выигрыша $u_i(s)$ каждого игрока i непрерывна в множестве $\{s: s \in Q^{(i)} \setminus \partial Q^{(i)}, s_{-i} \in B_{-i}\}$.

Определение 7.2'. Игра $G(S_i, u_i)_{i=1}^N$ **гарантирует выигрыш на границе безопасных стратегий в B (payoff secure at boundary secure strategies in B)**, если для каждого игрока i для каждого $s \in Q^{(i)} \cap \partial Q^{(i)} \cap B$ и для каждого $\varepsilon > 0$ существует $s'_i \in B_i$ такое, что $v_i(s'_i, s'_{-i}) \geq u_i(s) - \varepsilon$ для всех s'_{-i} в некоторой окрестности s_{-i} в B_{-i} .

Определение 7.5'. Игра $G(S_i, u_i)_{i=1}^N$ **предоставляет лучшую безопасную альтернативу в выпуклом множестве в B (provides better secure alternative in convex set in B)**, если (i) каждый игрок i при заданных стратегиях других игроков $s_{-i} \in B_{-i}$ имеет лучшую безопасную альтернативу в некотором непустом множестве $\tilde{Q}_i(s_{-i}) \subseteq Q_i(s_{-i}) \cap B_i$, и (ii) для всех $s \notin \times_{k=1}^N \tilde{Q}_k(s_{-k})$ таких, что $s_{-i} \in B_{-i}$ для некоторого i , существует такой игрок j , что $s_j \notin \overline{co(\tilde{Q}_j(s_{-j}))}$.

Определение 7.6'. BSA-Игра $G(S_i, u_i)_{i=1}^N$ называется **квазивогнутой на BSA-множествах в B** , если функция выигрыша каждого игрока i совпадает со своей квазивогнутой оболочкой на BSA-множествах при $s_{-i} \in B_{-i}$, т.е. $\forall i, \forall s: s_i \in \tilde{Q}_i(s_{-i}), s_{-i} \in B_{-i}$ выполняется $u_i(s) = \hat{u}'_i(s)$, где $\hat{u}'_i(s_i, s_{-i}) = \inf\{h(s_i): co(\tilde{Q}_i(s_{-i})) \rightarrow \mathbb{R}: \text{квазивогнутая}, u_i(s_i, s_{-i}) \leq h(s_i), s_i \in \tilde{Q}_i(s_{-i})\}$.

Тогда теорему 7.1 можно переформулировать следующим образом.

Теорема 7.1'. Если существует такое множество стратегий $B = \times_{i=1}^N B_i$, где $B_i \subseteq S_i$ – непустые замкнутые выпуклые множества, что игра $G(S_i, u_i)_{i=1}^N$ компактна, непрерывна во внутренних безопасных стратегиях в B , гарантирует выигрыш в граничных безопасных стратегиях в B , предоставляет лучшую безопасную альтернативу в выпуклом множестве в B и квазивогнута на BSA-множествах в B , то в игре G существует РБС $s^* \in B$.

Доказательство. Буквально повторяет доказательство теоремы 7.1, за исключением того, что в качестве модифицированной игры следует рассматривать $\hat{\hat{G}}_B = (B_i, \hat{v}_i)_{i=1}^N$, где $\hat{v}_i(s_i, s_{-i})$ – квазивогнутая оболочка функции $v_i(s_i, s_{-i})$ на множестве $s_i \in S_i$. Также во всех соответствующих местах доказательства следует выбирать отклонения $s'_i \in Q_i(s_{-i}) \cap B_i$ вместо $s'_i \in Q_i(s_{-i})$. $\square$

Чтобы использовать теоремы 7.1 и 7.1', необходимо исследовать структуру безопасных стратегий игроков. Такой шаг кажется неизбежным при поиске РБС. Тем не менее в некоторых хорошо известных играх множества безопасных стратегий устроены достаточно просто. Показано на задачах, что предложенные теоремы хорошо работают. Соответствующий материал предполагается опубликовать в отдельной статье.

8. Большая теорема. Второй вариант

Второй подход к построению большой теоремы, как альтернативы первому. При этом хотелось бы избежать недостатков теоремы 7.1. Первый, самый очевидный из них, – её сложность и неочевидность конструкций как в системе требующихся определений и понятий, так и при доказательстве результата.

Более формальная претензия к ней заключается в том, что она использует значения функции безопасного выигрыша за пределами множеств безопасных стратегий игроков и даже их границ. Для применения теоремы 7.1 требуются значения этих функций в малых окрестностях особых точек границы множеств безопасности, где безопасные выигрыши должны быть непрерывны по окружению, т.е. угрозы в них не должны появляться скачком. А хотелось бы, чтобы для применения теоремы было достаточно условия сильных угроз или его небольшого усиления и значений выигрышей игроков строго в рамках их безопасного множества. И даже если за рамками безопасных множеств есть точки недосягаемого предела функций безопасного выигрыша, равного максимуму выигрыша на безопасном множестве, то всё равно при наличии равновесия Нэша в безопасной области (= РБС)

и условия сильных угроз такие особые точки не будут препятствием для существования этого равновесия. Все эти пределы будут «убиваться» условием сильных угроз.

Основная идея второго варианта большой теоремы заключается в совмещении подходов теоремы Рени ($\Rightarrow$ теорема 6.1) и следствия Рени ($\Rightarrow$ теорема 6.2). Условия следствия просты и легко проверяемы, но они намного жёстче условий теоремы Рени и не всегда выполняются. Условия теоремы не так очевидны и применять их сложнее, а если сложность проверки условий теоремы существования и сложность прямой непосредственной проверки её вывода сравнимы, то такая теорема не очень полезна.

Казалось бы, объединяющий подход тавтологичен: зачем объединять в одном утверждении две формулировки, одна из которых является следствием другой? Но на практике было бы удобно иметь такой способ проверки, при котором можно было бы сначала убедиться, что более слабые условия выполняются почти везде, кроме некоторых особых точек, и только для этих особых отдельных случаев уже необходимо было бы проверять более сложное условие.

Условия и теоремы, и следствия Рени сформулированы как свойства игры (функций выигрыша) в целом. Чтобы применить комбинацию этих условий, достаточно показать, что их можно применять и к отдельным профилям, и что для отдельных профилей условие теоремы так же вытекает из условия следствия, как и для всей игры.

В основу теоремы также необходимо включить обобщённое условие квазивогнутости. В теореме 7.1, чтобы обеспечить его выполнение, в определении лучшей безопасной альтернативы было включено условие «в выпуклом множестве». Оно формулировалось так: « $\forall s \notin \times_{i=1}^N Q_i(s_{-i}), \exists j: s_j \notin co(Q_j(s_{-j}))$ ». Теперь необходимо извлечь это свойство из определения лучшей безопасной альтернативы и сформулировать более удобным образом, как свойство совокупности множеств безопасности игроков. Это можно сделать в следующей лемме.

Лемма 8.1. Пусть дана игра $G(S_i, u_i)_{i=1}^N$ с множествами безопасных стратегий $Q_i(s_{-i})$, тогда два условия являются эквивалентными:

$$\begin{aligned} & - \forall s \notin \times_{i=1}^N Q_i(s_{-i}), \exists j: s_j \notin \overline{co(Q_j(s_{-j}))}; \\ & - \{s: s_i \in Q_i(s_{-i}), \forall i\} = \{s: s_i \in \overline{co(Q_i(s_{-i}))}, \forall i\}. \end{aligned}$$

Доказательство. Рассмотрим любой профиль s^* .

$$\begin{aligned} s^* \notin \times_{i=1}^N Q_i(s_{-i}) & \Leftrightarrow s^* \notin \{s: s_i \in Q_i(s_{-i}), \forall i\}. \\ \exists j: s_j \notin \overline{co(Q_j(s_{-j}))} & \Leftrightarrow s^* \notin \{s: s_i \in \overline{co(Q_i(s_{-i}))}, \forall i\}. \end{aligned}$$

Значит, первое условие леммы можно переписать следующим образом.

$$s^* \notin \{s: s_i \in Q_i(s_{-i}), \forall i\} \Rightarrow \{s: s_i \in \overline{co(Q_i(s_{-i}))}, \forall i\}.$$

Поскольку множество безопасных профилей $\{s: s_i \in Q_i(s_{-i}), \forall i\} \subseteq \{s: s_i \in \overline{co(Q_i(s_{-i}))}, \forall i\}$, то первое условие будет эквивалентно следующему:

$$s^* \in \{s: s_i \in Q_i(s_{-i}), \forall i\} \Rightarrow s^* \in \{s: s_i \in \overline{co(Q_i(s_{-i}))}, \forall i\}.$$

А это и означает, что второе условие леммы выполнено. $\square$

Доказанная лемма позволяет выделить и определить нужное свойство множеств безопасности игроков.

Определение 8.1. В игре $G(S_i, u_i)_{i=1}^N$ с множествами безопасности игроков $Q_i(s_{-i})$ **множество безопасных профилей нерасширяемо** (выпуклыми оболочками безопасных стратегий), если $\{s: s_i \in Q_i(s_{-i}), \forall i\} = \{s: s_i \in \overline{co(Q_i(s_{-i}))}, \forall i\}$.

Ещё два усиливающих теорему условия – ослабление на подмножества безопасных стратегий $\tilde{Q}_i(s_{-i}) \subseteq Q_i(s_{-i})$ и на подмножества стратегий $B_i \subseteq S_i$ – будут сформулированы как локальный вариант теоремы вместе с уточнёнными для этого случая вариантами определений. Базовый вариант теоремы будет иметь следующий вид.

Теорема 8.1. Игра $G(S_i, u_i)_{i=1}^N$ компактна; квазивогнута на безопасных стратегиях; её множество безопасных профилей нерасширяемо; с сильными угрозами; для любого профиля $s \in S$: либо гарантирует выигрыш на безопасных стратегиях и взаимно-верхнеполунепрерывна в безопасных стратегиях; либо гарантирует лучший ответ в безопасных стратегиях; тогда в игре G существует РБС.

Доказательство. Сведём условия теоремы 8.1 к условиям теоремы 6.1 (построенной на условиях теоремы Рени). Для этого

надо сделать две вещи. Во-первых, доказать, что можно рассматривать условия гарантирования лучшего ответа, взаимной верхней полунепрерывности и гарантирования выигрыша на отдельных профилях, а не только в игре в целом, как доказано Рени в [14], и показать, что и для отдельных профилей из первого и второго следует третье. Для этого надо привести соответствующую часть доказательства из статьи Рени и убедиться, что оно действительно и на отдельных точках. Во-вторых, необходимо свести условия теоремы 8.1 к случаю с выпуклыми множествами безопасности и потом доказать, что найденное для них равновесие попадает в множества безопасности и для новой теоремы.

1. В условии теоремы «либо – либо» одна из альтернатив, взятая в целом для игры, строго сильнее другой. Если выполняется первая, то выполняется и вторая (но первую для задач проверять легче). То есть если мы покажем, что для всех профилей игры верна вторая альтернатива, то эта часть будет доказана. Приведём рассуждения из доказательства утверждения Рени применительно к нашему случаю. Пусть для профиля s^* игра гарантирует выигрыш на безопасных стратегиях и взаимно-верхнеполунепрерывна на безопасных стратегиях. То есть в профиле $s^* \forall i, \forall \varepsilon > 0$: игрок i может гарантировать себе выигрыш $u_i(s^*) - \varepsilon$ в s^* . Если s^* не является равновесием Нэша, то рассмотрим $\forall (s^*, u^*) \in \overline{\Gamma(\bar{G})}$. Тогда либо $\exists i: u_i(s^*) > u_i^*$, либо $\forall i: u_i(s^*) = u_i^*$. В первом случае, согласно взаимной верхней полунепрерывности, берём $s'_i = s_i^*$. Во втором случае, согласно той же взаимной верхней полунепрерывности и потому, что s_i^* не является равновесием Нэша, $\exists i, \exists s'_i: u_i(s'_i, s_{-i}^*) > u_i^*$. В обоих случаях, согласно свойству гарантирования выигрыша, игрок i может гарантировать себе в (s'_i, s_{-i}^*) , а значит, и в s^* , выигрыш $u_i(s'_i, s_{-i}^*) > u_i^*$. Свойство гарантирования лучшего ответа для s^* доказано.

Итак, для любого профиля из более сильного условия теоремы вытекает более слабое, а этого более слабого достаточно для выполнения утверждения теоремы. Более слабое условие, таким образом, выполняется для всех профилей игры и, соответственно, для игры в целом.

2. Обобщённая казивогнутость. В теореме 6.1 результат был доказан для случая выпуклых множеств безопасных стратегий.

Рассмотрим игру, обрезанную на выпуклых замыканиях безопасных множеств: $\bar{G}(S_i, \bar{u}_i), \bar{u}_i(s) = \begin{cases} \hat{u}_i(s), s_i \in \overline{co(Q_i(s_{-i}))}, \\ C_{\min}, s_i \notin \overline{co(Q_i(s_{-i}))}, \end{cases}$ где $\hat{u}_i(s)$ – квазивогнутая оболочка выигрыша $u_i(s)$ игрока i на множестве $\overline{co(Q_i(s_{-i}))}$. Для этой игры теорема 6.1 верна, т.е. в этой игре существует равновесие Нэша s^* . Теперь остаётся доказать, что это равновесие Нэша попадает в множества безопасных стратегий всех игроков: $s_i^* \in Q_i(s_{-i}^*), \forall i$. Это следует из того, что множество безопасных стратегий расширяемо: $\{s: s_i \in \overline{co(Q_i(s_{-i}))}, \forall i\} = \{s: s_i \in Q_i(s_{-i}), \forall i\}$. Действительно, если допустим, что $\exists i: s_i^* \notin \overline{co(Q_i(s_{-i}^*))}$, то, по лемме 8.1 найдётся игрок j , для которого $s_j^* \in \overline{co(Q_j(s_{-j}^*))}$. Для этого игрока $\bar{u}_j(s^*) = C_{\min}$ и, по условию сильных угроз, $\exists j: \bar{u}_j(s_j', s_{-j}^*) > \bar{u}_j(s^*)$. Получено противоречие.

Таким образом, все отличия теоремы 8.1 от условий теоремы 6.1 устранены и новое утверждение доказано сведением к уже доказанному предыдущему. $\square$

Локальный вариант теоремы 8.1. Практически работающий вариант теоремы должен быть локальным, так как глобально, на всём множестве S_i , требования теоремы не выполняются почти нигде, ни на одной из рассмотренных задач. Для того чтобы получить его, надо внести в имеющуюся формулировку два изменения, две поправки в множества, на которых ищутся равновесные точки. Совершить два перехода. Во-первых, перейти от множества безопасных стратегий к некоторому его подмножеству: $Q_i(s_{-i}) \subseteq S_i \rightarrow \tilde{Q}_i(s_{-i}) \subseteq Q_i(s_{-i})$. Во-вторых, для окружения рассматривать не всё множество стратегий, а только его подмножество: $S_{-i} \rightarrow B_{-i} \subseteq S_{-i}$.

Локальный вариант условия в прямоугольном подмножестве $s \in B = \times_{i=1}^N B_i, B_i \subseteq S_i$ был введён в определении 2.10 как *условие сильных угроз в B* , и на его основе сформулирована теорема 2.3 как локальный вариант теоремы 2.2. Локальный вариант двух малых теорем 6.1 и 6.2 по теореме и следствию Рени был дан как теоремы 6.1' и 6.2'. Для первого варианта большой теоремы случай подмножества безопасных стратегий был оформлен через определения 7.4, 7.5 *лучшей безопасной альтернативы*

на BSA множествах для базового варианта теоремы и уточнён через определение 7.6 квазивогнутости на BSA множествах как следствие этой теоремы. Окончательный локальный вариант этой большой теоремы потребовал уточнения определений 3.1', 3.2', 3.5', 3.6' и был представлен как теорема 7.1'.

Аналогичным образом соберём и перепишем для случая «на безопасном подмножестве $\tilde{Q}$ в множестве B » все определения, прямо или косвенно требуемые условиями теоремы 8.1. Собираем в единый блок все упоминаемые там понятия, определения: 2.8 сильной угрозы (для профиля, игрока, игры), 6.2 гарантированного безопасного ответа (для окружения, игрока), 6.4 гарантированного лучшего ответа в безопасных стратегиях (для стратегии, = для профиля & игрока), 6.5 гарантированного лучшего ответа в безопасных стратегиях (для игры), 6.6 гарантированного выигрыша в безопасных стратегиях (для стратегии), 6.7 гарантированного выигрыша в безопасных стратегиях (для игры), 6.8 взаимной верхней полунепрерывности в безопасных стратегиях (для игры), 7.3 квазивогнутости в безопасных стратегиях (для игры), 8.1 нерасширяемости множества безопасных профилей (для игры).

Итак, пусть для игры $G(S_i, u_i)_{i=1}^N$ заданы прямоугольные подмножества стратегий $B = \times_{i=1}^N B_i$, $B_i \subseteq S_i$ и подмножества безопасных стратегий $\tilde{Q}_i(s_{-i}) \subseteq Q_i(s_{-i}) \cap B_i \subseteq S_i$, $\forall i, \forall s_{-i} \in B_{-i}$, а также соответствующие подмножества безопасных профилей $\tilde{Q}^{(i)} = \{s: s_i \in \tilde{Q}_i(s_{-i}), \forall i\}$.

Определение 2.8'. Угроза игроку i в профиле $s \in (S_i, B_{-i})$ является **сильной по отношению к подмножеству безопасных стратегий $\tilde{Q}$** , если существует безопасная стратегия $s' = (s'_i, s_{-i})$, $s_i \in \tilde{Q}_i(s_{-i})$ такая, что $u(s') = v(s') > v(s)$. Для игрока i в игре G выполняется **условие сильных угроз относительно подмножества безопасных стратегий $\tilde{Q}$ в B** , если для него содержащиеся в любом профиле $s \in (S_i, B_{-i})$, не содержащимися в подмножестве безопасных стратегий, угрозы являются сильными по отношению к подмножеству безопасных стратегий $\tilde{Q}$. Игра G называется **игрой с сильными угрозами относительно подмножества безопасных стратегий $\tilde{Q}$ в B** , если для

всех игроков в ней выполняется условие сильных угроз относительно подмножества безопасных стратегий $\tilde{Q}$ в B .

Определение 6.2'. Игрок i имеет гарантированный безопасный ответ на подмножестве безопасных стратегий $\tilde{Q}$ в B при окружении $s_{-i} \in B_{-i}$, если $\exists s'_i \in B_i, \exists$ открытая окрестность $U_{s_{-i}} \subseteq B_{-i}, \forall s'_{-i} \in U_{s_{-i}}: s' = (s'_i, s'_{-i}) \in \tilde{Q}^{(i)}$. Игрок i имеет гарантированный безопасный ответ на подмножестве безопасных стратегий $\tilde{Q}$ в B в игре G , если он его имеет при любых окружениях.

Определение 6.4'. Безопасная стратегия s_i игрока i в профиле $s \in (S_i, B_{-i})$, не являющемся РБС, имеет гарантированный лучший ответ на подмножестве безопасных стратегий $\tilde{Q}$ в B , если $\forall (s, u) \in \overline{\Gamma(\bar{u})}, \exists \alpha > u_i, \exists s'_i \in \tilde{Q}_i(s_{-i}), \exists$ открытая окрестность $U_{s_{-i}} \subseteq B_{-i}: \forall s'_{-i} \in U_{s_{-i}}: s' = (s'_i, s'_{-i}) \in \tilde{Q}^{(i)}$ (безопасен для i) и $u_i(s') \geq \alpha > u_i$.

Определение 6.5'. Игра $G(S_i, u_i)_{i=1}^N$ гарантирует лучший ответ на подмножестве безопасных стратегий $\tilde{Q}$ в B , если $\forall s \in B \exists$ игрок i такой что: он имеет гарантированный безопасный ответ на подмножестве безопасных стратегий $\tilde{Q}$ в B , если его стратегия s_i не безопасна, и гарантированный лучший ответ на подмножестве безопасных стратегий $\tilde{Q}$ в B , если его стратегия s_i безопасна.

Определение 6.6'. Безопасная стратегия s_i игрока i в профиле $s \in (S_i, B_{-i})$ имеет гарантированный выигрыш на подмножестве безопасных стратегий $\tilde{Q}$ в B , если $\forall \varepsilon > 0, \exists s'_i \in \tilde{Q}_i(s_{-i}), \exists$ открытая окрестность $U_{s_{-i}} \subseteq B_{-i}: \forall s'_{-i} \in U_{s_{-i}}: s' = (s'_i, s'_{-i}) \in \tilde{Q}^{(i)}$ (безопасен для i) и $u_i(s') \geq u_i(s) - \varepsilon$.

Определение 6.7'. Игра $G(S_i, u_i)_{i=1}^N$ гарантирует выигрыш на подмножестве безопасных стратегий $\tilde{Q}$ в B , если для $\forall s \in B, \forall$ игрока i : имеется гарантированный безопасный ответ на подмножестве безопасных стратегий $\tilde{Q}$ в B , если его стратегия s_i не безопасна, и гарантированный выигрыш на подмножестве безопасных стратегий $\tilde{Q}$ в B , если его стратегия s_i безопасна.

Определение 6.8'. Игра $G(S_i, u_i)_{i=1}^N$ взаимно-верхнеполунепрерывна в подмножестве безопасных стратегий (в.в.п.н.п.б.с.)

в профиле $s \in B$, если для соответствующей этой игре обрезанной на подмножестве безопасных стратегий игры $\bar{G}(S_i, \bar{u}_i)_{i=1}^N$ для $\forall (s, u) \in \bar{\Gamma}(\bar{u})$ выполняется условие: $u_i(s) \leq u_i, \forall i \Rightarrow u_i(s) = u_i, \forall i$. Игра $G(S_i, u_i)_{i=1}^N$ **в.в.п.н.п.б.с. в B** , если она в.в.п.н.п.б.с. любом профиле $s \in B$.

Определение 7.3'. Игра $G(S_i, u_i)_{i=1}^N$ **квазивогнута на подмножестве безопасных стратегий $\tilde{Q}$ в B** , если функция выигрыша каждого игрока i совпадает с ее квазивогнутой оболочкой на подмножестве безопасных стратегий этого игрока, т.е. $\forall i, \forall s: s_i \in \tilde{Q}_i(s_{-i})$ выполняется $u_i(s) = \hat{u}_i(s)$.

Определение 8.1'. В игре $G(S_i, u_i)_{i=1}^N$ с подмножествами безопасности игроков $\tilde{Q}_i(s_{-i})$ **подмножество безопасных профилей нерасширяемо** (выпуклыми оболочками безопасных стратегий), если:

$$\{s: s_i \in \tilde{Q}_i(s_{-i}), \forall i\} = \overline{\{s: s_i \in co(\tilde{Q}_i(s_{-i})), \forall i\}}.$$

Теперь есть всё необходимое, чтобы сформулировать локальный вариант теоремы 8.1, т.е. её окончательный вариант, хорошо работающий на прикладных задачах.

Теорема 8.1'. *Игра $G(S_i, u_i)_{i=1}^N$: компактна; квазивогнута на подмножестве безопасных стратегий $\tilde{Q}$ в B ; её подмножество безопасных профилей нерасширяемо; с сильными угрозами относительно подмножества безопасных стратегий $\tilde{Q}$ в B ; для любого профиля $s \in S$: либо гарантирует выигрыши на подмножестве безопасных стратегий $\tilde{Q}$ в B и взаимно-верхнеполунепрерывна на подмножестве безопасных стратегий $\tilde{Q}$ в B ; либо гарантирует лучший ответ на подмножестве безопасных стратегий $\tilde{Q}$ в B ; – тогда в игре G существует РБС.*

Комментарий. Утверждение теоремы остаётся корректным при переходе от множества безопасных стратегий к их подмножеству, – это следует из того, что при переходе от определения 2.8 к 2.8' это более узкое множество оказывается в точно таком же отношении ко всем не входящим в него профилям, как более широкое множество в прежнем глобальном варианте теоремы. Все последующие понятия и рассуждения строятся на конструкции обрезанной функции, которая изначально намеренно была

сформулирована безотносительно к тому, что отсекаемые ей значения функции выигрыша имеют именно небезопасную природу. Поэтому все рассуждения сохраняют свою силу и при переносе на более узкое множество при условии, что выполняется определение 2.8'.

Доказательство. Рассмотрим игру $\hat{G}'_B(B_i, \hat{u}'_i(s))_{i=1}^N$, где $\hat{u}'_i(s)$ – функция выигрыша обрезанная на BSA-множестве $\tilde{Q}_i(s_{-i})$. Эта игра компактна, квазивогнута (так как исходная игра G квазивогнута на BSA-множестве). По условию теоремы исходная игра G : либо гарантирует выигрыш на подмножестве безопасных стратегий $\tilde{Q}$ в B и взаимно-верхнеполунепрерывна на подмножестве безопасных стратегий $\tilde{Q}$ в B ; либо гарантирует лучший ответ на подмножестве безопасных стратегий $\tilde{Q}$ в B . Из этого следует, что игра $\hat{G}'_B$ гарантирует лучший ответ.

Игра $\hat{G}'_B$ компактна, квазивогнута, гарантирует лучший ответ; следовательно, по теореме 3.1. (Рени): $\exists s^* \in B$ – равновесие Нэша этой игры.

Рассмотрим профиль s^* в исходной игре, произвольное отклонение s'_i произвольного игрока i от него, профиль (s'_i, s^*_{-i}) . Рассмотрим безопасный выигрыш в этом профиле $v_i(s'_i, s^*_{-i})$. Так как s^* – равновесие Нэша в игре $\hat{G}'_B$ и по условию теоремы игра G – с сильными угрозами относительно подмножества безопасных стратегий $\tilde{Q}$ в B , то из этого следует, что $v_i(s'_i, s^*_{-i}) \leq v_i(s^*)$. То есть s^* – равновесие Нэша игры угроз игра $\tilde{G}(S_i, v_i)_{i=1}^N$. При этом если $s'_i \notin \tilde{Q}_i(s^*_{-i})$, то $v_i(s'_i, s^*_{-i}) < v_i(s^*)$. Профиль s^* безопасен. Из этого, по теореме 2.1, следует, что s^* является РБС исходной игры G . $\square$

В заключение рассмотрения теоремы 8.1' можно добавить, что хотя её формулировка выглядит очень тяжеловесно, но тем не менее она весьма полезна. Дело в том, что локальная теорема требует выполнения намного меньших условий, чем теорема 8.1. А именно, условие сильных угроз должно выполняться на множествах $s \in (S_i, B_{-i}) \quad \forall i$, а для всех других условий теоремы достаточно выполнения на локальном множестве B . Чего и хотелось изначально добиться. Это позволяет рассматривать игры на маленьком подмножестве, имеющем внутри себя хорошие свойства

и обладающем преимуществом сильных угроз при любых односторонних отклонениях за его пределы. За этими границами игра может вести себя как угодно плохо (ну разве что компактность игры там требуется), но если в игровой задаче удастся найти такое удобное подмножество, которое можно интерпретировать как область неантагонистичности игры, то к ней можно применять теорему и делать вывод о существовании в игре РБС.

9. Заключение

Целью статьи является продемонстрировать метод конструирования теорем существования РБС из имеющихся в литературе теорем существования равновесий Нэша (в чистых стратегиях). Приведен весь ряд необходимых утверждений – от метатеоремы и исходных теорем из источников до рабочих теорем, которые могут применяться к задачам. Чтобы объединить и привести в одной публикации всю цепочку теоретических результатов, потребовался достаточно большой объем статьи. Тестирование рабочих теорем на задачах планируется вынести в отдельную публикацию.

Общая схема получения рабочих теорем следующая. В разделе 2 приведены теоремы, опубликованные ранее в статьях [3, 4]. В теореме 2.1 представлено необходимое и достаточное условие существования РБС, которое проясняет необходимость использования игры угроз с безопасным выигрышем, взятом в качестве целевой функции. Метатеорема 2.1 представляет центральный теоретический результат, положенный в основу метода. В теореме 2.3 представлен локальный вариант теоремы, который является рабочим, т.е. достаточен для применения к задачам.

Существенным новым материалом, опубликованным в данном разделе, является описание метода приведения исходной теоремы (который в источниках может быть сформулирован различными способами) к стандартному виду, что позволило построить предлагаемый метод в форме метатеоремы. Все различия исходных теорем были свернуты внутри переменной части, обозначенной как (#). После этого все соответствующие конкретные формулировки, приведенные к стандартному виду, могут далее

вставляться в гнездо внутри метатеоремы. Применение метода, приведение исходной теоремы к стандартному виду и получение из него конкретной теоремы существования РБС было пояснено примером 2.1.

В разделе 3 приведены две исходные теоремы из работы [14], теорема 3.1 и следствие 3.1. Они изложены нарочито подробно, с обильным цитированием источника, что необходимо для иллюстрации метода.

В разделе 4 приводится подробная интерпретация условий теорем Рени, в сравнении с так же подробно приведенным случаем теоремы Дебре (теорема 4.1), приведенным к стандартному виду (теорема 4.2), и полученной теоремой существования РБС (теорема 4.4). Эта теорема была первоначально сформулирована и опубликована в статье [12]. Далее в завершение раздела 4 и разделе 5 приводится подробный анализ теоремы Рени, на ряде примеров, условие теоремы трактуется как условие отсутствия точек перескока или точек, гарантирующих лучший ответ.

В разделе 6 строятся формально, по методу метатеоремы два критерия существования РБС по исходной теореме Рени (теорема 6.1) и по следствию Рени (теорема 6.2). По ним сформированы локальные варианты теорем 6.1' и 6.2'.

Первоначально предполагалось этим и ограничиться. Но при попытках применить эти теоремы к прикладным задачам выяснились многочисленные **трудности и проблемы**. Оказалось, что требуемые построенными формально теоремами условия для задач (Хотеллинга и Таллока – Скапердаса) не выполняются, ни одно из них (!).

1. **Квазивогнутость** (на безопасных стратегиях) не выполняется даже для самой простой из моделей – задачи Хотеллинга, так как не выполняется условие выпуклости для множеств безопасных стратегий: они часто даже не являются связными. Множество безопасных профилей в задаче Хотеллинга не выпукло за счёт безопасных профилей с нулевым выигрышем.

2. **Гарантированность выигрыша** также не выполняется, например для игры цен в задаче Хотеллинга при определённых условиях: при сильно несимметричном расположении игроков на отрезке и случая, когда максимум безопасных стратегий одного

игрока лежит на границе безопасных стратегий, а другого – на максимуме целевой функции.

3. **Компактность** множества безопасных профилей и стратегий не выполняется почти на всех задачах за счёт тех областей безопасных профилей, которые не содержат свою границу.

4. **Верхняя полунепрерывность** также не выполняется, так как встречаются точки недостижимых пределов, принадлежащие замыканию графика целевой функции, но не принадлежащие самому графику.

Таким образом, чтобы довести теоремы до применимости, пришлось провести их существенную доработку. При этом ставилась цель, чтобы новая усложнённая теорема была применима ко всем тестовым задачам. Были внесены следующие изменения.

А. Было введено *подмножество множества безопасных стратегий* $\tilde{Q}_i(s_{-i})$ вместо $Q_i(s_{-i})$. Это минимальное, самое простое изменение можно было ввести чисто формально.

Б. Были введены (для теоремы 7.1) дополнительные условия и определения *непрерывности во внутренних безопасных стратегиях* и *гарантирования выигрышей на границе безопасных стратегий* для корректной работы с граничными точками множеств безопасных стратегий.

В. Вместо *обрезанной функции* выигрыша была использована *функция безопасного выигрыша*, которая вела себя непрерывно, по крайней мере на рассмотренных задачах, т.е. *угрозы* в окрестности особой краевой точки, где нарушалось условие гарантированного выигрыша, *были малы* (это не обязательно должно быть так).

Г. В некоторых случаях альтернативой функции безопасного выигрыша (для несимметричной задачи Хотеллинга) оказалось использование в качестве исходной теоремы не *следствия Рени*, как это фактически общепринято, а *исходной теоремы Рени*. То есть соответствующий аналог теоремы 3.1 сильнее следствия 3.1.

Д. Очень много пришлось поработать над условием выпуклости множеств безопасных стратегий, заменив исходные *неквазивогнутые функции* их минимальными *квазивогнутыми оболочками*, введя сложное условие на выпуклость внутрь определения лучшей гарантированной альтернативы (*лучшая безопасная альтернатива в выпуклом множестве*).

Е. В качестве альтернативного подхода (для теоремы 8.1) был применён *комбинированный вариант* одновременного применения условий *следствия Рени* (для почти всех точек, где он выполнялся) и более сложных условия *теоремы Рени* (там, где более простые и одновременно более жёсткие условия не проходили).

При реализации этого плана и сами применяемые конструкции, и доказательство результатов существенно усложнились, тем не менее удалось получить две работающие теоремы. Если попытаться сравнить их, то однозначного преимущества ни одной из них не видно, у каждой есть и достоинства, и недостатки: в теореме 7.1 приходится обращаться к значениям функций за пределами безопасных множеств; в теореме 8.1 приходится применять тяжёлые для проверки условия теоремы Рени к отдельным профилям, где не проходят условия следствия Рени, но по крайней мере только для отдельных плохих точек. В итоге получено две мощных теоремы с хорошей применимостью. По мощности они одинаково проходят этот тест на прикладных задачах (в локальном варианте). Но получить идеал всё-таки не удалось – получено два компромиссных, как бы дополнительных друг по отношению к другу варианта. Итак, в разделах 7 и 8 приведены формулировки и доказательства теорем, модифицированные под требования задач (7.1 и 8.1), а также их локальные (рабочие) варианты (7.1' и 8.1'). Тестирование полученных теорем на задачах будет рассмотрено в отдельной статье. Все предложенные в статье теоремы можно свести в таблицу:

	Мета-теорема	Дебре	Рени, теорема	Рени, следствие
Исходные теоремы	2.2	4.1, 4.2	3.1	Следствие 3.1
Теоремы сущ. РБС (формальн.)	2.3	4.3	6.1	6.2
Локальные варианты			6.1'	6.2'
Модиф. под задачи вариант			7.1, следствие 7.1	8.1
Рабочий вариант	2.3	4.3	7.1'	8.1'

Литература

1. ИСКАКОВ М.Б. *Равновесие в безопасных стратегиях* // Управление большими системами. – 2004. – Вып. 9: «Лаборатория активных систем: 30 лет». – С. 145–157.
2. ИСКАКОВ М.Б. *Равновесие в безопасных стратегиях* // Автоматика и телемеханика. – 2005. – №3. – С. 139–153. (ISKAKOV M.B. *Equilibrium in Safe Strategies* // Automation and Remote Control. – 2005. – Vol. 66, No.3. – P. 465-478).
3. ИСКАКОВ М.Б. *Теоремы существования равновесия Нэша и равновесия в безопасных стратегиях* // Журнал новой экономической ассоциации. – 2022. – №4(56). – С. 12–27.
4. ИСКАКОВ М.Б., ИСКАКОВ А.Б. *Равновесие в безопасных стратегиях как развитие концепции равновесия Нэша* // Математическая теория игр и ее приложения. – 2023. – Т. 15, Вып. 1. – С. 48–72.
5. D'ASPREMONT C., GABSZEWICZ J., THISSE J.-F. *On Hotelling's "Stability in Competition"* // Econometrica. – 1979. – 47(5). – P. 1145–1150.
6. BERTRAND J. *Review of Cournot's 'Rechercher sur la theoric mathematique de la richesse'* // Journal des Savants. – 1883. – P. 499-508.
7. BAGH A., JOFRE A. *Reciprocal upper semicontinuity and better reply secure games: a comment* // Econometrica. – 2006. – Vol. 74. – P. 1715–1721.
8. BICH P. *Existence of pure Nash equilibria in discontinuous and non quasiconcave games* // Int. J. Game Theory. – 2009. – Vol. 38. – P. 395–410.
9. DASGUPTA P., MASKIN E. *The existence of equilibrium in discontinuous economic games, I: Theory* // Rev. Econ. Stud. – 1986. – Vol. 53(1). – P. 1–26.
10. DEBREU G. *A social equilibrium existence theorem* // Proc. Natl. Acad. Sci. USA. – 1952. – Vol. 38(10). – P. 886–893.

11. EDGEWORTH F.M. *Papers relating to Political Economy I*. – London: Macmillan, 1925.
12. ISKAKOV M., ISKAKOV A., D'ASPREMONT C. *Games for cautious players: the Equilibrium in Secure Strategies* // Games and Economic Behavior. – July 2018. – Vol. 110. – P. 58–70.
13. HOTELLING H. *Stability in Competition* // Econ. J. – 1929. – Vol. 39(153). – P. 41–57.
14. RENY P.J. *On the existence of pure and mixed strategy Nash equilibria in discontinuous games* // Econometrica. – 1999. – Vol. 67(5). – P. 1029–1056.
15. ROTHSCCHILD M., STIGLITZ J.E. *Equilibrium in competitive insurance markets: an essay on the economics of imperfect information* // Q. J. of Econ. – 1976. – Vol. 90. – P. 629–649.
16. SIMON L. *Games with Discontinuous Payoffs* // Review of Economic Studies. – 1987. – Vol. 54. – P. 569–597.
17. SKAPERDAS S. *Contest success functions* // Economic Theory. – 1994. – Vol. 7. – P. 283–290.
18. TULLOCK G. *The welfare costs of tariffs, monopoly and theft* // Western Econ. J. – 1967. – Vol. 5. – P. 224–232.
19. TULLOCK G. *Efficient rent seeking* // BUCHANAN J.M., TOLLISON R.D., TULLOCK G. E. *Toward a theory of the rent-seeking society* – College Station, TX: Texas A&M University Press, 1980. – P. 97–112.
20. WILSON C. *A model of insurance markets with incomplete information* // J. Econ. Theory. – 1977. – Vol. 16. – P. 167–207.

THE META-THEOREM FOR THE EXISTENCE OF EQUILIBRIUM IN SECURE STRATEGIES

Mikhail Iskakov, V.A. Trapeznikov Institute of Control Sciences of RAS, Moscow, Cand.Sc. (mih_iskakov@mail.ru).

Alexey Iskakov, V.A. Trapeznikov Institute of Control Sciences of RAS, Moscow, Cand.Sc. (isk_alex@mail.ru)

Abstract: The paper is a continuation of the cycle of papers in 2018-2023 devoted to the theoretical justification of equilibrium in secure strategies (EinSS) as a concept of solving non-cooperative games in pure strategies. A method for constructing EinSS existence theorems from known Nash equilibrium (NE) existence theorems is presented. In particular, theorems for the existence of Nash equilibria are formulated in a standard form and are inserted as a condition into the meta-theorem for the existence of EinSS. According to this method, two theorems for the existence of EinSS are derived and proven based on the theorem of Reny (1999) on the existence of Nash equilibria. The general scheme of deriving existence theorems is as follows. Section 2 summarizes the theorems published in the author's previous papers. Section 3 presents two original theorems from Reny's paper. Section 4 gives a detailed interpretation of the conditions of Reny's theorems, compared to the conditions of Debre's theorem. Section 5 gives a detailed analysis of Reny's theorem. Using a number of examples, the condition of the theorem is interpreted as the condition that there are no jump points or points that guarantee the best answer. Section 6 constructs formally, by the method of meta-theorem, two existence criteria for EinSS that use the original NE existence theorems. In Sections 7 and 8, two theorems are formulated and proved, which are specifically refined for solving applied problems (Hotelling's spatial competition, Tullock's rent competition, Bertrand – Edgeworth oligopoly). All considered theorems are summarized in a final table.

Keywords: equilibrium in secure strategies, Nash equilibrium, existence theorems, Reny's existence theorem, Hotelling's spatial competition, Tullock – Skaperdas rent-seeking contest.

УДК 519.833.2

ББК 22.18

DOI: 10.25728/ubs.2024.111.1

*Статья представлена к публикации
членом редакционной коллегии Ф.Т. Алескеровым.*

Поступила в редакцию 28.01.2024.

Опубликована 30.09.2024.

ИДЕНТИФИКАЦИЯ КВАДРАТИЧНЫХ КОМПЛЕКСНОЗНАЧНЫХ ДИНАМИЧЕСКИХ ОКРЕСТНОСТНЫХ МОДЕЛЕЙ НА КЛАСТЕРИЗОВАННЫХ ДАННЫХ И БЕЗ КЛАСТЕРИЗАЦИИ

Седых И. А.¹, Макаров К. Н.²

*(ФГБОУ ВО Липецкий государственный технический
университет, Липецк)*

*Окрестностные модели и их модификации применяются для моделирования различных распределенных систем и процессов. В исследовании рассмотрена квадратичная комплекснозначная динамическая окрестностная модель, в которой параметры, входы и состояния являются комплексными числами, и дано ее определение. Модель функционирует в дискретном времени. Показан пример состоящей из трех узлов комплекснозначной динамической окрестностной модели, для которой приведены граф структуры и функции пересчета состояний в общем виде. Рассмотрен также частный случай функций пересчета для квадратичной модели. Приведён алгоритм идентификации комплекснозначной динамической окрестностной модели, параметры которой находятся методом наименьших квадратов. Показан общий вид матриц системы линейных уравнений для нахождения параметров квадратичной модели. Приведены матрицы и выполнена идентификация для рассмотренного примера окрестностной модели. Найдены среднеквадратическая и средняя приведенная ошибки идентификации. В работе рассмотрена также идентификация комплекснозначной динамической окрестностной модели на кластеризованных данных. Кластеризация выполнена с использованием комплексных наборов данных методом *k*-средних. Предложенные алгоритмы идентификации реализованы в виде программы в пакете *Mathcad*, с помощью которой проведено сравнение результатов идентификации квадратичной комплекснозначной динамической окрестностной модели на кластеризованных данных и без кластеризации.*

Ключевые слова: динамическая окрестностная модель, квадратичная модель, комплексные числа, идентификация, кластеризация, метод *k*-средних.

¹ Ирина Александровна Седых, д.т.н., профессор (sedykh-irina@yandex.ru).

² Кирилл Николаевич Макаров, студент (kirik0-1@yandex.ru).

1. Введение

Окрестностный подход, появившийся в 90-е годы XX века, является развитием теории математического моделирования дискретных распределенных систем [7]. Окрестностные модели отличаются гибким заданием связей между элементами с помощью окрестностей и обладают наглядностью представления структуры в виде графа с несколькими видами дуг [6]. Ранее были введены и исследованы статические и динамические, линейные, билинейные и полилинейные, одноуровневые и иерархические окрестностные модели с постоянными и переменными окрестностями [10]. Разработаны методы их параметрической идентификации и смешанного управления. При этом входы, состояния и выходы, а также параметры моделей являлись действительными числами.

В различных сферах деятельности, например, в металлургии, робототехнике, космонавтике, при анализе данных, в искусственном интеллекте, в квантовых вычислениях актуальным является изучение, моделирование, анализ и прогнозирование поведения сложных систем, а также работа с большим объемом данных.

В некоторых случаях данные, поступающие на вход системы, и ее состояния являются комплексными числами, например, в электротехнике [21]. Комплексные данные расширяют область применения, возможности моделирования и анализа систем. Ранее разработаны методы кластеризации данных и идентификации окрестностных моделей для вещественных наборов данных [4]. В работе рассматриваются модификации данных методов для комплекснозначных наборов данных.

В статье рассмотрена квадратичная комплекснозначная динамическая окрестностная модель, для которой разработаны и реализованы алгоритмы идентификации на некластеризованных и кластеризованных комплекснозначных данных. Все расчеты проводились с помощью пакета Mathcad.

2. Комплекснозначная динамическая окрестностная модель

Окрестностные модели обобщают многие дискретные модели, отличаются гибкостью описания структуры связей между узлами системы, что позволяет улучшить управление объектом. Их применяют для моделирования сложных пространственно-распределенных систем. Простейшим классом окрестностных являются линейные модели. Также разработан класс билинейных окрестностных моделей, обобщающий линейные окрестностные и традиционные билинейные дискретные модели. Далее были введены четкие и нечеткие недетерминированные динамические окрестностные модели, которые позволяют моделировать параллельные стохастические процессы.

На примере рассмотрим комплекснозначную динамическую окрестностную модель (КДОМ), состоящую из трех узлов a_1, a_2, a_3 (см. рис. 1), которую можно описать как $CNS = (N, X, V, G, X[0])$, где $N = (A, O)$ – это структура окрестностной модели, $A = \{a_1, \dots, a_n\}$ – множество узлов, O – множество окрестностных связей; $X = X[t]$, $V = V[t] \in \mathbb{C}^n$ – векторы состояний и входов КДОМ в текущий момент времени t ; G – функция пересчета состояний модели; $X[0] \in \mathbb{C}^n$ – вектор состояний в начальный момент времени $t = 0$ [7].

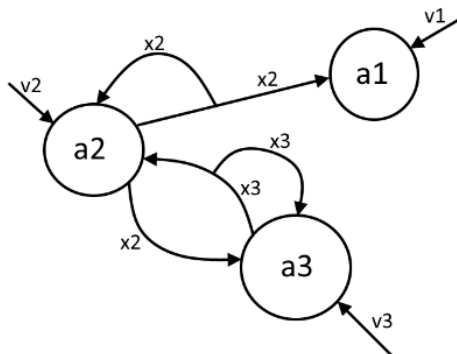


Рис. 1. Граф структуры окрестностной модели, состоящей из трех узлов a_1, a_2, a_3

Структуру КДОМ можно представить как ориентированный граф с двумя видами дуг. Состояние каждого узла зависит от данных, которые поступают от узлов, входящих в его окрестность.

Общий вид функции g_i для узла a_i комплекснозначной динамической окрестностной модели описывается формулой [7]

$$(1) \quad x_i[t] = g_i(V_i[t], X_i[t]),$$

где $x_i[t] \in C$ – состояние модели в узле a_i в момент времени $t + 1$,

$$X_i[t] = (x_{i_1}[t] \dots x_{i_{r_i}}[t])^T \in C^{r_i}, V_i[t] = (v_{i_1}[t] \dots v_{i_{p_i}}[t])^T \in C^{p_i},$$

где r_i – количество окрестностных связей по состояниям; p_i – количество окрестностных связей по входам для узла a_i .

Общий вид системы (1) для рассматриваемого примера представлен формулой

$$(2) \quad \begin{cases} x_1[t + 1] = g_1(v_1[t], x_2[t]), \\ x_2[t + 1] = g_2(v_2[t], x_2[t], x_3[t]), \\ x_3[t + 1] = g_3(v_3[t], x_2[t], x_3[t]), \end{cases}$$

где $v_i[t]$ – вход узла a_i в момент времени t ; $x_i[t]$ – состояние в узле a_i в момент времени t .

В квадратичном случае система (2) представляется формулой

$$(3) \quad \begin{cases} x_1[t + 1] = a_1 + b_{11}v_1[t] + c_{12}x_2[t] + e_{111}v_1^2[t] + \\ + q_{122}x_2^2[t] + d_{112}v_1[t]x_2[t], \\ x_2[t + 1] = a_2 + b_{22}v_2[t] + c_{22}x_2[t] + c_{23}x_3[t] + \\ + e_{222}v_2^2[t] + q_{222}x_2^2[t] + q_{233}x_3^2[t] + d_{222}v_2[t]x_2[t] + \\ + d_{223}v_2[t]x_3[t] + q_{223}x_2[t]x_3[t], \\ x_3[t + 1] = a_3 + b_{33}v_3[t] + c_{32}x_2[t] + c_{33}x_3[t] + \\ + e_{333}v_3^2[t] + q_{322}x_2^2[t] + q_{333}x_3^2[t] + d_{332}v_3[t]x_2[t] + \\ + d_{333}v_3[t]x_3[t] + q_{323}x_2[t]x_3[t]; \end{cases}$$

где $a_i, b_{ij}, c_{ij}, d_{ijk}, e_{ijk}, q_{ijk} \in C$ – параметры модели, $i, j, k = 1, 2, 3$.

3. Идентификация

Идентификация объектов – это построение оптимальных математических моделей по реализациям их входных и выходных

данных [21]. Идентификация динамических объектов состоит в определении их структуры и параметров по наблюдаемым данным – входному воздействию и выходной величине [8]. Идентификация осуществляется с помощью настраиваемой модели той или иной структуры, параметры которой могут изменяться [17].

3.1. АЛГОРИТМ ИДЕНТИФИКАЦИИ КВАДРАТИЧНОЙ КОМПЛЕКСНОЗНАЧНОЙ ДИНАМИЧЕСКОЙ ОКРЕСТНОСТНОЙ МОДЕЛИ

Пусть задана квадратичная комплекснозначная динамическая окрестностная модель со структурой, приведенной на рис. 1, и функциями пересчета состояний (3). Для ее идентификации необходимо для каждого узла a_i найти решение переопределенной системы линейных уравнений

$$(4) \quad L_i A_i = F_i,$$

где $A_i \in C^{(1+p_i+r_i) \times 1}$,

$$L_i^T = \begin{bmatrix} 1+i & \dots & 1+i \\ v_i^1 & \dots & v_i^M \\ x_i^1 & \dots & x_i^M \\ v_i^1 \otimes v_i^1 & \dots & v_i^M \otimes v_i^M \\ x_i^1 \otimes x_i^1 & \dots & x_i^M \otimes x_i^M \\ v_i^1 \otimes x_i^1 & \dots & v_i^M \otimes x_i^M \end{bmatrix} \in C^{(1+p_i+r_i) \times M};$$

$$v_i^k = \begin{bmatrix} v_{i_1}^k \\ \vdots \\ v_{i_{p_i}}^k \end{bmatrix}, x_i^k = \begin{bmatrix} x_{i_1}^k \\ \vdots \\ x_{i_{r_i}}^k \end{bmatrix}; F_i = \begin{bmatrix} x_i^1[t+1] \\ \vdots \\ x_i^M[t+1] \end{bmatrix} \in C^{M \times 1}; \quad i = 1, \dots, n;$$

$$k = 1, \dots, M,$$

где A_i – матрица параметров; L_i – матрица исходных данных (общий вид), в которой $v_{i_{p_i}}$ – вектор входных данных i -го узла модели для p_i -й строчки в текущий момент времени t ; $x_{i_{r_i}}$ – состояния i -го узла модели для r_i -й строчки в текущий момент времени t ; F_i – матрица состояний i -го узла модели в $(t+1)$ -й момент времени; M – объем выборки; « $\otimes$ » – произведение Кронекера.

Матрицы A_i и L_i из системы (4) для каждого узла примера окрестностной модели (3) будут иметь вид

$$A_1 = \begin{bmatrix} a_1 \\ b_{11} \\ c_{12} \\ e_{111} \\ q_{122} \\ d_{112} \end{bmatrix}; L_1^T = \begin{bmatrix} 1+i & \dots & 1+i \\ v_1^1 & \dots & v_1^M \\ x_2^1 & \dots & x_2^M \\ (v_1^1)^2 & \dots & (v_1^M)^2 \\ (x_2^1)^2 & \dots & (x_2^M)^2 \\ v_1^1 \cdot x_2^1 & \dots & v_1^M \cdot x_2^M \end{bmatrix}.$$

A_2, A_3, L_2^T, L_3^T задаются аналогично.

Решение системы (4) находится методом наименьших квадратов. Пусть известны исходные данные входов и состояний в текущий момент времени, состояний в следующий момент времени ($V[t], X[t], X[t+1]$) рассматриваемой динамической модели. Для нахождения параметров для каждого узла a_i необходимо задать матрицы L_i, F_i [19]. Далее находится псевдообратная матрица по формуле

$$(5) \quad L_i^+ = \left(\overline{L_i}^T \cdot L_i \right)^{-1} \cdot \overline{L_i}^T.$$

Параметры A_i вычисляются по формуле (6):

$$(6) \quad A_i = L_i^+ \cdot F_i.$$

Для поиска ошибки идентификации необходимо получить модельные значения функции пересчета состояния F'_i по формуле (7) для найденных параметров A_i и сравнить с исходными данными F_i [18]. Для нахождения среднеквадратической ошибки используется формула (8), средней приведенной – формула (9):

$$(7) \quad F'_i = L_i \cdot A_i;$$

$$(8) \quad E_i = \frac{1}{M} \sum_{j=1}^M |F_{ij} - F'_{ij}|^2;$$

$$(9) \quad \delta_i = \frac{1}{M} \frac{\sum_{j=1}^M |F_{ij} - F'_{ij}|}{\max_j |F_{ij}|} \cdot 100\%.$$

Рассмотренный алгоритм идентификации реализован в пакете Mathcad.

3.2. ПРИМЕР ИДЕНТИФИКАЦИИ

Рассмотрим пример идентификации квадратичной комплекснозначной динамической окрестностной модели (3) для первого узла.

Исходные данные для рассматриваемого примера сгенерированы по закону равномерного распределения. Действительные и мнимые части комплексных чисел генерировались отдельно.

Далее по формулам (4)–(9) проводим идентификацию квадратичной комплекснозначной динамической окрестностной модели.

Сравним F_i с F'_i для квадратичной модели (см. рис. 2–4).

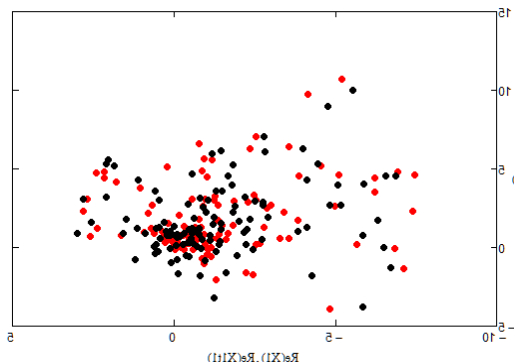


Рис. 2. График сравнения F_1 с F'_1 квадратичной модели

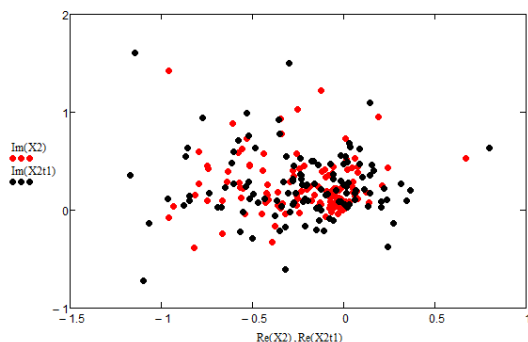


Рис. 3. График сравнения F_2 с F'_2 квадратичной модели

Ошибки идентификации полученной квадратичной модели приведены в таблице 1.

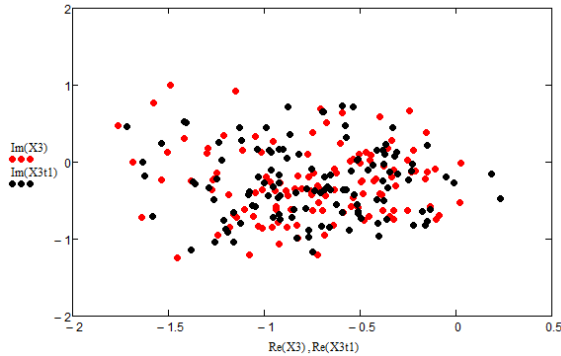


Рис. 4. График сравнения F_3 с F'_3 квадратичной модели

Таблица 1. Ошибки идентификации

Квадратичная ошибка	Приведенная ошибка
$E(X_1) = 0,583$	$\delta_{\text{пр}}(X_1) = 1,589\%$
$E(X_2) = 0,046$	$\delta_{\text{пр}}(X_2) = 3,205\%$
$E(X_3) = 0,047$	$\delta_{\text{пр}}(X_3) = 1,659\%$
$E_{\text{ср}} = 0,225$	$\delta_{\text{пр.ср}} = 2,151\%$

4. Идентификация комплекснозначной динамической окрестностной модели на кластеризованных данных

Рассмотрим идентификацию комплекснозначной динамической окрестностной модели на кластеризованных данных с помощью разработанной программы. При этом кластеризация выполняется с использованием метода k -средних.

Данный метод был модифицирован для кластеризации комплекснозначных наборов данных и использовался в работе в качестве иллюстрации применения кластеризации к комплекснозначным данным, так как он является простым в реализации и позволяет достичь заданной точности без больших вычислительных затрат.

4.1. КЛАСТЕРИЗАЦИЯ КОМПЛЕКСНОЗНАЧНЫХ ДАННЫХ МЕТОДОМ k -СРЕДНИХ

Кластерный анализ, или кластеризация, – это задача группировки набора объектов таким образом, чтобы объекты в одной группе, называемой кластером, были более похожи по различным признакам друг на друга, чем на объекты в других группах (кластерах) [1, 8, 13, 14, 16, 20].

Существуют различные группы методов кластеризации: эвристические, оптимизационные и иерархические. После проведения ряда вычислительных экспериментов в работе был выбран метод кластеризации k -средних.

Метод k -средних – это метод кластерного анализа, целью которого является разделение m наблюдений на k кластеров, при этом каждое наблюдение относится к тому кластеру, к центру (центроиду) которого оно ближе всего [2, 4, 5, 6, 11, 15]. В качестве меры близости используется евклидово расстояние [3, 9, 10, 12].

Ниже приведена схема кластеризации методом k -средних для комплекснозначных данных, см. рис. 5.

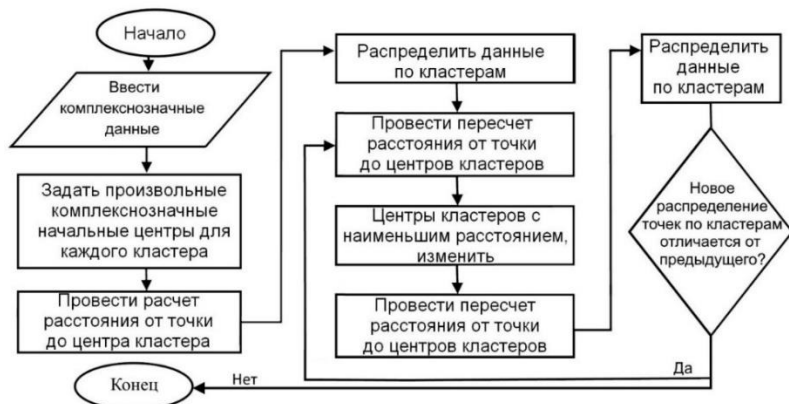


Рис. 5. Схема метода k -средних

4.2. ИДЕНТИФИКАЦИЯ НА КЛАСТЕРИЗОВАННЫХ ДАННЫХ

Рассмотрим идентификацию комплекснозначной динамической окрестностной модели на кластеризованных данных.

Сначала выполним разбиение исходных данных на три кластера и найдем псевдообратные матрицы (5) и параметры (6) для каждого узла и каждого кластера по формулам (4)–(9).

Далее находим модельные значения F'_i после кластеризации для каждого кластера по формуле (7). На рис. 6–8 приведены графики сравнения F_i с F'_i после кластеризации для квадратичной модели.

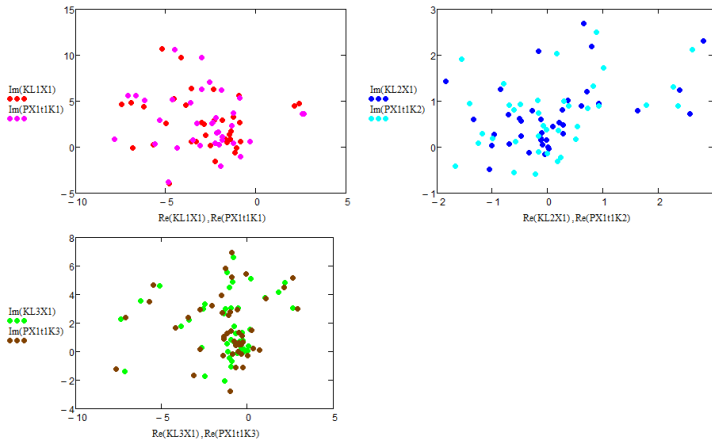


Рис. 6. Графики сравнения F_1 с F'_1 после кластеризации для квадратичной модели

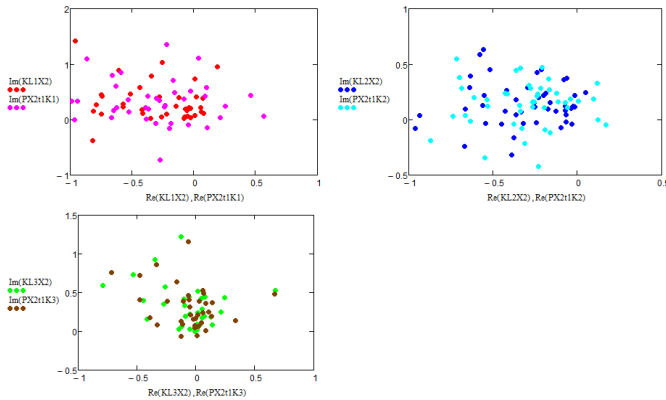


Рис. 7. Графики сравнения F_2 с F'_2 после кластеризации для квадратичной модели

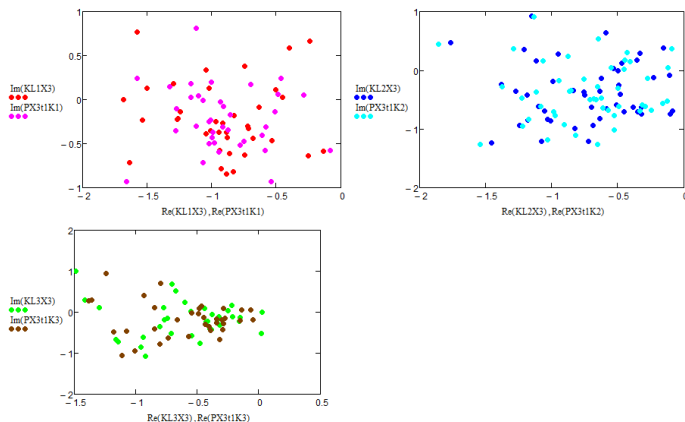


Рис. 8. Графики сравнения F_3 с F'_3 после кластеризации для квадратичной модели

Результаты ошибок идентификации комплекснозначной динамической окрестностной модели на кластеризованных данных приведены в таблице 2.

Таблица 2. Ошибки идентификации после кластеризации

Квадратичная ошибка	Приведенная ошибка
$E(X_1) = 0,251$	$\delta_{\text{пр}}(X_1) = 0,995\%$
$E(X_2) = 0,048$	$\delta_{\text{пр}}(X_2) = 2,954\%$
$E(X_3) = 0,04$	$\delta_{\text{пр}}(X_3) = 1,468\%$
$E_{\text{ср}} = 0,113$	$\delta_{\text{пр.ср}} = 1,806\%$

В результате исследования видно, что в данном примере в среднем после кластеризации ошибки идентификации комплекснозначной динамической окрестностной модели уменьшаются.

Следует отметить, что идентификация КДОМ с использованием метода кластеризации k -средних для комплекснозначных данных не всегда давала лучший результат по сравнению с идентификацией некластеризованных данных. На некоторых сгенери-

рованных выборках точки располагались примерно на многомерной плоскости. В этих случаях улучшения при кластеризации не происходило.

5. Заключение

В работе рассмотрены динамические окрестностные модели, идентификация квадратичной комплекснозначной моделей, метод кластеризации. В ходе работы была подробно исследована комплекснозначная динамическая окрестностная модель. На примере показана ее идентификация. Приведен алгоритм кластеризации комплекснозначных данных методом k -средних. Представлены результаты идентификации квадратичной комплекснозначной динамической окрестностной модели до и после кластеризации на сгенерированных данных. В дальнейшем планируется применение разработанных методов на прикладном примере и исследование идентификации КДОМ с применением других методов кластеризации.

Литература

1. БАЮК Д.А., БАЮК О.А., БЕРЗИН Д.В. и др. *Практическое применение методов кластеризации, классификации и аппроксимации на основе нейронных сетей*. – М.: Прометей, 2020. – 448 с.
2. БЕССМЕРТНЫЙ И.А., НУГУМАНОВА А.Б., ПЛАТОНОВ А.В. *Интеллектуальные системы. Учебник и практикум для вузов*. М.: Изд-во Юрайт, 2023. – 243 с. // Образовательная платформа Юрайт. – URL: <https://urait.ru/bcode/511999> (дата обращения: 24.05.2023).
3. ВОРОНОВ М.В., ПИМЕНОВ В.И., НЕБАЕВ И.А. *Системы искусственного интеллекта. Учебник и практикум для вузов*. – М.: Изд-во Юрайт, 2023. – 256 с. // Образовательная платформа Юрай. – URL: <https://urait.ru/bcode/519916> (дата обращения: 24.05.2023).
4. КРАСНОПЕРОВ К.Ю., СЕМЕНОВ В.А. *Анализ применения теории комплексных чисел в обработке и хранении данных // Современные научные исследования и инновации*. – 2023. –

- № 9. – URL: <https://web.snauka.ru/issues/2023/09/100742> (дата обращения: 03.02.2024).
5. НАЗАРОВ Д.М., КОНЫШЕВА Л.К. *Интеллектуальные системы: основы теории нечетких множеств. Учебник и практикум для вузов.* – М.: Изд-во Юрайт, 2023. – 186 с. // Образовательная платформа Юрайт. – URL: <https://urait.ru/bcode/514414> (дата обращения: 24.05.2023).
 6. ПЛАТОНОВ А.В. *Машинное обучение. Учебное пособие для вузов.* – М.: Изд-во Юрайт, 2023. – 85 с. // Образовательная платформа Юрайт. – URL: <https://urait.ru/bcode/520544> (дата обращения: 24.05.2023).
 7. СЕДЫХ И.А. *Линейные и квадратичные нечеткие иерархические окрестностные модели производства холоднокатаной стали* // Вестник ВГУ. Серия: Системный анализ и информационные технологии. – 2019. – №1. – С. 67–73.
 8. СЕДЫХ И.А., МАКАРОВ К.Н. *Нечеткая кластеризация комплекснозначных данных* // Вести учебных заведений Черногоземья. – 2023. – Т. 19, №2(72). – С. 46–57
 9. СКОРУБСКИЙ В.И., ПОЛЯКОВ В.И., ЗЫКОВ А.Г. *Математическая логика. Учебник и практикум для вузов.* – М.: Изд-во Юрайт, 2023. – 211 с. // Образовательная платформа Юрайт. – URL: <https://urait.ru/bcode/511996> (дата обращения: 24.05.2023).
 10. СТАНКЕВИЧ Л.А. *Интеллектуальные системы и технологии. Учебник и практикум для вузов.* – Саратов: Саратовский государственный технический университет имени Ю.А. Гагарина, ЭБС АСВ, 2019. — 160 с. // IPR SMART: – URL: <https://www.iprbookshop.ru/118365.html> (дата обращения: 24.05.2023).
 11. ТРЕГУБОВ В.Н., КАТКОВА М.А. *Информационное пространство логистического кластера: теория и методология формирования на основе облачных технологий.* – М.: Изд-во Юрайт, 2023. – 495 с. // Образовательная платформа Юрайт. – URL: <https://urait.ru/bcode/530657> (дата обращения: 24.05.2023).
 12. ХАЛИН В.Г. и др. *Системы поддержки принятия решений. Учебник и практикум для вузов* / Под редакцией В.Г. Халина,

- Г.В. Черновой. – М.: Изд-во Юрайт, 2023. – 494 с. // Образовательная платформа Юрайт. – URL: <https://urait.ru/bcode/511245> (дата обращения: 24.05.2023).
13. ARTHUR D., VASSILIVITSKII S. *k-means++: the advantages of careful seedings* // Proc. of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms, 2007. – 1027–1035 p.
 14. BLEI D.M., NG A.Y. JORDAN M.I. *Latent Dirichlet allocation* // Journal of Machine Learning Research, 2003. – P. 993–1022.
 15. BRADLEY P.S., FAYYAD U.M. *Refining initial points for k-means clustering* // Proc. of the 15th Int. Conf. of Machine Learning, San Francisco, 1998. – 99 p.
 16. CALINSKI T., HARABASZ J. *A dendrite method for cluster analysis* // Taylor & Francis Online. – 1974. – Vol. 3, Iss. 1. – P. 1–27.
 17. CHARU C.A., CHANDAN K.R. *Data Clustering: algorithms and applications*. – Chapman and Hall / CRC, 2013. – 652 p.
 18. DANGETI P. *Unsupervised feature selection. Computational Methods of Feature Selection*. – Packt Publishing, 2017. – 442 p.
 19. Dy J.G. *Statistics for Machine Learning*. – NewYork: Chapman and Hall / CRC, 2007. – 440 p.
 20. EDWIN M.K., RAYMOND T.Ng. *Algorithms for Mining Distance-Based Outliers in Large Datasets*. – Vancouver: VLDB, 1998. – P. 392–403.
 21. KASSAMBRA A. *Practical guide to cluster analysis in r*. – STHDA, 2017. – 187 p.

IDENTIFICATION OF QUADRATIC COMPLEX-VALUED DYNAMIC NEIGHBORHOOD MODELS ON CLUSTERED DATA AND WITH OUTCLUSTERING

Irina Sedykh, Lipetsk State Technical University, Lipetsk, Doctor of Technical Sciences, professor (sedykh-irina@yandex.ru).

Kirill Makarov, Lipetsk State Technical University, Lipetsk, student (kirik0-1@yandex.ru)

Abstract: Neighborhood models and their modifications used to model various distributed systems and processes. The study considers a quadratic complex-valued dynamic neighborhood model in which the parameters, inputs and states are complex

numbers, and its definition is given. The model functions in discrete time. An example of a complex-valued dynamic neighborhood model consisting of three nodes shown, for which the graph of the structure and the functions of the intersection of states given in general form. A special case of recalculation functions for a quadratic model is also considered. An algorithm for identifying a complex-valued dynamic neighborhood model whose parameters are determined by the least squares method given. A general view of the matrices of a system of linear equations for finding the parameters of a quadratic model shown. Matrices are given and identification performed for the considered example of a neighborhood model. The root-mean-square and average reduced identification errors are found. The paper also considers the identification of a complex-valued dynamic neighborhood model on clustered data. Clustering performed using complex data sets by the k-means method. The proposed identification algorithms implemented in the form of a program in the Mathcad package, with the help of which the results of identification of a quadratic complex-valued dynamic neighborhood model on clustered data and without clustering are compared.

Keywords: dynamic neighborhood model, quadratic model, complex numbers, identification, clustering, k-means method.

УДК 519.6

ББК 22.19

DOI: 10.25728/ubs.2024.111.2

*Статья представлена к публикации
членом редакционной коллегии А.К. Погодаевым.*

*Поступила в редакцию 29.08.2023.
Опубликована 30.09.2024.*

ОПРЕДЕЛЕНИЕ ЦЕНТРАЛЬНОСТИ ГРАФА АЛГОРИТМОМ PAGERANK С УЧЕТОМ ВЕСОВ СВЯЗЕЙ

Егоркин А. А.¹

(ФГБОУ ВО Российский государственный социальный
университет, Москва)

Работа посвящена нахождению центральности узлов взвешенных графов с учетом веса связей. Актуальность этой задачи обусловлена тем, что игнорирование весов дуг графа при нахождении центральности его узлов недопустимо для ряда прикладных задач, в первую очередь относящихся к задачам из финансовой сферы. В классической постановке алгоритма PageRank происходит потеря части информации о весах связей при формировании матрицы переходных вероятностей из матрицы смежности. Данный эффект был продемонстрирован в настоящей статье. Предложен метод определения центральности узлов сети, базирующийся на алгоритме PageRank, который позволяет учесть веса всех связей. В качестве примера рассматривался граф финансовых транзакций. Узлами графа являются клиенты коммерческого банка в том числе сам банк, а дугами – денежные переводы между узлами. Качество ранжирования определялось путем сравнения различных мер центральности с внешним параметром, который характеризует важность узла и не связан с сетевыми характеристиками графа. По результатам исследования было показано, что предлагаемая мера центральности лучшим образом ранжирует наиболее важные узлы графа по сравнению с иными мерами центральности. Также была продемонстрирована сходимость предлагаемого алгоритма.

Ключевые слова: центральность PageRank, цепи Маркова, взвешенный направленный граф.

1. Введение

На данный момент разработано большое количество мер центральности, позволяющих оценить значимость или важность узлов в графе на основании различных критериев [5]. К классическим мерам центральности традиционно относят центральность по близости, центральность по посредничеству,

¹ Антон Александрович Егоркин, аспирант (2-5@bk.ru).

центральность по степени, центральность по Кацу [13] и PageRank [15].

При этом перечисленные метрики центральности, как правило, используют одну из характеристик узла графа, например, количество соседей, кратчайших путей, проходящих через узел или расстояние до других узлов графа.

Так, в алгоритме PageRank, базирующемся на учете степени влияния соседей, не учитывается вес связи либо учитывается локально, так как при построении матрицы переходных вероятностей указанная информация в почти не используется. В настоящей статье рассматриваются прикладные задачи, для которых необходимо учитывать не только влияние соседей, но и веса связей, а также подходы к решению подобных задач.

1.1. БАЗОВЫЕ ПРИНЦИПЫ ОПРЕДЕЛЕНИЯ ЦЕНТРАЛЬНОСТИ НА ОСНОВЕ ЦЕПЕЙ МАРКОВА

Андрей Андреевич Марков был основоположником исследования случайных процессов, где состояние в будущем зависит только от текущего состояния и названных впоследствии его именем (марковский процесс или марковская цепь). Впервые данный процесс был сформулирован Марковым А.А. в 1907 г.

Данный процесс может быть сформулирован следующим образом. Рассмотрим сеть или граф (размера n на n), где матрицей $\mathbf{P} = (p_{ij})$ описываются вероятности перехода из одного узла-состояния в другой узел ($\sum_{i=1}^n p_{ij} = 1$ и $p_{ij} \geq 0$). Тогда для марковского процесса будет характерно:

- (1) $\mathbf{x}(k+1) = \mathbf{x}(k) \times \mathbf{P}$, $k \rightarrow \infty$,
- (2) $\mathbf{x} = \mathbf{x} \times \mathbf{P}$,

где $\mathbf{x}$ – вектор вероятностей распределения элементов по системе.

Существует предел с положительными элементами:

- (3) $\lim_{k \rightarrow \infty} \mathbf{P}^k$.

При достаточно большом k распределение элементов в сети не будет зависеть от начального вектора $\mathbf{x}$, а будет описываться лишь сходимостью произведений матрицы вероятностей переходов $\mathbf{P}$.

Из уравнения (2) следует, что задача сводится к нахождению левого собственного вектора при собственном значении, равном 1, которое, исходя из теоремы Перрона – Фробениуса, будет положительным.

Для того чтобы выполнялись условия формул (2)–(3), необходимо, чтобы матрица **P** была стохастической, неприводимой и апериодической, т.е. регулярной [1]. Данные требования являются следствием теоремы Перрона – Фробениуса. В литературе подобные матрицы иногда называются SIA (Stochastic Indecomposable Aperiodic) [6, 11].

Однако для реальных сетей условия неразложимости и апериодичности не всегда выполняются, особенно для интернет-сетей с «висячими узлами» – страницами без исходящих ссылок. С указанных страниц пользователь не может перейти на другие страницы в рамках заданной сети ($p_{ij} = 0 \forall j$). В таких случаях предполагается, что пользователь может случайным образом перейти на любую другую страницу, т.е. строка матрицы переходов заполняется значениями $1/n$ [14].

1.2. АЛГОРИТМ PAGERANK

Суть алгоритма заключается в том, что блуждание осуществляется не только по ссылкам самой сети, пользователь случайным образом может перемещаться в пространстве сети. Параметр $\alpha \in [0 \dots 1]$ – фактор телепортации – отвечает за случайное перемещение.

$$(4) \quad \mathbf{x}(t+1) = \mathbf{x}(t)[\alpha \cdot \mathbf{P} + (1 - \alpha) \cdot \mathbf{J}] .$$

Чем больше компонента $(1 - \alpha) \cdot \mathbf{J}$, отвечающая за случайное блуждание по графу (где **J** – квадратная матрица размера $n \times n$, состоящая из единиц), тем быстрее сходится процесс (4), но за счет потери информации о структуре исходного графа. И наоборот, чем меньше параметр α , тем медленнее сходится алгоритм (4), поскольку блуждание осуществляется по исходному графу. Оптимальный баланс обеспечивается при $\alpha = 0,85$ [10].

Существуют модификации алгоритма PageRank, в том числе персонализированный PageRank или PageRank с перезапуском [7, 17], которые предлагают различные способы «обработки» висячих узлов.

2. Потеря информации при формировании стохастической матрицы

Рассмотрим три матрицы смежности:

$$(5) \quad A_1 = \begin{bmatrix} 0 & 100 & 100 & 200 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 \end{bmatrix},$$

$$(6) \quad A_2 = \begin{bmatrix} 0 & 10 & 10 & 20 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 \end{bmatrix},$$

$$(7) \quad A_3 = \begin{bmatrix} 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 \end{bmatrix}.$$

Для преобразования матрицы смежности A в матрицу переходных вероятностей P можно использовать следующие подходы:

При наличии связи между узлами заполнять их значением $1/\sum_{i=1}^n \delta_{ij}$, где δ_{ij} – символ Кронекера:

$$(8) \quad P_1 = \begin{bmatrix} 0 & 0,33 & 0,33 & 0,33 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0,5 & 0,5 \\ 0 & 0 & 0,5 & 0 & 0,5 \\ 0 & 0,33 & 0,33 & 0,33 & 0 \end{bmatrix}.$$

При наличии связи между узлами заполнять их значением $1/\sum_{i=1}^n a_{ij}$:

$$(9) \quad P_2 = \begin{bmatrix} 0 & 0,25 & 0,25 & 0,5 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0,5 & 0,5 \\ 0 & 0 & 0,5 & 0 & 0,5 \\ 0 & 0,33 & 0,33 & 0,33 & 0 \end{bmatrix}.$$

Матрица переходных вероятностей P_1 может быть получена из любой матрицы смежности A_1, A_2, A_3 при заполнении элементов матрицы P_1 значениями $1/\sum_{i=1}^n \delta_{ij}$.

Аналогично и для матрицы переходных вероятностей P_2 , которая может быть получена из любой матрицы смежности A_1 , A_2 , A_3 при заполнении элементов матрицы P_2 значениями $1/\sum_{i=1}^n a_{ij}$.

Таким образом, происходит потеря важной информации об исходной матрице смежности. Для задач анализа интернет-сетей важен в первую очередь сам факт наличия ссылки и ее направление, но для многих других задач также важен и вес связи или дуги в терминах теории графов.

Существуют модификации PageRank для матриц вероятностей типа P_2 , где взвешенность учитывается на уровне отдельных узлов. Однако из-за стохастического характера этих матриц «глобальное» распределение весов не может быть учтено в рамках всего графа.

В каких задачах важно учитывать вес связей между узлами? Например, в задачах оценки финансовых рисков [9, 18] в которых вес дуги может представлять собой сумму транзакций в денежном выражении между клиентами, являющимися узлами графа [4]. В данном случае это является критически важной информацией, поскольку платеж на один рубль или один миллиард рублей – это существенно разные денежные потоки. В описанных выше подходах данная информация будет потеряна при создании стохастической матрицы вероятностей перехода.

3. Предложение по учету весов связей в алгоритме PageRank

Для решения описанной выше проблемы предлагается сначала выполнить кластеризацию дуг графа на основе их весов, далее рассчитать показатели центральности для сформированных кластеров, а затем агрегировать полученные вектора центральностей.

Как показано в [3], веса дуг графа финансовых транзакций подчиняются степенному закону распределения [19, 21]. В работе [2] нами было продемонстрировано, что логарифмирование исходных данных, подчиняющихся степенному закону распределения [8, 16], способствует более равномерной кластеризации данных. При этом результат не зависит от выбора основания

логарифма, и при использовании, например, алгоритма кластеризации *k-means* [12] кластеризация в одномерном случае осуществляется вокруг среднегеометрических значений соответствующего кластера.

Графическая интерпретация агрегирования векторов центральности представлена на рис. 1.

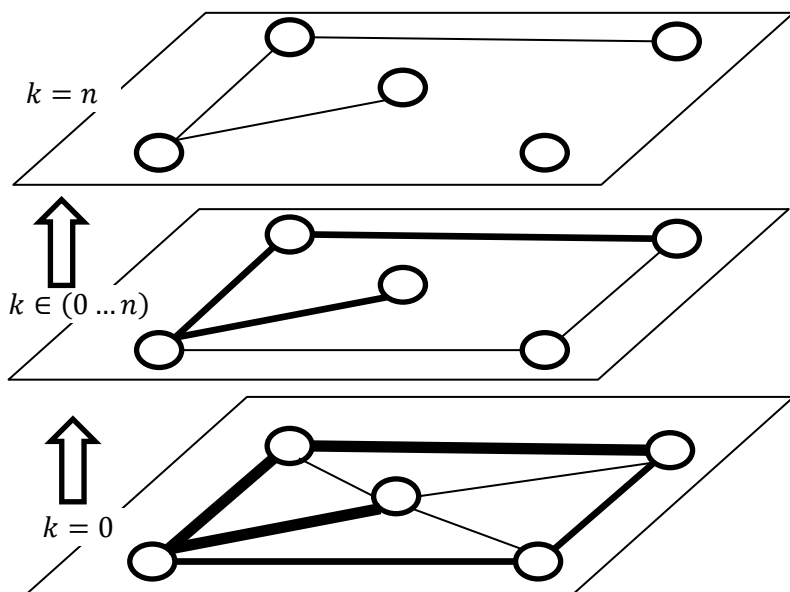


Рис. 1. Графическая интерпретация алгоритма

На шаге $k = 0$ определяется вектор центральности исходного графа s_0 . На следующих шагах $k \in [1, \dots, n - 1]$ определяется вектор центральности s_k графа, для которого выполняется условие, что веса дуг больше определенного порога $a_{ij} > b_k$. С каждым последующим шагом порог b_k увеличивается, а количество дуг уменьшается. Появляются висячие узлы, центральность которых минимальна и определяется исключительно случайным блужданием. Таким образом, для шага $k \in [1, \dots, n - 1]$ интерес представляет не абсолютная величина центральности, а ее отклонение от минимального значения.

На каждом из шагов центральность определяется для взвешенного графа. Принимается, что веса всех дуг в одном кластере одинаковы и равны среднегеометрическому значению $\overline{a_k}$.

Исходя из описанных выше предпосылок, предлагаемый алгоритм можно описать следующим образом:

$$(10) \mathbf{c} = \mathbf{c}_0 + \sum_{k=1}^{n-1} [\mathbf{c}_k - \min(\mathbf{c}_k)] \frac{\overline{a_k} - \overline{a_{k-1}}}{\overline{a_0}}$$

где $\mathbf{A} = \|a_{ij}\|$ – матрица смежности; $\mathbf{c}_k$ – вектор центральности PageRank матрицы смежности, в которой веса дуг больше порога b_k : $a_{ij} > b_k$; $\mathbf{c}_0$ – вектор центральности PageRank исходной матрицы смежности; $\overline{a_k}$ – среднегеометрическое значение весов дуг матрицы смежности с весами более порога b_k ; $\overline{a_{k-1}}$ – среднегеометрическое значение весов дуг матрицы смежности с весами более порога b_{k-1} , при этом $b_k > b_{k-1}$; $\overline{a_0}$ – среднегеометрическое значение весов дуг исходной матрицы смежности; n – количество шагов алгоритма.

Порядок определения порогов b_k в простейшем случае осуществляется следующим образом:

$$(11) b_k = 10^{d_k},$$

$$\text{где } d_k = \lg \min(a_{ij} > 0) + k \cdot \frac{\lg \max(a_{ij}) - \lg \min(a_{ij} > 0)}{n}$$

Итоговое значение вектора $\mathbf{c}$ нормируется на сумму значений вектора $\mathbf{c}$.

4. Пример использования алгоритма

Рассмотрим матрицу смежности

$$\mathbf{A} = \begin{bmatrix} 0 & 10 & 20 & 200 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 20 & 1 \\ 0 & 0 & 80 & 0 & 1 \\ 0 & 10 & 1 & 1 & 0 \end{bmatrix}.$$

Используя в алгоритме $n = 3$, имеем следующие промежуточные вычисления (таблица 1).

Таблица 1. Промежуточные расчеты

k	Матрица смежности	Вид графа	b_k	$\bar{a}_k$	c_k
0	$\begin{bmatrix} 0 & 10 & 20 & 200 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 20 & 1 \\ 0 & 0 & 80 & 0 & 1 \\ 0 & 10 & 1 & 1 & 0 \end{bmatrix}$		-	6,3	$\begin{bmatrix} 0,03 \\ 0,14 \\ 0,24 \\ 0,24 \\ 0,35 \end{bmatrix}$
1	$\begin{bmatrix} 0 & 10 & 20 & 200 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 20 & 0 \\ 0 & 0 & 80 & 0 & 0 \\ 0 & 10 & 0 & 0 & 0 \end{bmatrix}$		5,8	29,4	$\begin{bmatrix} 0,05 \\ 0,08 \\ 0,41 \\ 0,42 \\ 0,04 \end{bmatrix}$
2	$\begin{bmatrix} 0 & 0 & 0 & 200 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 80 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$		34,2	126,5	$\begin{bmatrix} 0,13 \\ 0,13 \\ 0,35 \\ 0,25 \\ 0,13 \end{bmatrix}$

Итоговое значение центральности получается путем агрегирования векторов c_k в соответствии с формулой (10):

$$c = \begin{bmatrix} 0,03 \\ 0,14 \\ 0,24 \\ 0,24 \\ 0,35 \end{bmatrix} + \begin{bmatrix} 0,04 - 0,04 \\ 0,08 - 0,04 \\ 0,41 - 0,04 \\ 0,42 - 0,04 \\ 0,04 - 0,04 \end{bmatrix} \frac{29,4 - 6,3}{6,3} + \begin{bmatrix} 0,13 - 0,13 \\ 0,13 - 0,13 \\ 0,35 - 0,13 \\ 0,25 - 0,13 \\ 0,13 - 0,13 \end{bmatrix} \frac{126,5 - 29,4}{6,3}.$$

После нормирования итоговый вектор центральности имеет вид

$$c = \begin{bmatrix} 0,003 \\ 0,03 \\ 0,55 \\ 0,39 \\ 0,02 \end{bmatrix}.$$

В стандартном алгоритме PageRank (вектор c_k на шаге 0) наиболее важным узлом является узел №5. В предложенной мере центральности наиболее важными узлами являются узлы №3 и №4, так как в них заходят дуги с наибольшим весом.

5. Результаты применения алгоритма на реальных данных

В качестве примера применения алгоритма рассмотрим граф финансовых транзакций коммерческого банка. Узлами графа являются клиенты банка, дугами – денежные переводы между узлами за период с 2022 по 2023 год. Вес дуги – сумма транзакций в денежных единицах [21].

Рассмотрим следующие меры центральности для направленного графа [20]:

- центральность по близости (Closeness);
- центральность по посредничеству (Betweenness);
- центральность по степени (Degree);
- PageRank;
- предлагаемая мера (Financial PageRank).

Для того чтобы оценить качество ранжирования узлов рассматриваемыми мерами, сравним полученные значения центральностей с внешними данными, которые характеризуют степень важности узла. Узлами в графе являются экономические агенты, целесообразно использовать в качестве внешних данных экономические показатели, например такие, как размер собственных средств, валюта баланса, объем выручки и т.п. Так как узлы в графе в большей своей части являются пользователями услуг банка, то в качестве экзогенных данных будем использовать объем бизнеса, который клиент имеет с банком.

Объем бизнеса (*Volume*) определяется как средние дневные остатки по следующим банковским продуктам: гарантии, выданные кредиты, привлеченные депозиты и средства на текущих

счетах. Поскольку не все узлы графа имеют перечисленные выше продукты, то в анализе используем только тех клиентов у кого объем $Volume > 0$.

Рассмотрим граф, описанный выше. Граф имеет следующие характеристики:

- количество узлов – 4,5 тыс. шт.;
- количество дуг – 6,5 тыс. шт.

Определим для узлов данного графа все меры центральности, перечисленные выше. Выберем из графа узлы, которые имеют показатель $Volume > 0$, и отсортируем их по убыванию. Количество узлов, у которых $Volume > 0$, равно 2,1 тыс. шт.

Ценностью меры центральности является способность ранжирования узлов. Для оценки эффективности ранжирования, разобьем узлы графа на корзины по показателю $Volume$ нарастающим итогом. Далее оценим, какое количество верно отранжированных узлов попадет в соответствующую корзину для каждой меры центральности.

Показатель $Volume$ имеет неравномерное распределение. На первые 2,5% узлов, упорядоченных по данному показателю, приходится 63% всего суммарного показателя $Volume$, на первые 10% узлов – 82%. Исходя из этого, основное требование к мере центральности в данном примере – это корректное ранжирование наиболее важных узлов – узлов с высокими показателями $Volume$ (таблица 2).

Таблица 2. Сравнение мер центральности

№	Доля узлов, упорядоченных по показателю Volume	Доля Volume	Доля правильно определенных узлов				
			Betweenness	Closeness	Degree	Financial Pagerank	Pagerank
1	2.5%	63%	30%	30%	43%	58%	34%
2	5.0%	73%	39%	32%	50%	57%	40%
3	10%	82%	50%	40%	54%	57%	43%
4	15%	88%	55%	44%	63%	57%	49%
5	25%	94%	61%	51%	68%	64%	55%
6	50%	99%	68%	61%	72%	68%	65%
7	100%	100%	100%	100%	100%	100%	100%

В приведенных результатах видно, мера центральности *Financial Pagerank* лучше ранжирует узлы с высокими показателями *Volume*. Среди узлов с самыми крупными значениями объема бизнеса (строка №1 в таблице 2) предложенная мера корректно определяет 58% из первых 2,5% узлов графа, что является лучшим результатом среди рассматриваемых мер центральности.

Далее по мере увеличения количества узлов, упорядоченных по *Volume*, увеличиваются и доли корректно предсказываемых узлов по всем мерам центральности. Начиная с первых 15% узлов графа лучшими ранжирующими способностями обладает мера центральности по степени (*Degree*).

За счет того, что анализируемая мера учитывает весь спектр весов дуг графа, распределение центральностей по сравнению со стандартным алгоритмом *PageRank* становится более полярным, т.е. наиболее важные узлы имеют большее абсолютное значение центральности, а менее важные – меньшее. Данный эффект демонстрирует следующий график, показывающий сравнение значений центральностей (рис. 2).

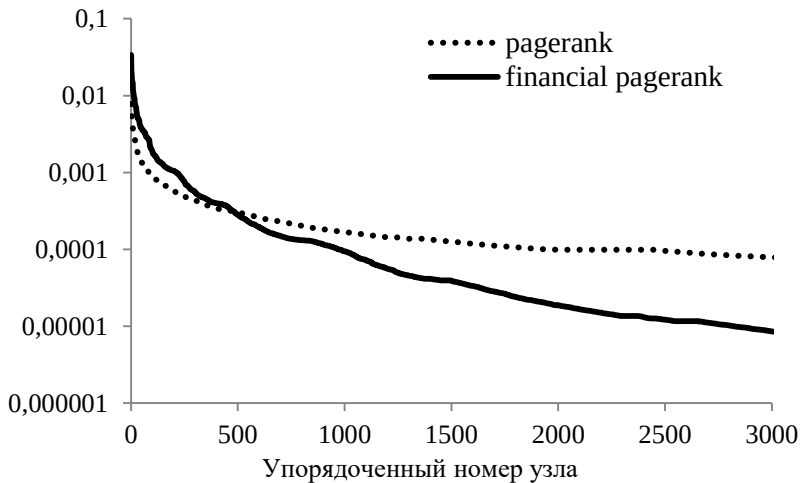


Рис. 2. Сравнение значений центральности
(в логарифмических координатах)

5.1 СХОДИМОСТЬ АЛГОРИТМА

Ключевой внешней переменной алгоритма является количество кластеров n , на которое целесообразно разбивать дуги графа. Нахождение оптимального значения n является отдельной задачей, например, в статье [2] был описан подход к определению оптимального количества кластеров, исходя из поиска локального максимума коэффициента силуэта.

При этом ожидаемо, что при увеличении показателя n алгоритм должен показывать значения, близкие к $n - 1$. Ниже приведен график корреляции Спирмена между центральностями Financial PageRank, рассчитанными для количества кластеров n и $n - 1$.

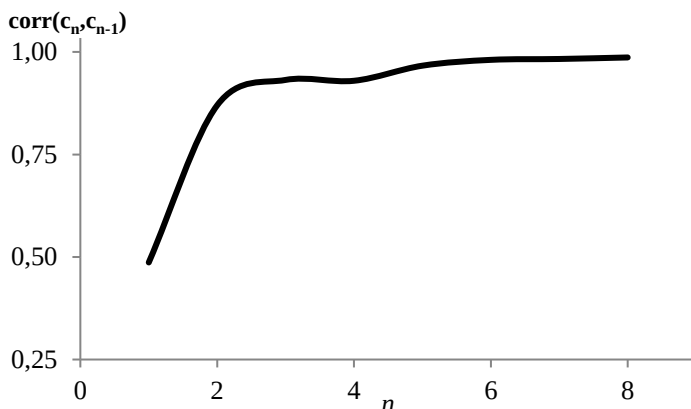


Рис. 3. Корреляция Спирмена для количества кластеров n и $n - 1$

На графике мы видим, что начиная с количества кластеров $n \geq 5$ дальнейшее увеличение количества кластеров не ведет к изменению итоговой оценки.

6. Заключение

Учет весов связей при определении центральности узлов в ряде прикладных задач может иметь критически важный характер. Как нами было показано, это особенно важно в области финансовых вычислений. В представленной работе предложен

подход к реализации данной задачи, показавший возможность применения на реальных данных.

За счет учета весов дуг, предложенная мера центральности имеет большую полярность в абсолютных значениях центральности по сравнению с мерой *PageRank*. Это помогает предложенной мере центральности наилучшим образом осуществлять ранжирование наиболее важных узлов.

Отдельно требуется рассмотреть задачу сходимости предложенного алгоритма при увеличении количества кластеров. Из приведенных расчетов видно, что итоговая центральность сильно отличается от центральности на первом шаге, которая представляет собой центральность *PageRank*. Но при достаточном количестве кластеров (пять и более) дальнейшее увеличение количества кластеров не приводит к существенному изменению предлагаемой меры центральности.

Определение оптимального количества кластеров является отдельной задачей. В [2] были предложен подход к поиску наилучшего количества кластеров, но он не учитывал всех аспектов исследуемой меры центральности.

Также к недостаткам предложенной меры центральности следует отнести сложность ее расчета, особенно при большом количестве кластеров.

Литература

1. ГАНТМАХЕР Ф.Р. *Теория матриц*. – М.: Физматлит. 2004. – 559 с.
2. ЕГОРКИН А.А. Особенности использования алгоритма классификации *k-means* для данных, подчиненных степенному закону распределения // Современная наука: актуальные проблемы теории и практики. Серия: Естественные и технические науки. 2023. – №9. – С. 65–69.
3. ЕГОРКИН А.А., АГАЕВ Р.П. *Использование степенного закона в финансовых задачах* // Наука. Производство. Образование: Сборник научных трудов Всероссийской научно-технической конференции, Российский государственный социальный университет, 2023. – С. 58–64

4. ПЫРКИНА О.Е., ЗАДАДАДЕВ С.А. *Разработка критериев безопасности сложных систем электронного денежного оборота с помощью графовой аналитики* // Управление развитием крупномасштабных систем (MLSD'2021): Труды Четырнадцатой международной конференции: Институт проблем управления им. В.А. Трапезникова РАН, 2021. – С. 765–771.
5. ЧЕБОТАРЕВ П.Ю., ГУБАНОВ Д.А. *How to choose the most appropriate centrality measure?* // ArXiv. – 2020. – P. 1–25.
6. ALI M., ISLAM M., SALMAN A. *Average Convergence for Directed & Undirected Graphs in Distributed Systems* // Computer Systems Science and Engineering, 2021. – P. 399–413.
7. AVRACHENKOV K., HOFSTAD R., SOKOL M. *Personalized PageRank with Node-dependent Restart* // Research report, 2014. – 12 p.
8. BARABASI A.L., BONABEAU E. *Scale Free Networks* // Scientific American. – 2003. – P. 60–69.
9. BARDOSCIA M., BARUCCA P., BATTISTON S. et al. *The physics of financial networks* // Nat Rev Phys. – 2021. – No. 3. – P. 490–507.
10. BRIN S., PAGE L. *The anatomy of a large-scale hypertextual web search engine* // Proc. of the 7th Int. Conf. on World Wide Web. ACM, 1998. – P. 107–117.
11. CHEVALIER P-Y, GUSEV V.V., HENDRICKX J.M. et al. *Sets of Stochastic Matrices with Converging Products: Bounds and Complexity.* // ArXiv. – 2017. – 19 p.
12. GHADIRI M., SAMADI S., VEMPALA S. *Socially fair k-means clustering.* // ACM Conf. on Fairness, Accountability, and Transparency. – 2021. – P. 438–448.
13. KATZ L. *A new status index derived from sociometric analysis* // Psychometrika. – 1953. – P. 39–43.
14. LANGVILLE N., MEYER C.D. *Google's PageRank and Beyond: The Science of Search Engine Rankings.* – Princeton University Press, 2006. – 224 p.
15. NEWMAN M.E.J. *Networks: An Introduction* // Oxford University Press, 2012. – 241 p.

16. NEWMAN M.E.J., AARON C., COSMA R.S. *Power-law distributions in empirical data* // SIAM Review. – 2009. – No. 4 – P. 661–703.
17. PANCHENDRARAJAN R., SAXENA A. *Topic-based influential user detection: a survey* // Applied Intelligence. – 2023. – 53 p.
18. PEI Y., FANG L., IPENBURG W. et al. *Subgraph anomaly detection in financial transaction networks*. // ACM Int. Conf. on AI in Finance, 2020. – P. 1–8.
19. SAXENA A. *Evolving Models for Dynamic Weighted Complex Networks*. In *Principles of Social Networking*. – Springer, 2023. – 246 p.
20. SAXENA A., IYENGAR S. *Centrality Measures in Complex Networks: A Survey* // ArXiv. – 2020. – 75 p.
21. SAXENA A., PEI Y., VELDSINK J. et al. *The Banking Transactions Dataset and its Comparative Analysis with Scale-free Networks* // ACM Int. Conf. on Advances in Social Networks Analysis and Mining, 2021. – P. 283–296.

PAGERANK CENTRALITY, ACCOUNTING FOR THE WEIGHTS OF LINKS

Anton Egorkin, Russian State Social University, Moscow, Graduate Student (2-5@bk.ru).

Abstract: The work is devoted to finding the centrality of nodes of weighted graphs. The relevance of this task is due to the fact that ignoring the weights of the arcs of the graph when finding the centrality of its nodes is unacceptable for a number of applied tasks, primarily related to tasks from the financial sphere. In the classical formulation of the PageRank algorithm, part of the information about the weights of connections is lost when forming a matrix of transient probabilities from the adjacency matrix. This effect has been demonstrated in this article. A method for determining the centrality of network nodes is proposed, based on the PageRank algorithm, which allows taking into account all the weights of the links. The graph of financial transactions was considered as an example. The nodes of the graph are the clients of a commercial bank, including the bank itself, and the arcs are money transfers between nodes. The ranking quality was determined by comparing various centrality measures with an external node parameter unrelated to the network characteristics of the transaction graph. According to the results of the study, it was shown that the proposed centrality measure ranks the most important nodes of the

graph in the best way compared to other centrality measures. The convergence of the proposed algorithm was also demonstrated.

Keywords: PageRank, centrality, Markov chains, weighted directed graph.

УДК 519.8

ББК 22.18

DOI: 10.25728/ubs.2024.111.3

*Статья представлена к публикации
членом редакционной коллегии Г.А. Угольниковым.*

Поступила в редакцию 20.04.2024.

Опубликована 30.09.2024.

КОМПЛЕКСНАЯ ОЦЕНКА ИНФОРМАЦИОННЫХ РИСКОВ. II: АЛГОРИТМ ИДЕНТИФИКАЦИИ СТРУКТУРЫ ДЕРЕВА КОМПЛЕКСНОЙ ОЦЕНКИ ИНТЕГРАЛЬНОГО ИНФОРМАЦИОННОГО РИСКА

Рей А. С.¹, Широкий А. А.²

(ФГБУН Институт проблем управления
им. В.А. Трапезникова РАН, Москва)

При оценке рисков информационной безопасности весьма важным представляется учет характерных для информационных систем видов неопределенности. Существующие методы и алгоритмы оценки информационных рисков могут не учитывать некоторые из них, вследствие чего полученные оценки рисков могут быть искажены. В связи с этим естественным образом возникает задача разработки нового или адаптации уже существующего метода для оценки рисков сложных систем с учётом всех характерных для рассматриваемого класса систем видов неопределенности. В настоящей работе развивается ранее предложенная идея применения для оценки информационных рисков метода комплексного оценивания, предполагающего агрегацию оценок информационной системы по стандартным критериям информационной безопасности – конфиденциальности, целостности и доступности. В первой части работы было показано, что этот метод при соответствующих модификациях позволяет учесть все нужные виды неопределенности. В этой части работы предлагается алгоритм идентификации структуры дерева комплексной оценки на основе принципа сворачивания связанных критериев. Работоспособность алгоритма продемонстрирована на примере построения деревьев оценки рисков конфиденциальности, целостности и доступности для SMART-систем на основе «Интернета вещей».

Ключевые слова: сложные информационные системы, интегральный риск, комплексная оценка, системы Интернета вещей.

1. Введение

В первой части работы (см. [16]) авторами были проанализированы укрупненные группы подходов и методов за период 2019–23 гг., направленные на оценку информационных рисков. Группы и их элементы были классифицированы по основанию учета характерных для сложных информационных систем видов неопределенности, а именно:

¹ Анастасия Сергеевна Рей, м.н.с. (a.rey@ipu.ru).

² Александр Александрович Широкий, к.ф.-м.н., с.н.с. (shiroky@ipu.ru).

А. неопределённость значений отдельных факторов оценки состояния системы в целом;

В. неопределённость взаимного влияния элементов системы друг на друга;

С. неопределённость зависимости риска системы в целом от значений локальных (точечных) рисков.

В обзоре был проведен анализ качественных и количественных методов, таких как безмодельный подход [27, 31, 38], статистический подход [18, 28], подход нечеткой математики [13, 14], энтропийный подход [20], использование открытых стандартов [8], теоретико-графовый подход [1, 20, 22, 24, 32, 33, 36], «когнитивная игра» [11, 15], социотехнический подход [7, 26], попарное сравнение [35], анализ иерархий [5, 30], сценарный подход [19, 21], ранжирование [34, 37], анализ уровня защиты и блокады [23], методы агрегирования [2, 3, 4, 6].

Анализ показал, что в настоящее время большинство исследований сфокусировано на оценке локальных рисков, небольшое число работ рассматривает неопределенность значений отдельных факторов оценки, а взаимное влияние элементов системы друг на друга и, как следствие, на ее интегральный риск не рассматривается практически совсем. Методов, позволяющих учесть одновременно неопределенности всех трех видов, в литературе обнаружено не было.

Тем не менее среди рассмотренных подходов и методов можно выделить ряд наиболее перспективных (в обсуждаемом смысле). К таковым авторы относят количественный метод «когнитивной игры» [11, 15] и качественный метод комплексного оценивания (МКО) [3, 6]. Последний представляет больший интерес, поскольку область применения качественных методов обычно шире, чем количественных. Поэтому главной задачей второй части настоящей работы является его адаптация для оценки информационных рисков сложных систем. В случае успеха в дальнейшем можно будет перейти к модификациям метода для учета неопределенности вышеперечисленных видов.

Для решения поставленной задачи был разработан алгоритм идентификации структуры дерева комплексной оценки на основе принципа сворачивания связанных критериев, в контексте оценки информационных рисков впервые предложенного в [12].

Работоспособность алгоритма продемонстрирована на примере построения деревьев оценки рисков конфиденциальности, целостности и доступности для SMART-систем на основе «Интернета вещей».

2. Материалы и методы

2.1. ОБЩАЯ ПОСТАНОВКА ЗАДАЧИ

Рассмотрим некоторую информационную систему S . Её безопасность оценивается в следующих аспектах: конфиденциальность (C), целостность (I), доступность (A) [25]. Будем считать, что значения соответствующих критериев лежат в следующих множествах:

$$(1) \quad CE = \{1, \dots, n_C\}; \quad IE = \{1, \dots, n_I\}; \quad AE = \{1, \dots, n_A\}, \\ n_C, n_I, n_A \in \mathbb{N}.$$

Пусть комплексная оценка информационной безопасности представляется одним из значений множества $SE = \{1, \dots, n_S\}$, $n_S \in \mathbb{N}$. Тогда процедура её получения сводится к определению некоторого отображения

$$(2) \quad E: CE \times IE \times AE \rightarrow SE,$$

где $SE \subset \mathbb{N}$ – множество возможных значений (уровней) интегрального риска.

Решение задачи (1) обсуждалось в работе [11]. При этом предполагалось, что значения $K_C \in CE$, $K_I \in IE$, $K_A \in AE$ уже тем или иным способом получены, т.е. известны. Древовидная структура для интегральной оценки строится из соображения частичной зависимости рисков доступности и целостности, так как утрата целостности необходимо ведет к утрате доступности. Следовательно, локальные критерии K_I и K_A следует свернуть в первую очередь. Утрата же конфиденциальности не влияет ни на целостность, ни на доступность, поэтому соответствующий критерий K_C следует сворачивать с результатом свёртки K_I и K_A (см. рис. 1).

Развивая этот подход, опишем способ идентификации под-деревьев комплексного оценивания, корнями которых являются критерии K_C , K_I и K_A . Вначале определим субкритерии, исходя из следующих соображений.

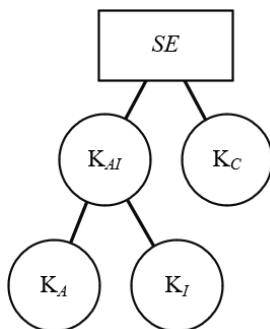


Рис. 1. Структура бинарного дерева комплексного оценивания интегрального информационного риска [11]

Каждая информационная система имеет особенности, которые потенциально снижают её конфиденциальность, целостность или доступность. Объединим особенности, известные лицу, принимающему решения, во множество уязвимостей U . Внешние или внутренние воздействия (целенаправленные или нет) на одну или несколько уязвимостей из множества U будем называть атаками.

Пусть A – множество атак. Будем называть выполнимостью атаки функцию

$$(3) \quad f: A \times 2^U \rightarrow \{0, 1\}.$$

Отметим, что

$$(4) \quad \forall a \in A \exists! u_a \subset 2^U: f(a, u) = \begin{cases} 1, & u \in u_a, \\ 0, & u \notin u_a. \end{cases}$$

Множество u_a будем называть множеством выполнимости атаки a . Зададим функцию

$$(5) \quad s: 2^U \rightarrow \mathbb{R},$$

характеризующую тяжесть атаки. Если в рассматриваемой системе уязвимости не ранжированы тем или иным способом, то её значение для произвольной атаки $a \in A$ равно мощности соответствующего ей множества u_a , т.е.

$$(6) \quad s(u_a) = |u_a| \in \mathbb{N}.$$

Если атаки a_i, a_j такие, что $u_{a_i} \cap u_{a_j} = \emptyset$, то будем их называть независимыми, в противном случае – зависимыми. Зададим функцию

$$(7) \quad \mathfrak{d}: 2^U \times 2^U \rightarrow \mathbb{R},$$

характеризующую степень зависимости атак друг от друга. Если в рассматриваемой системе уязвимости не ранжированы тем или иным способом, то для некоторых двух атак $a_i, a_j \in A$

$$(8) \quad \mathfrak{d}(u_{a_i}, u_{a_j}) = |u_{a_i} \cap u_{a_j}| \in \mathbb{N} \cup \{0\}.$$

Поскольку в конкретной информационной системе ряд уязвимостей может отсутствовать, а также учитывая возможность развития системы во времени и, соответственно, изменения состава уязвимостей, использовать их в качестве критериев-вершин дерева комплексного оценивания неразумно: построенное таким образом дерево будет требовать перестройки при изменении системы, в том числе при выполнении действий, направленных на снижение её риска (удаление уязвимостей). Следовательно, роль вершин должны выполнять критерии, рассчитываемые как значения функции, характеризующей тяжесть атаки для рассматриваемой системы. Структура же дерева определяется взаимной зависимостью типовых атак, по аналогии с [8].

В следующем параграфе представлен алгоритм построения дерева комплексного оценивания, не зависящий от наличия или отсутствия конкретных уязвимостей в защищаемой системе.

2.2. АЛГОРИТМ ИДЕНТИФИКАЦИИ СТРУКТУРЫ ДЕРЕВА КОМПЛЕКСНОЙ ОЦЕНКИ ИНТЕГРАЛЬНОГО ИНФОРМАЦИОННОГО РИСКА

Предлагаемый алгоритм состоит из шести шагов и позволяет построить дерево комплексного оценивания интегрального риска для заданного класса информационных систем. При описании для простоты будем считать, что все типичные уязвимости для рассматриваемого класса систем эквивалентны в смысле их вклада в интегральный риск. Отметим, что алгоритм легко обобщается на случай ранжированных уязвимостей.

1. Определить множество уязвимостей U как перечень типичных уязвимостей для рассматриваемого класса информационных систем, а также множество типичных атак $A = A_C \cup A_I \cup A_A$, где A_C , A_I и A_A порождаются списками типов атак, влияющих на конфиденциальность, целостность и доступность защищаемой системы соответственно.

2. Построить бинарные матрицы E_C, E_I, E_A размерностей $|2^{A_C}| \times |2^U|, |2^{A_I}| \times |2^U|, |2^{A_A}| \times |2^U|$ соответственно, описывающие, с одной стороны, возможность осуществления атаки конкретного типа при наличии той или иной комбинации уязвимостей, а с другой – зависимость типов атак друг от друга через общие комбинации зависимостей. Если атака $a \in A_C$ (или некоторая их комбинация $P \in 2^{A_C}$) может быть проведена с использованием уязвимости $u \in U$ (или комбинации уязвимостей $Q \in 2^U$), то в соответствующей клетке матрицы E_C должна стоять 1, в противном случае – 0. Матрицы E_I и E_A заполняются аналогично.

3. Для матриц E_C, E_I, E_A рассчитать векторы S_C, S_I, S_A размерностей $|2^{A_C}|, |2^{A_I}|, |2^{A_A}|$ соответственно, элементы которых являются суммами строк матриц E_C, E_I, E_A . Элементами этих векторов будут числа, отражающие число комбинаций уязвимостей, с использованием которых может быть реализована конкретная атака (или их комбинация). Отметим, что в случае ранжированных уязвимостей в клетках матриц E_C, E_I, E_A вместо нулей и единиц будут содержаться некоторые числа, характеризующие опасность уязвимости. В этом случае суммы, являющиеся элементами векторов S_C, S_I, S_A , затруднительно содержательно интерпретировать. Однако учитывая, что получаемые с их помощью значения мы в дальнейшем будем использовать лишь для ранжирования, это не является проблемой.

4. Аналогично п. 2 построить матрицы D_C, D_I, D_A с тем отличием, что единица в клетке матрицы $D_C (D_I, D_A)$ ставится в случае, если соответствующая комбинация уязвимостей была отмечена в матрице $E_C (E_I, E_A)$ единицей хотя бы для одной атаки из соответствующей текущей строке комбинации.

5. Аналогично п. 3 рассчитать векторы B_C, B_I, B_A как суммы строк матриц D_C, D_I, D_A соответственно.

6. Рассчитать векторы $Y_C = \frac{S_C}{B_C}, Y_I = \frac{S_I}{B_I}, Y_A = \frac{S_A}{B_A}$. Элементы этих векторов будут рассматриваться как показатели зависимости соответствующих атак.

7. Построить поддеревья комплексного оценивания, корнями которых являются критерии K_C, K_I и K_A , пошагово по следующему принципу: на текущем шаге агрегируются те пары под-

структур, соответствующее значение показателя зависимости которых является максимальным среди всех пар допустимых к агрегации на текущем шаге подструктур. Агрегация подразумевает добавление в дерево новой вершины-критерия с двумя рёбрами, соединяющими её с агрегируемым подструктурами. Например, на первом шаге мы можем агрегировать только листья будущего дерева (вершины, соответствующие типовым атакам, определённым на шаге 1). На втором шаге агрегации могут подвергаться как не агрегированные ранее вершины, соответствующие атакам, так и новые вершины, добавленные на предшествующих шагах и являющиеся корнями некоторых поддеревьев.

8. Установить значения для каждого из листьев, построенных на предыдущем шаге поддеревьев комплексного оценивания. В качестве базового значения можно использовать первые $|A_C|$, $|A_I|$, $|A_A|$ элементов векторов S_C, S_I, S_A соответственно, характеризующие опасность атаки для рассматриваемого класса информационных систем. Эти значения необходимо привести к порядковой шкале с n градациями. Стандартный вариант шкалы обычно включает три уровня: высокий, средний или низкий уровень риска [9].

9. Задать монотонные матрицы сверток размерностью $n \times n$ для определения значений агрегированных критериев, соответствующих вершинам дерева КО, не являющимся листьями. Допустимо использовать правила построения матриц с помощью метода порогового агрегирования [4] (см. рис. 2).

K_1	1	3	2	1
	2	3	2	2
	3	3	3	3
		3	2	1
	K_2			

Рис. 2. Пример монотонной матрицы свёртки

Следует упомянуть о возможных предельных случаях, с которыми можно столкнуться при осуществлении такой оценки:

1. Если атаки в пределах поддерева независимы ($u_{a_i} \cap u_{a_j} = \emptyset \forall i \neq j: a_i, a_j \in A_C$ для поддерева конфиденциаль-

ности, для других поддеревьев запись будет аналогичной), то построить дерево комплексной оценки на основе взаимной зависимости критериев не получится и следует использовать другие методы, например, рассчитать значение соответствующего локального критерия как взвешенную сумму значений (для поддерева конфиденциальности $K_C = \sum_{a_c \in A_C} k_{a_c} \cdot s(u_{a_c})$, где k_{a_c} – некоторые коэффициенты).

2. Если в пределах поддерева соответствующие атакам множества подмножеств уязвимостей, на которых атака выполнима, непусты и совпадают ($u_{a_i} = u_{a_j} \neq \emptyset \forall i \neq j: a_i, a_j \in A_C$ для поддерева конфиденциальности, для других поддеревьев запись будет аналогичной), то для получения значения соответствующего критерия следует воспользоваться методом порогового агрегирования [4].

3. Наконец, возможна ситуация, когда в пределах поддерева окажутся три или более атак с равномоными попарными пересечениями их множеств выполнимости. В этом случае решение задачи порождения поддерева КО не будет единственным и выбор конкретного поддерева следует сделать на следующем этапе, когда при рассмотрении конкретной информационной системы будет возможно провести ранжирование уязвимостей.

3. Результаты и обсуждение

3.1. ИЛЛЮСТРАЦИЯ РЕЗУЛЬТАТОВ

Для иллюстрации описанного выше алгоритма составим синтетический пример. Рассмотрим некоторую сложную систему, характеризующуюся четырьмя видами уязвимостей $u_1, \dots, u_4$ и тремя видами атак $a_1, \dots, a_3$. Пусть

$$U = \{u_1, u_2, u_3, u_4\}, A = \{a_1, a_2, a_3\},$$

$$u_{a_1} = \{u_2, u_3, u_4\}, u_{a_2} = \{u_1, u_4\}, u_{a_3} = \{u_2, u_3\}.$$

Тогда матрица соответствия уязвимостей атакам будет иметь следующий вид (таблица 1).

Составим матрицу E для данной системы и рассчитаем значения вектора S . Результат в транспонированном виде (для удобства чтения) представлен в таблице 2.

Таблица 1. Матрица соответствия уязвимостей атакам

	u_1	u_2	u_3	u_4
a_1		1	1	1
a_2	1			1
a_3		1	1	

Таблица 2. Транспонированная матрица E и значения вектора S

	a_1	a_2	a_3	a_1, a_2	a_1, a_3	a_2, a_3	a_1, a_2, a_3
u_1		1		0	0	0	0
u_2	1		1	0	1	0	0
u_3	1		1	0	1	0	0
u_4	1	1		1	0	0	0
u_1, u_2	0	0	0	0	0	0	0
u_1, u_3	0	0	0	0	0	0	0
u_1, u_4	0	1	0	0	0	0	0
u_2, u_3	1	0	1	0	1	0	0
u_2, u_4	1	0	0	0	0	0	0
u_3, u_4	1	0	0	0	0	0	0
u_1, u_2, u_3	0	0	0	0	0	0	0
u_1, u_2, u_4	0	0	0	0	0	0	0
u_1, u_3, u_4	0	0	0	0	0	0	0
u_2, u_3, u_4	1	0	0	0	0	0	0
u_1, u_2, u_3, u_4	0	0	0	0	0	0	0
S	7	3	3	1	3	0	0

Теперь составим матрицу D и рассчитаем векторы B и Y для рассматриваемой системы. Результаты в транспонированном виде приведены в таблице 3.

Исходя из получившихся значений в соответствии с п. 7 алгоритма в первую очередь мы должны свернуть критерии-атаки a_1 и a_3 , а затем – результат их свёртки с a_2 . Получившееся дерево комплексной оценки приведено на рис. 3.

Далее можно переходить к получению оценки интегрального риска. Предположим, что в ходе аудита безопасности риск реализации атаки a_1 был оценён как высокий (3 балла), а риски реализации атак a_2 и a_3 – как низкие (1 балл).

Таблица 3. Матрица D и значения векторов B и Y

	a_1	a_2	a_3	a_1, a_2	a_1, a_3	a_2, a_3	a_1, a_2, a_3
u_1		1		1	0	1	1
u_2	1		1	1	1	1	1
u_3	1		1	1	1	1	1
u_4	1	1		1	1	1	1
u_1, u_2	0	0	0	0	0	0	0
u_1, u_3	0	0	0	0	0	0	0
u_1, u_4	0	1	0	1	0	1	1
u_2, u_3	1	0	1	1	1	1	1
u_2, u_4	1	0	0	1	1	0	1
u_3, u_4	1	0	0	1	1	0	1
u_1, u_2, u_3	0	0	0	0	0	0	0
u_1, u_2, u_4	0	0	0	0	0	0	0
u_1, u_3, u_4	0	0	0	0	0	0	0
u_2, u_3, u_4	1	0	0	1	1	0	1
u_1, u_2, u_3, u_4	0	0	0	0	0	0	0
B	7	3	3	9	7	6	9
Y	1	1	1	1 / 9	3 / 7	0	0

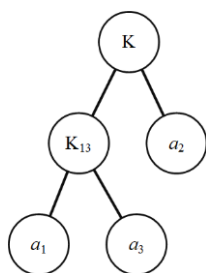


Рис. 3. Дерево комплексной оценки

Тогда при использовании стандартной матрицы свертки (см. рис. 2) получим $K_{13} = 3$ и $K = 3$, следовательно интегральный риск системы оценивается как высокий. Теперь предположим, что ущерб в результате реализации атаки a_1 не очень велик и матрица свёртки критерия K_{13} имеет следующий вид (см. рис. 4):

K_1	1	3	2	1
	2	3	2	1
	3	3	3	2
		3	2	1
	K_3			

Рис. 4. Пример матрицы свёртки критерия K_{13}

Матрицу свёртки критерия K оставим прежней. Тогда $K_{13} = 2$ и $K = 2$, и интегральный риск системы оценивается как средний.

3.2. ПОСТРОЕНИЕ ДЕРЕВА КОМПЛЕКСНОЙ ОЦЕНКИ ДЛЯ SMART-СИСТЕМ НА БАЗЕ «ИНТЕРНЕТА ВЕЩЕЙ»

Теперь применим предлагаемый алгоритм для построения дерева комплексной оценки интегрального риска SMART-системы на базе «Интернета вещей». Сформируем перечень типичных уязвимостей U и атак A на основе работ [17, 29]. Результаты представлены в таблицах 4 и 5.

Таблица 4. Перечень распространённых уязвимостей «Интернета вещей»

U	Уязвимость	Ссылка
u_1	Гетерогенная архитектура	[17]
u_2	Недостаточная физическая безопасность	[29]
u_3	Недостаточный сбор энергии	[29]
u_4	Ненужные открытые порты	[29]
u_5	Устаревшие протоколы	[17]
u_6	Слабое шифрование	[17, 29]
u_7	Ограниченный объем памяти и процессора	[17]
u_8	Небезопасные приложения	[17]
u_9	Недостаточный контроль доступа	[29]
u_{10}	Плохая аутентификация	[17, 29]
u_{11}	Слабые методы программирования	[29]
u_{12}	Неправильная возможность обновления ПО	[17, 29]
u_{13}	Недостаточные механизмы аудита	[29]

Таблица 5. Перечень атак в направлении «Интернета вещей»

А	Вид атаки	Затрагиваемые критерии ИБ (Конфиденциальность, Целостность, Доступность)	Комментарии
a_1	Заражение	К, Ц, Д	Вредоносное ПО, программы-вымогатели, вредоносные устройства, вредоносный код, заражение трояном
a_2	Подслушивание	К	Прослушивание, сканирование, сбор голосовых данных
a_3	Перегрузка возможностей системы	Д	Отказ в обслуживании, переполнение буфера, атака срыва стека, DoS, DDoS
a_4	Кража данных	К	Фишинг
a_5	Перенаправление	К, Д	Спуфинг (IP - , MAC - , DNA -), имперсонация, атака посредника
a_6	Модификация	Ц, Д	Атаки на модификацию фирменного системного или программного обеспечения
a_7	Ввод	Ц	Атака с вводом ложных данных
a_8	Изменение	Д	Изменение топологии сети и снижение производительности сети, напр. атака «воронки»
a_9	Подбор	К, Ц	Подбор ключа путем перебора всех возможных вариантов (Brute-force), перебор по словарю
a_{10}	Атака по сторонним каналам	К	Анализ мощности, атака по времени, атака по ошибкам вычисления
a_{11}	Истощение	Д	Атаки - вампиры, атаки, истощающие батарею, напр. карусельные атаки, растягивающие атаки
a_{12}	Зашумление	Д	Сигналы с помехами: однонаправленными, узкополосными, частичными, широкополосными

Таблица 5 также содержит сведения о критериях информационной безопасности, которые будут нарушены при успешном выполнении соответствующего типа атаки. Построим матрицы соответствия уязвимостей атакам для критериев конфиденциальности, целостности и доступности (рис. 5). Нули в столбцах, соответствующих уязвимости u_1 «гетерогенная архитектура» отражают тот факт, что эта уязвимость не может быть использована для реализации какой-либо атаки сама по себе, а только в комбинации с другими уязвимостями.

	u_1	u_2	u_4	u_5	u_6	u_8	u_9	u_{10}	u_{11}	u_{12}	u_{13}
a_1	0		1	1	1	1		1	1	1	1
a_2	0	1	1	1	1		1		1		1
a_4				1	1		1	1	1		1
a_5	0			1	1	1					1
a_9					1			1			
a_{10}		1			1						

а) Соответствие уязвимостей атакам на конфиденциальность

	u_1	u_2	u_4	u_5	u_6	u_8	u_{10}	u_{11}	u_{12}	u_{13}
a_1	0		1	1	1	1	1	1	1	1
a_6		1						1	1	1
a_7							1	1		1
a_9					1		1			

б) Соответствие уязвимостей атакам на целостность

	u_1	u_2	u_3	u_4	u_5	u_6	u_7	u_8	u_{10}	u_{11}	u_{12}	u_{13}
a_1	0			1	1	1		1	1	1	1	1
a_3	0			1	1		1			1	1	1
a_5	0				1	1		1				1
a_6		1								1	1	1
a_8	0		1						1			
a_{11}	0		1									
a_{12}	0		1									

в) Соответствие уязвимостей атакам на доступность

Рис. 5. Соответствие типичных уязвимостей SMART-системы на основе Интернета вещей типовым атакам на этот класс систем

Отметим, что для технических систем, к которым относятся SMART-системы на основе «Интернета вещей», установление таких соответствий не представляет трудности. Для других классов сложных систем идентификация соответствий может быть отдельной задачей, требующей привлечения экспертов и использования специальных методов согласования экспертных оценок.

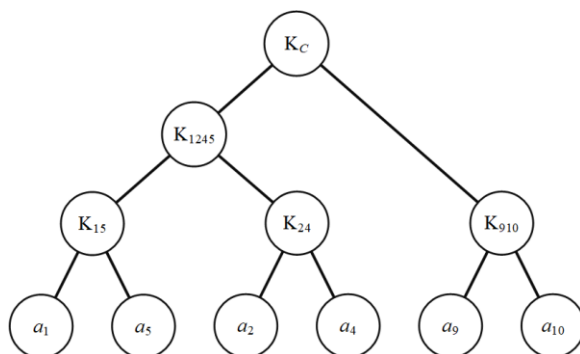
Построим матрицы E_C , E_I , E_A для поддеревьев конфиденциальности, целостности и доступности. Их размерность составит 63×2047 , ..., ... соответственно. Такую же размерность будут иметь и матрицы D_C , D_I , D_A . Рассчитаем векторы S_C , S_I , S_A , B_C , B_I , B_A и, наконец, Y_C , Y_I , Y_A . В таблице 6 приведён фрагмент вектора Y_C , включающий показатели зависимостей всех пар атак на конфиденциальность.

Таблица 6. Рассчитанные показатели попарной зависимости атак на конфиденциальность; округление до 4 знака после запятой

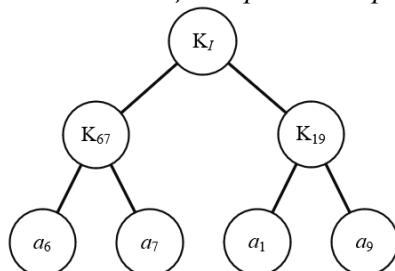
	Показатель зависимости Y_C	Порядок свёртки
a_1, a_2	0,0882	
a_1, a_4	0,0571	
a_1, a_5	0,0587	3
a_1, a_9	0,0059	
a_1, a_{10}	0,0019	
a_2, a_4	0,1080	2
a_2, a_5	0,0517	
a_2, a_9	0,0039	
a_2, a_{10}	0,0118	
a_4, a_5	0,0805	
a_4, a_9	0,0476	
a_4, a_{10}	0,0154	
a_5, a_9	0,0303	
a_5, a_{10}	0,0303	
a_9, a_{10}	0,2000	1

На основе значений элементов векторов Y_C , Y_I , Y_A построим поддерева комплексной оценки рисков конфиденциальности, целостности и доступности. Как видно из таблицы 6, в поддереве конфиденциальности пары (a_9, a_{10}) , (a_2, a_4) и (a_1, a_5) будут свёрнуты в первую очередь. Затем будут свёрнуты промежуточные критерии K_{15} и K_{24} с показателем зависимости, приблизительно равным 0,005 (для комбинации атак a_1, a_2, a_4, a_5), являющиеся результатом свёртки пар (a_1, a_5) и (a_2, a_4) соответственно. Наконец, комплексная оценка риска конфиденциальности получается как свёртка промежуточных критериев K_{910} и K_{1245} , являющихся в свою очередь свёртками соответственно (a_9, a_{10}) и (K_{15}, K_{24}) .

Аналогичным образом строятся поддерева комплексной оценки риска целостности и доступности. Результаты построения приведены на рис. 6, 7.



а) Поддерево комплексной оценки риска конфиденциальности



б) Поддерево комплексной оценки риска целостности

Рис. 6. Поддерева комплексной оценки риска конфиденциальности и целостности

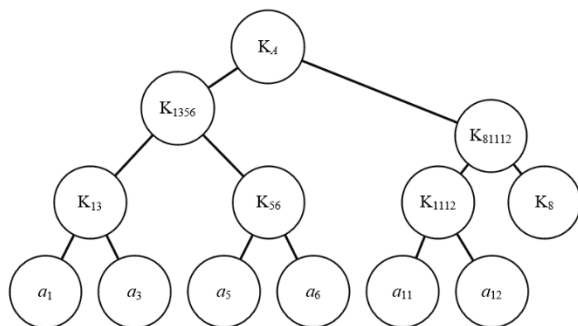


Рис. 7. Поддерево комплексной оценки риска доступности

4. Заключение

В настоящей работе описан алгоритм формирования структуры дерева комплексной оценки интегрального риска сложной информационной системы. В его основе лежит принцип агрегирования критериев по их взаимной зависимости. Предложенный в более ранней работе для свёртки стандартных критериев информационной безопасности – конфиденциальности, целостности и доступности, – он был распространён на задачу построения соответствующих им поддеревьев комплексной оценки. Базовыми критериями при этом выступают типичные для рассматриваемого класса систем атаки, а связь определяется на основе общих для них типовых уязвимостей. В данной работе алгоритм был применён для построения дерева комплексной оценки интегрального риска SMART-системы на основе «Интернета вещей», но предложенный подход может быть распространён на любые сложные системы с идентифицируемыми перечнями уязвимостей и атак.

Дальнейшим развитием данной работы видится модификация метода комплексного оценивания для учёта различных видов неопределённости.

Литература

1. АБДУЛОВА Е.А., КАЛАШНИКОВ А.О. *К вопросу управления рисками критической информационной инфраструктуры* // Управление развитием крупномасштабных систем (MLSD'2021): Труды 14-й Международной конференции. – 2021. – С. 1275–1282.
2. АЛЕКСЕЕВ А.О. *Исследование устойчивости механизмов комплексного оценивания к стратегическому поведению агентов (на примере согласования политики организации в области риск-менеджмента)* // Прикладная математика и вопросы управления. – 2019. – №4. – С. 136–154.
3. АЛЕКСЕЕВ А.О., КАТАЕВА Т.А. *Применение механизмов комплексного оценивания и матричных неанонимных обобщенных медианных механизмов согласования интересов агентов* // Вестник Южно-Уральского государственного университета. Серия «Компьютерные технологии, управление, радиоэлектроника». – 2021. – №3. – С. 75–89.
4. АЛЕСКЕРОВ Ф.Т., ЯКУБА В.И. *Метод порогового агрегирования трехградационных ранжировок* // Доклады академии наук. – 2007. – Т. 413, №2. – С. 181–183.
5. БАКЕЕВ Д.Ш., ТИШИНА Н.А. *Программная реализация оценки рисков безопасности информации на основе гибридного метода* // Приоритетные направления инновационной деятельности в промышленности. Сборник научных статей по итогам пятой международной научной конференции. – 2020. – Т. 2. – С. 6–12.
6. БАРКАЛОВ С.А., НОВИКОВ Д.А., НОВОСЕЛЬЦЕВ В.И. и др. *Модели управления конфликтами и рисками* / Под ред. Д.А. Новикова]. – Воронеж: Научная книга, 2008. – 495 с.
7. БЕЗЗАТЕЕВ С.В., ЕЛИНА Т.Н., МЫЛЬНИКОВ В.А. и др. *Методика оценки рисков информационных систем на основе анализа поведения пользователей и инцидентов информационной безопасности* // Научно-технический вестник информационных технологий, механики и оптики. – 2021. – Т. 21, №4. – С. 553–561.

8. ВЛАСОВА Е.А., КАРПОВ Ю.А., ТАРАСОВ Б.В. *Построение дерева сверток для комплексной оценки на основе матрицы парных сравнений критериев* // Вестник Воронежского государственного технического университета. – 2009. – Т. 5, №10. – С. 187–191.
9. *ГОСТ Р 51901.1-2002 Менеджмент риска. Анализ риска технологических систем.* – Официальное издание. – М.: ИПК Изд-во стандартов, 2002 год.
10. ЗИМА В.М., КРЮКОВ Р.О., КРАВЧУК А.В. *Методика оценивания информационных рисков на основе анализа уязвимостей* // Вопросы оборонной техники. Серия 16: Технические средства противодействия терроризму. – 2019. – №11-12. – С. 36–46.
11. КАЛАШНИКОВ А.О., АНИКИНА Е.В. *Модели управления информационными рисками сложных систем* // Информационная безопасность. – 2020. – Т. 23, №2. – С. 191–202.
12. КАЛАШНИКОВ А.О. *Управление информационными рисками организационных систем: механизмы комплексного оценивания* // Информационная безопасность. – 2016. – Т. 3, №1. – С. 315–322.
13. КИСЕЛЕВА Т.В., МАСЛОВА Е.В. *Классификация рисков ИТ-сервисов и способы оценивания вероятностей их возникновения* // ИТНОУ: информационные технологии в науке, образовании и управлении. – 2020. – №1(15). – С. 67–71.
14. КОЛОСОК И.Н., ГУРИНА Л.А. *Оценка рисков управления киберфизической ЭЭС на основе теории нечетких множеств* // Методические вопросы исследования надежности больших систем энергетики. – 2019. – Т. 1, №70. – С. 238–247.
15. НОВИКОВ Д.А. *«Когнитивные игры»: линейная импульсная модель* // Проблемы управления. – 2008. – №3. – С. 14–22.
16. РЕЙ А.С., КАЛАШНИКОВ А.О. *Комплексная оценка информационных рисков. I: Краткий обзор подходов и методов* // Управление большими системами: сборник трудов. – 2024. – №110. – С. 68–86.
17. ABDULLAH A.A., WALEED A., MALEBARY S. et al. *A review of cyber security challenges attacks and solutions for Internet of Things based smart home* // Int. J. Comput. Sci. Netw. Secur. – 2019. – Vol. 9, No. 9. – P. 139–146.

18. AKINROLABU O., NURSE J.R.C., MARTIN A. et al. *Cyber risk assessment in cloud provider environments: Current models and future needs* // Computers & Security. – 2019. – Vol. 87. – P. 101600.
19. BOLBOT V., THEOTOKATOS G., BOULOUGOURIS E. et al. *A novel cyber-risk assessment method for ship systems* // Safety science. – 2020. – P. 104908.
20. ERSHADI M.J., FOROUZANDEH M. *Information Security Risk Management of Research Information Systems: A hybrid approach of Fuzzy FMEA, AHP, TOPSIS and Shannon Entropy* // J. Digit. Inf. Manag. – 2019. – Vol. 17, No. 6. – P. 321.
21. GUNES B., KAYISOGLU G., BOLAT P. *Cyber security risk assessment for seaports: A case study of a container port* // Computers & Security. – 2021. – Vol. 103. – P. 102196.
22. HÄCKEL B. *Assessing IT availability risks in smart factory networks* // Business Research. – 2019. – Vol. 12, No. 2. – P. 523–558.
23. HAN C.H., HAN C.H. *Semi-quantitative cybersecurity risk assessment by blockade and defense level analysis* // Process Safety and Environmental Protection. – 2021. – Vol. 155. – P. 306–316.
24. HE W., LI H., LI J. *Unknown vulnerability risk assessment based on directed graph models: a survey* // IEEE Access. – 2019. – Vol. 7. – P. 168201–168225.
25. *ISO/IEC 27005:2022(EN) Information security, cybersecurity and privacy protection — Guidance on managing information security risks* [Электронный ресурс]. – Режим доступа: <https://www.iso.org/obp/ui/en/#iso:std:iso-iec:27005:ed-4:v1:en>.
26. KIOSKLI K., POLEMI N. *A Socio-Technical Approach to Cyber-Risk Assessment* // World Academy of Science, Engineering and Technology Int. Journal of Electrical and Computer Engineering. – 2020. – Vol. 14, No. 10. – P. 305–309.
27. KORNEEV N.V., KORNEEVA J.V., YURKEVICHYUS S.P. et al. *An Approach to Risk Assessment and Threat Prediction for Complex Object Security Based on a Predicative Self-Configuring Neural System* // Symmetry. – 2022. – Vol. 14, No. 1. – P. 102.
28. KRISPER M., DOBAJ J., MACHER. G. et al. *RISKEE: A risk-tree based method for assessing risk in cyber security* // European Conf. on Software Process Improvement. – 2019. – P. 45–56.

29. NESHENKO N., BOU-HARB E., CRICHIGNO J. et al. *Demystifying IoT Security: An Exhaustive Survey on IoT Vulnerabilities and a First Empirical Look on Internet-Scale IoT Exploitations* // IEEE Communications Surveys & Tutorials. – 2019. – Vol. 21, No. 3. – P. 2702–2733.
30. NTAFOUKAS K., MCCRUM D.P., PASQUALE L. *A Socio-Technical Approach to Cyber-Risk Assessment* // A cyber-physical risk assessment approach for Internet of Things enabled transportation infrastructure. – 2022. – Vol. 12, No. 18. – P. 9241.
31. PALKO D., BABENKO T., BIGDAN A. et al. *Cyber Security Risk Modeling in Distributed Information Systems* // Appl. Sci. – 2023. – Vol. 13, No. 4. – P. 2393.
32. RIOS E., REGO A., ITURBE E. et al. *Continuous quantitative risk management in smart grids using attack defense trees* // Sensors. – 2020. – Vol. 20. – P. 4404.
33. SCHMITZ C., PAPE S. *LiSRA: Lightweight security risk assessment for decision support in information security* // Computers & Security. – 2020. – Vol. 90. – P. 101656.
34. SUBRIADI A.P., NAJWA N.F. *The consistency analysis of failure mode and effect analysis (FMEA) in information technology risk assessment* // Heliyon. – 2020. – Vol. 6, No. 1. – e03161.
35. TUSHER H.M., MUNIM Z.H., NOTTEBOOM T.E. et al. *Cyber security risk assessment in autonomous shipping* // Maritime Economics & Logistics. – 2022. – Vol. 24, No. 2. – P. 208–227.
36. TUSHER H.M., MUNIM Z.H., NOTTEBOOM T.E. et al. *Development of the mechanism of assessing cyber risks in the internet of things projects* // 12th Conf. on Internet of Things, Smart Spaces, and Next Generation Networks and Systems., ruSMART-2019. St. Petersburg: Springer Int. Publishing. – 2019. – P. 481–494.
37. WANG Y., WANG Y.-H., QIN H. et al. *A systematic risk assessment framework of automotive cybersecurity* // Automotive Innovation. – 2021. – Vol. 4. – P. 253–261.
38. WANG Y., XUE W., ZHANG A. *Application of Big Data Technology in Enterprise Information Security Management and Risk Assessment* // Journal of Global Information Management (JGIM). – 2023. – Vol. 31, No. 3. – P. 1–16.

COMPLEX INFORMATION RISKS ASSESSMENT. II: ALGORITHM FOR IDENTIFYING THE STRUCTURE OF THE TREE OF INTEGRATED INFORMATION RISK ASSESSMENT

Anastasiya Rey, V.A. Trapeznikov Institute of Control Sciences of RAS, Moscow, Junior Researcher (a.rey@ipu.ru).

Alexander Shiroky, V.A. Trapeznikov Institute of Control Sciences of RAS, Moscow, Ph.D. in Physics and Mathematics Science, Senior Researcher (shiroky@ipu.ru).

Abstract: When assessing information security risks, it is essential to take into account the various types of uncertainties that are inherent in information systems. Current methods and algorithms for risk assessment may not account for all of these uncertainties, which can lead to inaccurate risk estimates. Therefore, it becomes necessary to develop a new or adapt an existing method for risk assessment that considers all types of uncertainty specific to the class of system under consideration. In this paper, we build on our previous idea of using an integrated assessment method to evaluate information risks. This method aggregates assessments of information systems based on standard information security criteria such as confidentiality, integrity, and availability. By incorporating these criteria, we aim to obtain more accurate and reliable risk estimates that take into account all relevant uncertainties. In the first part of the work, we demonstrated that this method, with appropriate modifications, allows for taking into account all necessary types of uncertainty. We propose an algorithm for identifying the structure of an integrated assessment tree based on the principle of combining related criteria. We demonstrate the efficiency of the algorithm by building risk assessment trees for confidentiality, integrity, and accessibility in SMART systems based on the Internet of Things using this approach.

Keywords: complex information systems, integral risk, complex assessment, accounting for uncertainty.

УДК 004.056.5

ББК 16.8

DOI: 10.25728/ubs.2024.111.4

*Статья представлена к публикации
членом редакционной коллегии И.В. Бычковым.*

Поступила в редакцию 18.04.2024.

Опубликована 30.09.2024.

СНИЖЕНИЕ РАЗМЕРНОСТИ ЗАДАЧИ НАХОЖДЕНИЯ КРИТИЧЕСКИХ УЗЛОВ СЕТИ

Крыгин А. А.¹, Тарасова С. М.²
(ФГБУН Институт проблем управления
им. В.А. Трапезникова РАН, Москва)

Одним из классов задач, решаемых при оценке устойчивости инженерной сети, являются задачи нахождения критических узлов. В ряде постановок эта задача формулируется как нахождение такого подмножества узлов заданной мощности (критических узлов), при выходе из строя которых всей сети будет нанесен максимальный ущерб. И наиболее частый способ оценки ущерба в такой постановке – определение количества связанных пар узлов в сети с исключенными критическими узлами. Для таких узлов, которые соответствуют минимуму количества связанных пар, требуются проведение дополнительных мер по повышению надежности и безопасности. Ряд методов решения задачи нахождения критических узлов использует сведение ее к эквивалентной задаче линейного программирования. Основной проблемой этого подхода является большая размерность задачи, и, как следствие, высокая вычислительная сложность ее решения. В работе проводится исследование различных характеристик вершин графовой модели сети, анализ значений которых позволит заранее установить факт принадлежности вершины к подмножеству критических или наоборот, к подмножеству некритических узлов. Благодаря этому можно сформировать дополнительные ограничения, снижающие размерность задачи линейного программирования и ее вычислительную сложность, что позволит находить критические узлы в инженерных сетях с большим количеством объектов за приемлемое время. В процессе исследования было решено множество различных подзадач, поэтому в работе описывается первая, подготовительная его часть.

Ключевые слова: устойчивость электрических сетей, поиск критических узлов, графовые модели сетей.

1. Введение

Наиболее простой математической моделью сети является ее представление в виде неориентированного и невзвешенного графа. Такая модель применима практически ко всем видам сетей и является базовой (начальной) в задачах оценки их устойчивости.

¹ Андрей Александрович Крыгин, к.т.н. (andreyakr@yandex.ru).

² Софья Михайловна Тарасова, техник (tarasva_sofia@mail.ru).

Объектам сети соответствуют вершины графа, а наличие какого-либо отношения (в зависимости от вида сети) между объектами соответствует ребрам. При этом все вершины считаются равноправными, отдельно не выделяются, например, вершины – источники ресурса и вершины-потребители.

Одним из классов задач, решаемых при оценке устойчивости сети, являются задачи нахождения критических ребер и узлов (Critical Link Disruptor (CLD) and Critical Node Disruptor (CND) optimization problems). Постановка этих задач заключается в нахождении такого подмножества узлов (или ребер), при выходе из строя которых всей сети будет нанесен максимальный ущерб.

В общем виде задачу нахождения критических узлов для графовой модели сети можно сформулировать так. Заданы связный, неориентированный граф без петель и кратных ребер $G(V, E)$, $|V| = n$, $|E| = m$, количество удаляемых вершин q и функция-критерий $D(G)$, областью определения которой является множество графов, а областью значений – вещественные числа. Обозначим W – подмножество V , $|W| \leq q$ и G_W – подграф графа G , получившийся при удалении из V подмножества W и инцидентных ребер. Тогда требуется найти такое подмножество W' , при котором $D(G_{W'})$ достигает своего минимума (или максимума, в зависимости от функции-критерия).

Чаще всего в качестве критерия $D(G)$ выбирают одну из трех функций:

- количество компонент связности (определяется максимум);
- размер наибольшей компоненты связности (определяется минимум);
- количество связных пар вершин (определяется минимум).

Наиболее часто в работах применяется третий критерий – количество связных пар вершин, который также используется и в данной работе.

Существуют и другие постановки задачи поиска критических узлов, например, задача нахождения критических узлов с учетом дистанции (Distance-Based Critical Node Problem)

[10, 11, 16, 21] и β -вершинная задача нахождения критических узлов (β -vertex disruptor problem) [9, 12, 17].

В качестве критерия $D(G)$ в постановке задачи нахождения критических узлов с учетом дистанции выбрано количество связанных пар вершин, расстояние между которыми не превосходит заданного параметра k . То есть минимизируется количество пар вершин, для которых в графе с удаленными q вершинами существует путь, содержащий не более k ребер. Во многих работах, посвященных этой задаче, используются различные методы сведения ее к эквивалентным задачам целочисленного линейного программирования.

В некотором смысле β -вершинную задачу нахождения критических узлов можно назвать обратной. Если граф сети после удаления q вершин остается связным, то количество связанных пар в нем равно $\frac{(n-q)(n-q-1)}{2}$. Обозначим M – количество связанных пар вершин в таком графе (связном или несвязном) и $\beta = \frac{2M}{(n-q)(n-q-1)}$. Тогда задачу можно сформулировать так: необходимо найти минимальное по мощности подмножество вершин, при удалении которых из графа $\beta < \beta_0$, где β_0 – заданный параметр. Эта задача тесно связана с «прямой» задачей; переход от одной задачи к другой в ее эквивалентной формулировке задачи линейного программирования, можно осуществить, поменяв местами целевую функцию и одно из ограничений.

Также активно исследуются задачи и для других критериев $D(G)$, например, в [13] рассматривается размер наибольшей компоненты связности.

1.1. Точные методы решения задачи нахождения критических узлов по критерию наименьшего количества связанных пар вершин

В [19] была доказана NP-полнота задачи нахождения критических узлов. На сегодняшний день наиболее эффективным методом нахождения строго оптимального (точного) решения задачи поиска критических узлов является сведение ее к эквивалентной

задаче линейного программирования (ЛП). За исключением метода полного перебора, каких-либо других точных методов решения этой задачи, включая алгоритмы на графах и сетях, не известно. Более того, в [18] доказана теорема о том, что нахождение приближенного решения, достаточно близкого к точному также является NP-трудной задачей, задана метрика «близости» и указана ее нижняя граница.

Рассмотрим постановку задачи ЛП для критерия минимального количества связных пар. Введем следующие булевские переменные:

$u_{i,j}$ – переменная, показывающая связаны ли вершины i и j . $u_{i,j} = 0$, если вершины i и j не связаны, $u_{i,j} = 1$, если вершины i и j связаны;

v_i – переменная, показывающая будет ли вершина i в оптимальном решении удалена из графа. $v_i = 0$, если вершина i не удаляется из графа, $v_i = 1$, если вершина i удаляется из графа.

Тогда постановку задачи в терминах ЛП можно задать так:

$$\begin{aligned} \min \sum_{i,j \in V} u_{i,j} \\ u_{i,j} + v_i + v_j \geq 1, \forall (i, j) \in E \\ u_{i,j} + u_{j,w} - u_{w,i} \leq 1, \forall (i, j, w) \in V \\ u_{i,j} - u_{j,w} + u_{w,i} \leq 1, \forall (i, j, w) \in V \\ -u_{i,j} + u_{j,w} + u_{w,i} \leq 1, \forall (i, j, w) \in V \\ \sum_{i \in V} v_i \leq q. \end{aligned}$$

Целевая функция $\sum_{i,j \in V} u_{i,j}$ равна количеству всех связных пар вершин и она минимизируется.

Первая группа неравенств $u_{i,j} + v_i + v_j \geq 1$, называемая в литературе «неравенством треугольника», записывается для каждого ребра и задает структуру графа. Если соединенные ребром вершины i и j не удалены, то они всегда связны по ребру. Это утверждение эквивалентно вышеприведенному неравенству.

Для трех произвольных вершин i, j и w справедливо утверждение: если i связана с j , и j связана с w , то i связана с w . Это утверждение эквивалентно выражению: $u_{i,j} + u_{j,w} + u_{w,i} \neq 2$. Систему неравенств, эквивалентную этому выражению можно записать так:

$$u_{i,j} + u_{j,w} - u_{w,i} \leq 1,$$

$$u_{i,j} - u_{j,w} + u_{w,i} \leq 1,$$

$$-u_{i,j} + u_{j,w} + u_{w,i} \leq 1.$$

Эта группа неравенств записывается для всех троек вершин.

Последнее неравенство $\sum_{i \in V} v_i \leq q$ задает общее количество удаляемых вершин.

В [22] доказана теорема о том, что исходная постановка задачи и ее формулировка в виде задачи ЛП эквивалентны. В [17] показаны методы сведения задачи нахождения критических узлов к задаче ЛП для оставшихся двух критериев (количество компонент и размер максимальной компоненты) и доказана эквивалентность исходной задачи и задачи ЛП.

В большинстве случаев у задачи нахождения критических узлов существует несколько оптимальных решений. Если требуется найти все решения (или заданное количество решений), то поступают следующим образом. Будем обозначать буквой Q (с индексом или без) одно из решений задачи, т.е. подмножество вершин, мощность которого не превосходит q такое, что не существует другого подмножества $R, |R| \leq q$ и $D(G_Q) > D(G_R)$. Находят первое решение Q_1 и фиксируют значение функции-критерия $D_{opt} = D(G_{Q_1})$. После этого в систему неравенств добавляют неравенство, исключающее найденное решение: $\sum_{i \in Q_1} v_i < q$ и повторно запускают решение задачи ЛП. Если на k -й итерации значение $D(G_{Q_k})$ станет больше D_{opt} , то это свидетельствует о том, что все оптимальные решения найдены.

1.2. Методы, близкие к точным

Из постановки задачи ЛП видно, что наибольшую долю неравенств составляют тройки $u_{i,j} + u_{j,w} + u_{w,i} \neq 2$, число этих

троек равно $C_n^3 = O(n^3)$. И, например, для графа с 120-ю вершинами размер файла описания задачи ЛП занимает около 4-х Гб. За 2 недели расчета (процессор Intel Xeon E5-2620v4@2.1ghz, оперативная память 64 Гб, программа решения задачи ЛП GLPK) не было найдено ни одного решения для такой задачи. Использование вычислительных кластеров позволяет решать задачу ЛП за приемлемое время (несколько часов) для графов с $n < 200$ [19]. Для определения критических узлов у графов бóльшей размерности или для решения задачи за меньшее время появились эвристические методы поиска решений, близких к точным.

О количестве работ в этой области говорит тот факт, что появились работы, каталогизирующие, сопоставляющие и сравнивающие между собой уже имеющиеся методы. В [14] рассматривается несколько десятков методов решения задачи нахождения критических узлов в различных постановках. Обзорная работа [20] посвящена приближенным методам решения задачи для вышеописанных трех основных критериев.

В рассмотренных работах одна из основных схем решения задачи методами, близкими к точным, включает в себя два этапа. На первом этапе предлагается алгоритм нахождения «первоначального решения» из q вершин. Первоначальное решение – это подмножество из q вершин, которое не обязательно является точным решением, будем обозначать первоначальные и текущие решения буквой S . Далее это решение «обогащают» методом парного обмена.

В данном случае функция-критерий $D(G_S) = \sum_{i,j \in V \setminus S} u_{i,j}$, ищется минимум. Пусть S – текущее решение, в начале работы инициализируем S как первоначальное решение. Алгоритм парного обмена заключается в следующем:

1. Выбирается произвольная вершина $i \in S$.
2. Среди всех вершин из $V \setminus S$ определяется вершина j такая, что значение D будет минимальным, если в решении S вместо вершины i брать вершину j и $D(G_{S \cup \{j\} \setminus \{i\}}) < D(G_S)$. Если такой вершины нет, то переход к п.1.
3. Из S исключается i и добавляется j .

4. Выбирается следующая вершина из S .

5. Переход к п.2.

6. Алгоритм заканчивает работу по истечении заданного времени или когда для любой вершины $i \in S$ и любой вершины $j \in V \setminus S$ выполняется $D(G_{S \cup \{j\} \setminus \{i\}}) > D(G_S)$.

С помощью таких методов удалось добиться того, что в 40% случаев полученное решение совпадало с точным.

Скорость работы алгоритма и близость полученного решения к точному напрямую зависят от близости первоначального решения к точному решению. Хорошие результаты дают методы, в которых это решение получается с помощью исходной задачи ЛП, но с ослабленными условиями, что позволяет решать ее за приемлемое время.

В [18] используется тот факт, что решение целочисленной задачи ЛП во много раз медленней, чем решение той же задачи ЛП, но без ограничения на целые значения переменных. Моделирование показало, что, например, для графа из 35 вершин решение не целочисленной задачи занимает около 2–3 секунд, а ограничение на целочисленность увеличивает это время до 32 минут, т.е. более чем в 600 раз. Авторы [18] решают задачу ЛП, заменив условие $v_i \in \{0, 1\}$ на условие $0 \leq v_i \leq 1$ и аналогично для переменных $u_{i,j}$. После этого первоначальное решение набирается из переменных v_i , значения которых наиболее близки к единице.

Также есть работы, в которых используется следующий факт: время решения задачи ЛП возрастает с ростом q . Первоначальное решение получалось следующим образом. Требуемое количество раз решается задача ЛП для малого ($q = 1, 2$) количества удаляемых вершин, после чего полученные решения объединяются.

Помимо описанной двухэтапной схемы решения существуют и другие подходы. Например, в [8] применяется необычный способ: условия задачи сначала усиливаются, на подмножество критических вершин накладываются дополнительные требования, а потом задача в новой постановке решается простым приближенным методом. Для этого водится новая метрика фрагментации

графа и далее формулируется смешанно-целочисленная задача квадратичного программирования, которая одновременно минимизирует количество связанных пар вершин и максимизирует фрагментацию графа. Эта задача приближенно решается с помощью жадного алгоритма. В [15] описывается способ обучения нейронной сети для определения критических вершин. Авторы [6] применяют генетический алгоритм и предлагают три различных подхода его начальной инициализации. В [23] используется мембранный эволюционный алгоритм.

Все рассмотренные в этом разделе алгоритмы являются эвристическими и не дают гарантии, что найденные решения являются оптимальными, а также не дают гарантированной оценки близости полученных решений к точному.

1.3. Постановка цели исследования.

В данном исследовании рассматриваются точные методы и методы, близкие к точным, в которых используется решение вышеописанной задачи ЛП для нахождения критических узлов сети по критерию наименьшего количества связанных пар вершин. Исследуется возможность сократить размерность задачи ЛП, используя характеристики вершин. К проведению этого исследования подтолкнуло следующее рассуждение. Рассмотрим некоторый граф, имеющий вершину степени 1. Очевидно, что удаление такой вершины никак не влияет на связность остальных вершин. Следовательно, за исключением одного случая, эта вершина не будет критической и в задаче ЛП можно заменить нулем переменную v_i , соответствующую этой вершине. Исключение происходит, когда при удалении q вершин все вершины графа G_q имеют степень 0 или 1. То есть граф распадается на изолированные вершины и пары смежных вершин. Тогда при нахождении оптимального решения для $q + 1$ удаляемых вершин в него попадут вершины степени 1.

Целью исследования является подбор двух характеристик вершины, обладающих следующим свойством. Упорядочим все вершины графа по значению первой характеристики. Характеристика должна быть такая, что несколько первых вершин почти

всегда будут критическими и в задаче ЛП соответствующие переменные v_i можно заменить единицами. Аналогично последние несколько вершин, упорядоченные по значению второй характеристики, почти никогда не будут критическими и соответствующие переменные v_i можно заменить нулями. Кроме того эта пара характеристик должна максимально сокращать размерность задачи ЛП, среди всех других пар. Для найденных двух характеристик необходимо получить оценку количества вершин, для которых переменные v_i заменяются нулями и единицами. Будем называть такую пару характеристик *оптимальной* и задачу ЛП, в которой заменены нулями и единицами соответствующие переменные v_i , – *сокращенной*. Также будем считать, что в современных «решателях» задач ЛП первым стоит модуль оптимизации и уменьшения количества переменных. Тогда при добавлении в задачу ЛП равенств вида $v_i = 1$ или $v_i = 0$ будет происходить сокращение ее размерности.

Исследование получилось достаточно большим из-за того, что потребовалось решить ряд подготовительных задач. Возникла необходимость выбрать значения и обосновать этот выбор для двух параметров: количества критических вершин и доли решений. Только после этого стало возможным сформулировать корректную постановку исследуемой задачи. Общая схема исследования состоит из следующих пунктов:

1. Определение класса графов, для которых проводится исследование. Нет гарантии того, что для графов различных по своей природе инженерных сетей (например, транспортных и тепловых), искомая пара оптимальных характеристик будет одной и той же. Поэтому исследуемый класс графов нужно определить.

2. Определение количества критических вершин q . Для разных диапазонов значений q , по которым ищутся решения задачи ЛП, искомая пара оптимальных характеристик может различаться. Поэтому для постановки задачи необходимо выбрать этот диапазон и обосновать выбор.

3. Оценка доли решений. Предположим, что для заданного графа и количества критических вершин q имеется 100 решений,

т.е. 100 наборов из q вершин (одна и та же вершина может содержаться в нескольких наборах). И пусть вершина k содержится только в одном наборе. Тогда если объявить вершину k некритической, то будет найдено 99 решений. Возникает вопрос: допустима ли такая ситуация? Будем называть *долей решений* отношение количества решений в сокращенной задаче ЛП к количеству решений в исходной задаче. Тогда необходимо определить, какая доля решений допустима.

4. Постановка задачи.

5. Оценка верхней границы количества вершин, которые почти всегда являются критическими, и аналогично для вершин, которые почти никогда ими не являются. То есть оценка максимального допустимого количества равенств вида $v_i = 0$ и $v_i = 1$, которые можно добавить в задачу ЛП, не нарушая ограничений постановки, без учета значений каких-либо характеристик вершин.

6. Выбор оптимального алгоритма подбора характеристик. Предположим, что для заданного графа, q и двух пар характеристик для первой пары можно добавить 5 равенств вида $v_i = 1$ и 10 равенств вида $v_i = 0$, а для второй, соответственно, 4 и 15. Необходимо определить в каком из этих двух случаев сокращенная задача ЛП будет решена быстрее. Для этого нужно более детально оценить количество переменных и ограничений в сокращенной задаче ЛП для каждого случая.

7. Подбор оптимальной пары характеристик. Оценка для каждой оптимальной характеристики допустимого количества равенств вида $v_i = 0$ и $v_i = 1$, которые можно добавить в задачу ЛП и сравнение полученных значений с их верхними оценками.

8. Проверка результатов на отдельных, не рассматриваемых в исследовании примерах.

В данной работе описывается первая часть исследования, состоящая из выбора класса сетей, определении параметров, постановки задачи и оценки верхних границ (пункты 1–5).

2. Определение класса графов

Помимо указанной выше необходимости выбора класса графов для корректной постановки задачи справедливо следующее утверждение, которое накладывает дополнительное ограничение на выбираемый класс.

Утверждение 1. *В общем случае для произвольных сетей не существует характеристики, позволяющей сократить размерность.*

Доказательство. Введем понятие *полного* решения задачи нахождения критических узлов. Для заданного графа существуют две величины q_{begin} и q_{end} , такие, что для всех $q < q_{begin}$ удаление любых q вершин не нарушает связности оставшихся вершин. И существует такой набор из q_{end} вершин, что при их удалении оставшийся граф распадается на изолированные вершины и количество связных пар в нем равно нулю. Тогда полным решением задачи заключается в нахождении q_{begin} и q_{end} и всех решений для каждого $q_{begin} \leq q \leq q_{end}$.

Для доказательства утверждения достаточно привести пример задачи (т.е. граф и количество удаляемых вершин q) для которой не существует характеристики, сокращающей ее размерность

Рассмотрим следующий граф (рис. 1).

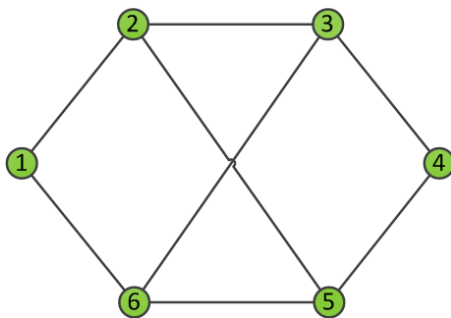


Рис. 1. Пример графа

Полное решение задачи для этого графа выглядит так. При удалении любой единичной вершины связность оставшихся вершин не нарушается. При удалении пары вершин (2, 6) или (3, 5) граф распадается на изолированную вершину 1 или 4 и оставшиеся три связных вершины. При удалении любой другой пары вершин связность оставшихся вершин не нарушается. При удалении тройки вершин (1, 3, 5) или (2, 4, 6) оставшиеся три вершины изолированы. При удалении любой другой тройки вершин связность оставшихся вершин не нарушается.

Покажем, что для этого графа нельзя сократить размерность задачи добавлением равенств вида $v_i = 1$ или $v_i = 0$, в которых вершины выбираются по значению некоторой характеристики.

В изображении данного графа имеются две оси симметрии (вертикальная и горизонтальная), следовательно, любые характеристики вершин 1, 4 совпадают и любые характеристики вершин 2, 3, 5, 6 также совпадают. То есть, с точки зрения выделения подмножества вершин для добавления равенств в задачу ЛП, вершины (1,4) и (2,3,5,6) неразличимы и добавляться могут только следующие группы равенств:

$$v_i = 0 \text{ или } 1; i \in (2, 3, 5, 6) \text{ и}$$

$$v_i = 0 \text{ или } 1; i \in (1, 4).$$

Перебирая эти варианты можно определить, что для $q = 3$ не существует группы равенств, добавление которых приводит к точным решениям.

Сформулируем условия, при которых размерность задачи ЛП будет уменьшаться. Вычислительная сложность современных алгоритмов решения задачи ЛП напрямую зависит от количества переменных и неравенств [4, 7]. Поэтому сокращение размерности будет происходить только при добавлении в задачу ЛП равенств вида $v_i = 1$ или $v_i = 0$. И с другой стороны, например, знание о том, что одна из трех вершин является критической, никак не позволит сократить размерность.

В графе, у которого есть теоретическая возможность снизить размерность задачи ЛП, должны быть вершины, которые почти всегда являются критическими, и вершины, которые по-

что никогда не являются критическими. Определить это можно следующим способом. Рассмотрим для некоторого графа полное решение задачи нахождения критических узлов и определим для каждой вершины частоту ее появления – долю решений, в которых есть эта вершина. Построим следующую диаграмму: на оси абсцисс отрезок $[0; 1]$ разделен на большое количество (20–50) равных отрезков, обозначим эту ось f . Для каждого отрезка на оси ординат откладывается доля вершин (количество таких вершин, разделенное на количество всех вершин), частота которых лежит в диапазоне этого отрезка; обозначим эту ось h . Будем называть эту диаграмму «диаграмма частот появления вершин в решениях». Тогда у графа есть теоретическая возможность добавлять равенства в задачу ЛП, если на краях (в районе 0 и в районе 1) диаграммы имеются ненулевые столбцы.

В качестве исследуемого класса графов были выбраны укрупненные графы магистральных электрических сетей (сетей передачи энергии) государств СНГ, взятые из открытых источников. В [5] и [2] приведены такие схемы для Белоруссии и Таджикистана. Магистральные электрические сети, связывающие в единое целое отдельные регионы и их крупные источники, имеют замкнутую структуру, что хорошо согласуется с критерием наименьшего количества связных пар вершин. В отличие, например, от распределительных сетей, в которых важна связность источников и потребителей ресурса.

Рассматривалось порядка 20 сетей с количеством узлов 30–60. Для каждой сети было найдено полное решение задачи нахождения критических узлов, общее время расчета составило около 7 месяцев. Из-за такой большой продолжительности расчета не было возможности увеличить количество рассматриваемых графов или их размерность. К сожалению, все последующие выводы основаны на этой малой выборке.

На рис. 2 показаны типовые диаграммы частот появления вершин в решениях для выбранного класса сетей. Наличие ненулевых столбцов в районе единицы и нуля свидетельствует о том, что для рассматриваемых сетей есть теоретическая возможность

уменьшить размерность задачи ЛП.

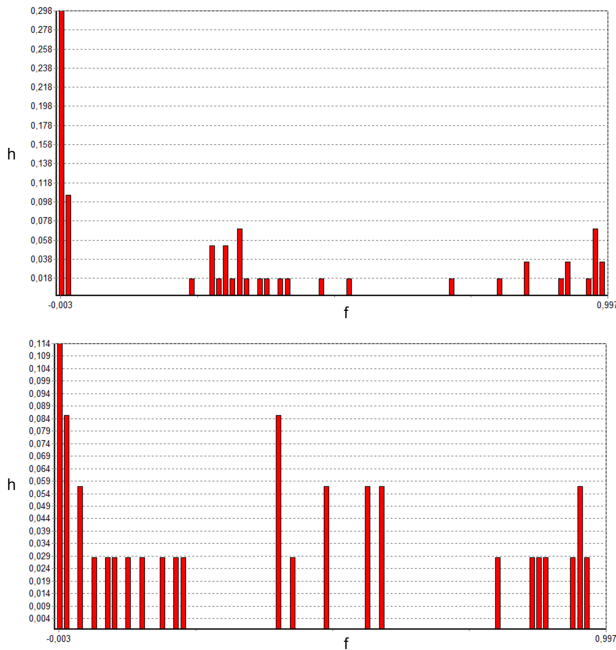


Рис. 2. Диаграммы частот появления вершин в решениях у электрических сетей

В дальнейшем характеристику частоты, с которой вершина встречается в полном решении, будем называть «частотой».

3. Оценка количества критических вершин

Рассмотрим два крайних случая: для заданного графа решается задача нахождения критических узлов при $q = 1$ и при всех $q \leq q_{end}$. Нет никаких гарантий того, что искомая оптимальная пара характеристик в этих случаях будет одной и той же. Будем считать, что задача решается для всех $q \leq q_0$, соответственно, необходимо оценить величину q_0 .

В рассматриваемом случае электрических сетей логично предположить, что после решения задачи ЛП для всех $q \leq q_0$

(нахождения всех решений для каждого q) строится объединение найденных критических вершин; обозначим $P(q_0)$ – объединение всех критических вершин, полученных при решении исходной задачи ЛП для всех $q \leq q_0$, $p = |P(q_0)|$ и $\overline{P(q_0)}, \bar{p}$ – аналогичные величины для сокращенной задачи. Для узлов сети, соответствующих множеству $P(q_0)$, проводятся мероприятия, обеспечивающие повышенные требования по надежности и безопасности. Будем считать, что после проведения этих мероприятий вероятность выхода из строя узла, соответствующего вершине из множества $P(q_0)$, пренебрежимо мала. Учитывая, что на поддержание повышенного уровня надежности и безопасности объектов электросетей требуются значительные ресурсы, оценим величину p как функцию от количества вершин $p(n)$ и с ее помощью оценим величину $q_0(n)$.

Величина $p(n)$ была оценена как $p(n) = \frac{n}{5}$ по следующим соображениям.

1. Будем считать, что количество объектов сети пропорционально площади, которую эта сеть охватывает. Количество потребителей также пропорционально этой площади, и на проведение вышеописанных мероприятий каждый из потребителей выделяет одинаковую сумму. Отсюда следует линейный вид $p(n)$.

2. Рассматривались различные государственные программы (касающиеся не только электросетей), в которых для достаточно большого множества объектов проводились мероприятия с целью улучшения показателей всего множества. Причем мероприятия заключались в улучшении, модернизации или замене отдельных объектов. В разных случаях доля таких объектов колебалась от 5 до 18%. Например, доля выселяемых по реновации домов в Москве составляет 6,5% [3], доля школ, попавших в программу обновления российских школ, — 18,3% [1].

3. Из-за небольшого количества как рассматриваемых сетей, так и вершин в каждой сети эта доля была увеличена до 20%.

Далее для каждого графа выполнялась следующая операция. С помощью последовательного перебора значений $q = 1, 2, \dots$ определялась величина $\tilde{q}$ такая, что мощность $P(\tilde{q})$ не превосхо-

дит $\frac{n}{5}$, а мощность $P(\tilde{q} + 1)$ – превосходит. И среди всех графов определялся максимум величины $\frac{\tilde{q}}{n}$. Оказалось, что этот максимум также равен 0,2; это означает, что у одного из графов есть единственное решение задачи ЛП для найденного $\tilde{q}$ и это решение включает в себя все вершины всех решений при $q < \tilde{q}$. В результате для дальнейшего исследования использовалась оценка $q_0 = \lceil n/5 \rceil$.

4. Оценка доли решений

Обозначим отношение количества решений в сокращенной задаче к количеству решений в исходной задаче μ , $0 \leq \mu \leq 1$. Будем называть решение исходной задачи *найденным*, если это решение также является решением сокращенной задачи, и *ненайденным* – в противном случае. Проведем оценку допустимых значений μ , используя следующее рассуждение.

Рассмотрим следующую ситуацию:

- предполагается, что количество узлов, которые могут одновременно выйти из строя, не превосходит q_0 ;
- в любом решении исходной задачи ЛП для количества вершин, не превосходящих q_0 (найденном или найденном), есть хотя бы одна вершина из множества $\overline{P(q_0)}$.

Тогда можно утверждать, что если одновременно выйдет из строя q_0 вершин, то эти вершины не являются решением задачи ЛП для q_0 вершин и никакое подмножество этих вершин также не является решением задачи ЛП для количества вершин, меньших q_0 . То есть любое подмножество вершин из множества $V \setminus \overline{P(q_0)}$ и мощностью не превосходящее q_0 не является критическим. Будем считать, что в этом случае электрическая сеть защищена в должной степени и из этого свойства оценим величину μ .

Пусть для некоторого q общее число всех решений равно d и количество найденных решений равно $\mu \cdot d$ (рис. 3). Обозначим среднюю частоту, с которой вершина встречается в решениях (рассматриваются только вершины, входящие хотя бы в одно из решений), за f . Пока будем считать, что частоты всех вершин

одинаковы и равны f . То есть во всех решениях встречается $f \cdot d$ экземпляров каждой вершины.

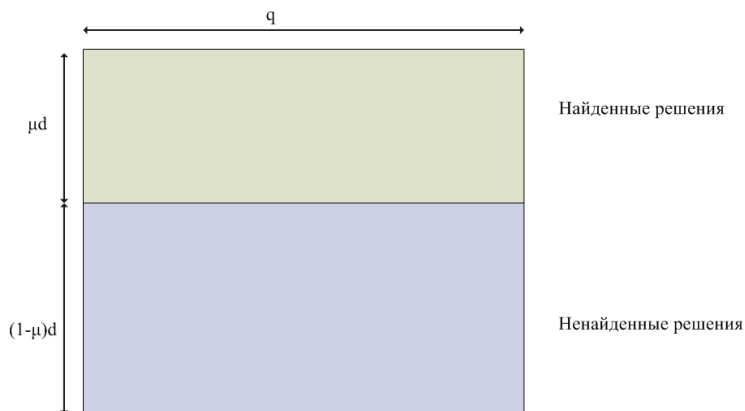


Рис. 3. Схема решений

По условию в каждом из ненайденных решений содержится хотя бы одна вершина из найденных решений. Для оценки μ удобно решать противоположную задачу. Предположим, что среди ненайденных решений существует решение R , ни одна из вершин которого не содержится в найденных решениях. Тогда все экземпляры вершин этого решения также находятся среди ненайденных решений. Из этого условия оценим снизу величину $1 - \mu$; тогда если выбрать μ таким, что $1 - \mu$ будет ниже этой оценки, то можно утверждать, что в любом ненайденном решении содержится хотя бы одна вершина из найденных решений.

Фактически нужно «упаковать» оставшиеся $(f \cdot d - 1) \cdot q$ экземпляров вершин в как можно меньшее количество решений так, чтобы все решения были различны. В каждом решении не может быть больше чем $q - 1$ вершин, совпадающих с вершинами из R . Соответственно, общее количество ненайденных решений можно оценить так:

$$(1 - \mu) \cdot d \geq 1 + \left\lceil \frac{q \cdot (f \cdot d - 1)}{q - 1} \right\rceil.$$

$$\text{Тогда если } \mu > 1 - b_0, \text{ где } b_0 = \frac{1 + \left[\frac{q \cdot (f \cdot d - 1)}{q - 1} \right]}{d},$$

то не существует найденного решения R , ни одна из вершин которого не содержится в найденных решениях.

При различных значениях частот будем рассуждать аналогично. Предположим, что существует найденное решение R , ни одна из вершин которого не содержится в найденных решениях. Тогда нужно «упаковать» $\sum_{i \in R} f_i \cdot d - q$ оставшихся экземпляров вершин в как можно меньшее количество решений. И сравнение с предыдущей оценкой приводит к сравнению выражений $f \cdot q$ и $\sum_{i \in R} f_i$.

Результаты этого сравнения следующие: $f \cdot q$ всегда меньше $\sum_{i \in R} f_i$. На качественном уровне это объясняется тем, что величина f усредняется по всем вершинам, среди которых много вершин с малой частотой, а в выражении $\sum_{i \in R} f_i$ усреднение проводится только по q вершинам и доля вершин с малой частотой в них меньше. Количественно оценить величину $\eta = \frac{\sum_{i \in R} f_i}{f \cdot q}$ в общем виде не представляется возможным, поэтому было проведено моделирование. Типичный график зависимости величины η от номера решения N показан на рис. 4.

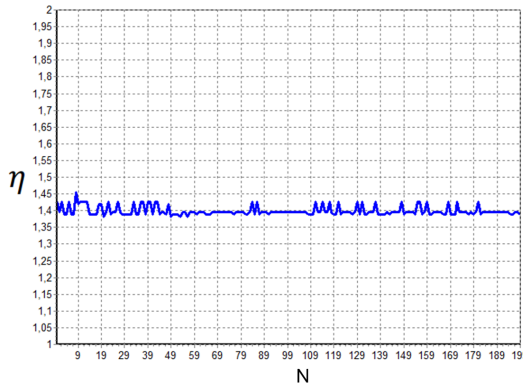


Рис. 4. График зависимости η от номера решения

Графики представляют собой незначительные колебания около горизонтальной линии. В зависимости от графа и количества удаляемых вершин величина η_0 , вокруг которой происходят колебания, меняется от 1,05 до 1,8, и для всех рассматриваемых графов и всех решений $\eta \geq 1,05$.

Окончательно величина μ была оценена как $\max(1 - b)$, где
$$b = \frac{1 + \left[\frac{1,05 \cdot q \cdot (f \cdot d - 1)}{q - 1} \right]}{d},$$
 f – средняя частота, d – количество решений, q – количество удаляемых вершин и максимум брался по всем рассматриваемым графам и всем $q \leq q_0(n)$. Было получено значение $\mu \approx 0,5$, т.е. в результате в задачу ЛП можно добавлять ограничения до тех пор, пока количество решений не сократится вдвое.

5. Постановка задачи

Суммируя результаты, полученные в предыдущих разделах, можно дать следующую постановку задачи.

Рассматривается задача нахождения критических узлов укрупненных магистральных электрических сетей по критерию наименьшего количества связных пар вершин, которая решается точными или приближенными методами с помощью сведения ее к эквивалентной задаче линейного программирования. При этом задача решается для всех $q \leq \lceil n/5 \rceil$, где q – количество удаляемых вершин в графе сети и n – количество вершин графа. Требуется разработать метод снижения размерности этой задачи ЛП. Если для заданного графа и q с помощью сокращенной задачи ЛП будет найдено не меньше половины решений, то такой результат считается приемлемым.

Так как каких-либо определенных свойств графов магистральных электрических сетей неизвестно (даже не имеющее отношения к задаче свойство безмасштабности в последние годы ставится под сомнение), то решение предполагает часто используемый подход, заключающийся в следующем. Разрабатывается эвристический метод и при необходимости уточняются (опреде-

ляются, аппроксимируются) параметры метода на некоторой выборке графов указанного класса сетей. Далее на другой выборке определяется точность метода. Идея предлагаемого метода заключается в подборе двух характеристик вершин, свойства которых описаны в разделе «Постановка цели исследования». Предлагаемый метод также является эвристическим и не дает гарантии того, что не существует графа электрической сети, для которого в сокращенной задаче ЛП будет найдено меньше половины решений.

6. Оценка количества вершин, которые почти всегда являются критическими

Исследование этого вопроса можно разделить на два этапа:

1. Определить для каждого из 20 рассматриваемых графов и для каждого $q \leq \lceil n/5 \rceil$, существует ли такие подмножества вершин, которые почти всегда являются критическими. Найти подмножество, обладающее следующими свойствами (будем обозначать это подмножество A и $a = |A|$): доля решений, содержащих A , больше 0,5, и доля решений, содержащих любое подмножество вершин мощности $a + 1$, меньше 0,5;

2. Оценить величину a для произвольного графа электрических сетей.

6.1. Определение подмножеств критических вершин

Для нахождения подмножеств A наибольшей мощности использовался метод полного перебора, который осуществлялся следующим образом. Для каждого графа и для каждого $1 < q \leq \lceil n/5 \rceil$ перебирались все возможные комбинации вершин, количество вершин в комбинации r изменялось в диапазоне $1 \leq r \leq q - 1$. Для каждой комбинации вычислялась частота ее вхождения в решения, и для каждого r определялась группа комбинаций, частота которых превышала $\mu = 0,5$. В результате для каждого графа и каждого q была получена группа подмножеств A .

6.2. Оценка величины $a = |A|$

Эта подзадача заключается в оценке a для произвольного графа магистральной электрической сети. по имеющимся группам подмножеств A . Так как в распоряжении имелись только группы подмножеств A , то использовались методы аппроксимации, что, конечно, не в лучшую сторону сказалось на точности оценки.

Очевидно, что a зависит от q , какой-либо другой зависимости a , например, от количества вершин и ребер графа, выявлено не было. Оказалось, что лучше оценивать зависимость от q величины $q - a$. На рис. 5 представлен график этой зависимости.

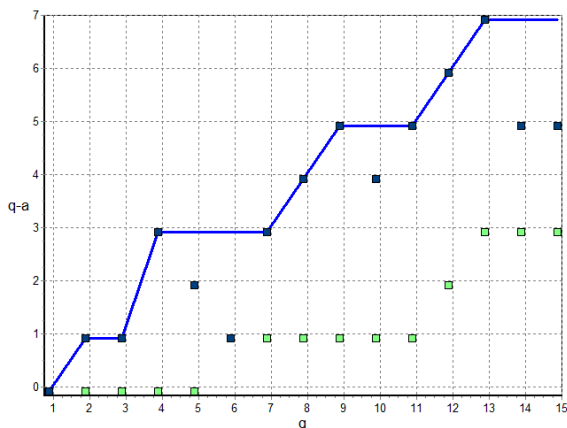


Рис. 5. График зависимости $q - a$ от q

Зелеными точками показаны усредненные по всем рассматриваемым графам значения величины $q - a$, а темно-синими — максимальные (среди всех рассматриваемых графов) значения этой величины. Из графика видно, что средние значения $q - a$ не убывают, следовательно, логично предположить, что и максимальные значения будут не убывать, а убывание в некоторых областях на графике вызвано недостаточным количеством рассматриваемых графов электрических сетей, которое, в свою очередь, вызвано длительным периодом решения для них задачи нахождения

дения критических узлов. Синим цветом показан график неубывающих максимальных значений $q - a$, для которого и строится аппроксимация.

Обозначим как $f^+(q)$ функцию, аппроксимирующую этот график, и чтобы определить ее вид, также рассмотрим график зависимости $\frac{q-a}{q}$ от q , изображенный на рис. 6; раскраска точек аналогична предыдущему рисунку.

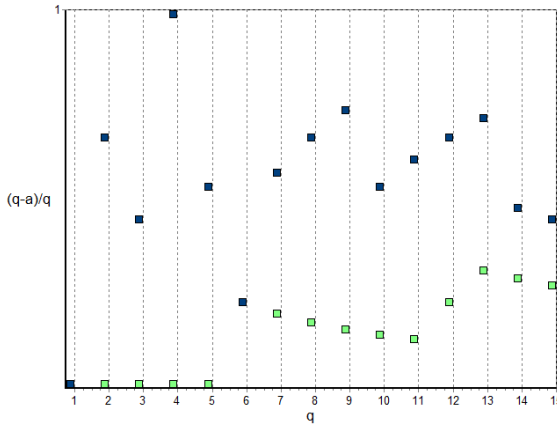


Рис. 6. График зависимости $\frac{q-a}{q}$ от q

Функция $f^+(q) \approx q - a$ не может иметь порядок строго выше q , так как в противном случае $a \approx q - f^+(q) \rightarrow -\infty$. Если бы $f^+(q)$ имела порядок строго ниже q , то на рис. 7 это проявилось бы в немонотонном убывании как средних, так и максимальных значений $\frac{q-a}{q}$, а этого не наблюдается. Поэтому логично аппроксимировать величину $q - a$ функцией порядка q . На рис. 7 коричневой линией показана аппроксимация величины $q - a$ функцией $f^+(q) = 2\lfloor \frac{q}{4} \rfloor + 1$.

В итоге будем оценивать величину a (наибольшее количество добавляемых в задачу ЛП ограничений вида $v_i = 1$) следующим образом: $a = q - 2\lfloor q/4 \rfloor - 1$.

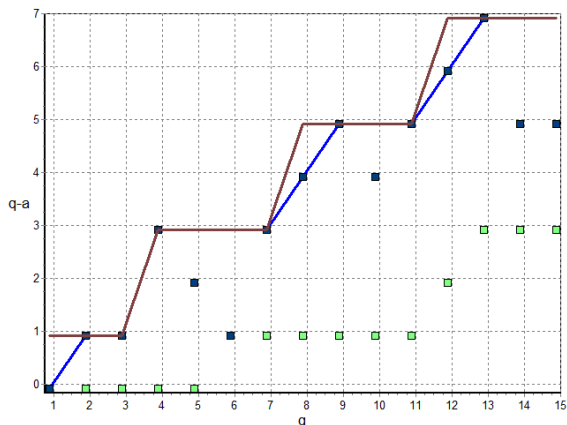


Рис. 7. График зависимости $q - a$ от q и его аппроксимация

7. Оценка количества вершин, которые почти никогда не являются критическими

Аналогично предыдущему разделу исследование этого вопроса можно разделить на два этапа:

1. Определить (для каждого q) подмножество вершин, которые никогда не являются критическими (будем обозначать это подмножество B и $b = |B|$). Найти подмножество, обладающее следующими свойствами: доля решений, не содержащих B , больше 0,5, и доля решений, не содержащее любое подмножество вершин мощности $b + 1$, меньше 0,5.

2. Оценить и по возможности спрогнозировать значение величины $b = |B|$;

7.1. Определение подмножеств B

Аналогично подмножествам A подмножества B также находились методом полного перебора, при этом для каждого графа и каждого q отдельно выделялись подмножества вершин, не входящих ни в одно решение. Это существенно сократило время перебора.

7.2. Оценка величины $b = |B|$

Очевидно, что b зависит и от q и от n . На рис. 8 представлен график зависимости $b(q, n)$.

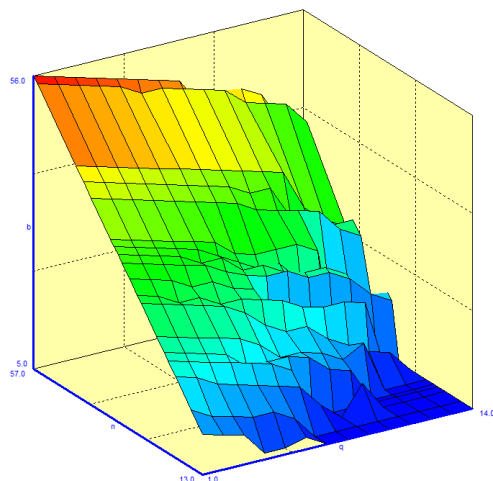


Рис. 8. График зависимости b от q и n

Этот график можно охарактеризовать как «плоскость с отклонениями». Причем отклонения увеличиваются при возрастании q и убывании n , т.е. в области $q \geq \lceil n/5 \rceil$, в которой количество критических узлов превышает полученную в разделе 4 оценку. Было установлено, что в области $q \leq \lceil n/5 \rceil$ выполняется соотношение $b \leq n - q$, причем в большинстве случаев $b = n - q$. Это объясняется тем, что при малых q количество решений мало и равенство будет в случае единственного решения задачи ЛП. В итоге будем оценивать величину b значением $n - q$.

8. Заключение

Рассматривается задача нахождения критических узлов инженерной сети – набора вершин, при удалении которых из графа сети, количество связных пар вершин в оставшемся графе будет

минимально. Один из наиболее часто используемых методов нахождения точного решения этой задачи или решения, близкого к точному, заключается в сведении ее к эквивалентной задаче линейного программирования (ЛП). Для этих методов проводится исследование возможности сокращения размерности задачи ЛП следующим путем. Необходимо подобрать две характеристики вершин таких, что первые несколько вершин, упорядоченных по значению первой характеристики, почти всегда являются критическими и последние несколько вершин, упорядоченных по значению второй характеристики, почти никогда не являются критическими. Если такие характеристики возможно подобрать, то в задаче ЛП можно заменить переменные, отвечающие за то, является ли данная вершина критической или нет, значениями 1 и 0 соответственно, тем самым сократив размерность задачи.

В работе представлена подготовительная часть исследования и получены следующие результаты, конкретизирующие задачу и сокращающие дальнейший перебор.

1. Показано, что в общем случае (для произвольного графа) вышеописанные характеристики вершин подобрать невозможно. Поэтому вместо всех возможных графов рассматривались графы магистральных электрических сетей, для которых показана теоретическая возможность существования таких характеристик.

2. Из общих практических соображений показано, что достаточно рассматривать наборы критических вершин, мощность которых не превышает пятую часть всех вершин.

3. При решении сокращенной задачи возможно будет найдено не все множество решений, т.е. не все наборы критических вершин. Также из общих соображений показано, что допустимо сокращать размерность до тех пор, пока количество найденных решений больше половины всех решений исходной задачи.

4. Сформулирована корректная постановка задачи.

5. Получены верхние оценки количества вершин, которые почти всегда являются критическими и почти никогда не являются критическими.

Литература

1. *Минпросвещения России.* – Режим доступа: <https://edu.gov.ru/modernization> (дата обращения: 22.05.2024).
2. *Перспективы применения комплексов альтернативной энергии на примере республики Таджикистан.* – Режим доступа: <https://research-journal.org/archive/10-64-2017-october/perspektivy-primeneniya-kompleksov-alternativnoj-energii-na-primere-respubliki-tadzhikistan> (дата обращения: 21.05.2024).
3. *Программа реновации. Итоги и планы на 2024 г.* – Режим доступа: <https://www.sobyanin.ru/programma-renovatsii-itogi-i-planu-na-2024>(дата обращения: 22.05.2024).
4. ХАЧИЯН Л.Г. *О точном решении систем линейных неравенств и задач линейного программирования* // Вычисл. матем. и матем. физ. – 1982. – №22(4). – С. 999–1002.
5. *Электроэнергия в Беларуси: распределение – статьи энергетической тематики.* – Режим доступа: https://energobelarus.by/blogs/Energy_dis-senting_opinion/19/ (дата обращения: 21.05.2024).
6. BECZI E., GASKO N. *Approaching the bi-objective critical node detection problem with a smart initialization-based evolutionary algorithm* // Peer Computer Science. – 2021. – Vol. 7. – P. 750–758.
7. BONNANS J.F., GILBERT J.C., LEMARECHAL C. et al. *Numerical optimization: Theoretical and practical aspects.* – Springer Berlin–Heidelberg, 2006. – Vol. 51. – P. 494–495.
8. CHEN W., JIANG M., JIANG C. et al. *Critical node detection problem for complex network in undirected weighted networks* // Physica A: Statistical Mechanics and its Applications. – 2020. – Vol. 538. – P. 11–45.

9. DINH T.N., XUAN Y., THAI M.T. et al. *On New Approaches of Assessing Network Vulnerability: Hardness and Approximation* // IEEE ACM Trans. on Networking. – 2012. – Vol. 20(2). – P. 609–619.
10. GLORY U.A., ARULSELVAN A., AKARTUNALI K. et al. *Efficient methods for the distance-based critical node detection problem in complex networks* // Computers and Operations Research. – 2021. – Vol. 131. – P. 108–121.
11. HOOSHMAND F., MIRARABRAZI F., MIRHASSANI S.A. *Efficient Benders decomposition for distance based critical node detection problem* // Omega. – 2020. – Vol. 93. – P. 16–31.
12. LALOU M., KHEDDOUCI H. *Network Vulnerability Assessment Using Critical Nodes Identification* // Int. Symposium on Networks, Computers and Communications (ISNCC). – 2023. – P. 1–6.
13. LIU C., GE G., ZHANG Y. *Identifying the cardinality-constrained critical nodes with a hybrid evolutionary algorithm* // Information Sciences. – 2023. – Vol. 642. – P. 24–41.
14. MEGZARI A., PRAVIJA RAJ P.V., OSAMY W. *Applications, challenges, and solutions to single- and multi-objective critical node detection problems: a survey* // J. Supercomput. – 2023. – Vol. 79. – P. 19770–19808.
15. MUNIKOTI S., DAS L., NATARAJAN B. *Scalable graph neural network-based framework for identifying critical nodes and links in complex networks* // Neurocomputing. – 2023. – Vol. 422. – P. 211–221.
16. SALEMI H., BUCHANAN A. *Solving the Distance-Based Critical Node Problem* // INFORMS Journal on Computing. – 2022. – Vol. 34(3). – P. 1309–1326.
17. SHEN S., SMITH J.C., GOLI R. *Exact interdiction models and algorithms for disconnecting networks via node deletions* // Discrete Optimization. – 2012. – Vol. 9. – P. 172–188.

18. SHEN Y., NGUYEN N.P., XUAN Y. et al. *On the Discovery of Critical Links and Nodes for Assessing Network Vulnerability* // IEEE/ACM Trans. on Networking. – 2013. – Vol. 21. – P. 963–973.
19. THAI M.T., DINH T.T., SHEN Y. *Hardness and Approximation of Network Vulnerability* // Handbook of Combinatorial Optimization. – 2013. – Vol. 5. – P. 1631–1666.
20. UGURLU O., AKRAM N., AKRAM V.K. *Critical nodes detection in IOT-based cyber-physical systems: Applications, methods, and challenges* // Emerging trends in IoT and integration with data science, cloud computing, and big data analytics. – 2022. – Vol. 2022. – P. 226–239.
21. VEREMYEV A., PROKOPYEV O.A., PASILIAO E.L. *Critical nodes for distance-based connectivity and related problems in graphs* // Networks. – 2015. – Vol. 66. – P. 170–195.
22. WALTEROS J.L., VEREMYEV A., PARDALOS P.M. et al. *Detecting critical node structures on graphs: A mathematical programming approach* // Networks. – 2018 – Vol. 73. – P. 48–88.
23. XU Y., GUO P. *MEA-CNDP: A Membrane Evolutionary Algorithm for Solving Biobjective Critical Node Detection Problem* // Computational Intelligence and Neuroscience. – 2021 – Vol. 2021. – P. 101–118.

REDUCTION THE DIMENSIONALITY OF THE TASK OF FINDING CRITICAL NODES IN THE NETWORK

Andrey Krygin, V.A. Trapeznikov Institute of Control Sciences of RAS, Moscow, Cand.Sc., (andreyakr@yandex.ru).

Sofia Tarasova, V.A. Trapeznikov Institute of Control Sciences of RAS, Moscow, technician (tarasva_sofia@mail.ru).

Abstract: One of the classes of problems solved in the assessment of the stability of an engineering network is the problem of finding critical nodes. In many formulations, this problem is posed as finding a subset of nodes of a given cardinality (critical nodes) such that the failure of which would cause maximum damage to the entire network. And the most common way to assess the damage in such a formulation is to determine the number of connected node pairs in the network with excluded critical nodes. For such nodes that correspond to the minimum number of connected pairs, additional measures are required to increase reliability and safety. Several methods of solving the problem of finding critical nodes use reducing it to an equivalent linear programming problem. The main problem of this approach is the large size of the problem, and consequently, the high computational complexity of its solution. The work conducts research on various characteristics of vertices of a graph model of a network, the analysis of the values of which will allow determining in advance the fact of belonging to the subset of critical or, conversely, to the subset of non-critical nodes. Thanks to this, it is possible to form additional constraints that reduce the dimensionality of the linear programming problem and its computational complexity, which will allow finding critical nodes in engineering networks with a large number of objects in an acceptable time. During the research, a large number of different subproblems were solved, so the work describes only the first, preparatory part of it.

Keywords: resilience of electrical networks, search for critical nodes, graph models of networks.

УДК 519.1

ББК 31.27-082

DOI: 10.25728/ubs.2024.111.5

*Статья представлена к публикации
членом редакционной коллегии А.А. Печниковым.*

Поступила в редакцию 08.04.2024.

Дата опубликования 30.09.2024.

КОГНИТИВНЫЕ КАРТЫ ДЛЯ АНАЛИЗА И МОДЕЛИРОВАНИЯ В СИСТЕМЕ СТРАТЕГИЧЕСКОГО ПЛАНИРОВАНИЯ ГОСУДАРСТВА

Авдеева З. К.¹,

*(ФГБУН Институт проблем управления
им. В.А. Трапезникова РАН, МГТУ им. Н.Э. Баумана, Москва)*

Коврига С. В.²

*(ФГБУН Институт проблем управления
им. В.А. Трапезникова РАН, Москва)*

Одной из основных задач научно-методологического обеспечения государственного стратегического планирования является развитие методологии индикативного планирования и мониторинга реализации документов стратегического планирования. Индикативное планирование является инструментом оценки достижения целей социально-экономического развития и обеспечения безопасности, в его основе лежит система индикаторов. Индикаторы дают возможность осуществлять измерение, мониторинг, оценку достижимости поставленных целей развития и, в случае отклонений, обеспечить коррекцию государственной политики в нужном направлении устойчивого социально-экономического развития. Выделены проблемы системной несогласованности документов стратегического планирования, отсутствия механизмов (1) проверки непротиворечивости целей, задач и мер, предусмотренных в этих документах, обеспеченности целей и задач наборами целевых показателей и индикаторов; (2) прослеживания структуры связей целей – задач – индикаторов. Исходя из обозначенных проблем предложена модель систематизации целей – задач – индикаторов с использованием аппарата когнитивных карт. Представлены аналитические основы прослеживания и реализации базовой задачи мониторинга параметров данной модели на базе анализа ее структурных свойств. Работоспособность предложенного научно-методического и информационно-аналитического инструментария демонстрируется с использованием ряда документов стратегического планирования Российской Федерации. Обозначены некоторые направления дальнейших исследований в части (1) расширения задачи мониторинга достижимости целей развития с включением модели внешней среды для отслеживания влияния на цели внешних сигналов в текущей обстановке; (2) развития методов композиции и анализа многоуровневой сети целей –

¹ Зинаида Константиновна Авдеева, к.т.н., с.н.с. (avdeeva@ipu.ru).

² Светлана Вадимовна Коврига, н.с. (kovriga@ipu.ru).

задач – индикаторов – мероприятий, формирующейся по срезу заданных стратегических целей в системе документов стратегического планирования.

Ключевые слова: стратегическое планирование, индикативное планирование, когнитивная карта, мониторинг, структурный анализ.

1. Введение

В современных экономических условиях все большее распространение приобретает индикативное планирование (ИП) применительно к стратегическому планированию (СП) на разных уровнях государственного управления (страна, регион, отрасль) или корпоративного управления (различные виды хозяйствующих субъектов).

ИП как форма косвенного государственного регулирования играет важную роль для различных хозяйствующих субъектов в качестве планов-прогнозов, позволяющих последним лучше ориентироваться в долгосрочной экономической конъюнктуре при разработке собственных планов [4, 5].

ИП является важным инструментом по реализации целей, поставленных в стратегическом плане развития с учетом конкретно складывающейся экономической ситуации [5, 7]. Основу ИП составляет система целевых показателей и индикаторов, на основе которых фиксируются параметры социально-экономического развития объекта (страны, региона, отдельных хозяйствующих субъектов) в средне- или долгосрочной перспективе для обеспечения возможности оценки достижимости поставленных целей его развития. ИП, являясь альтернативой директивному планированию, предполагает замену жестко заданных плановых показателей развития объекта СП на систему оценки параметров с учетом возможных изменений и непрямого регулирования возможных исполнителей. С точки зрения теории организационного управления ИП является примером информационного управления (воздействия), так как индикаторы являются ориентирами для хозяйствующих субъектов и их плановые значения оцениваются исходя из модельных ситуаций в некотором приближении.

В практике государственного управления в качестве целевых часто используются плановые и прогнозные значения показателей, установленные в программных документах СП [3].

Совокупность документов СП порождает сложную сетевую модель управления разными сферами и аспектами государственной политики¹, что порождает ряд проблем, связанных с системной несогласованностью и несбалансированностью документов СП по приоритетам, целям, задачам, показателям, срокам реализации, ресурсам; неразвитостью научно-методологического и информационно-аналитического обеспечения ИП (см., например, [6, 8]).

Поэтому в целях развития ИП в практике государственного СП (что закреплено на государственном уровне²) актуальными являются задачи

- формирования модели ИП, направленной на
 - построение системы сопряженных (по целям) документов СП;
 - систематизацию, структуризацию и сквозную взаимосвязку параметров: целей, задач, индикаторов, мероприятий в этой системе документов;
- разработку инструментария согласования, прослеживания, мониторинга и корректировки параметров системы документов СП.

В контуре стратегического управления в условиях сложной сетевой системы взаимосвязей разных сфер функционирования и развития такой системы, как государство, необходимо применять гетерогенный комплекс моделей, на верхнем уровне – модели, идентифицирующие наборы параметров, связанных с целями и задачами, и структуру их взаимосвязей, на других уровнях, по мере появления конкретных данных о процессах,

¹ В ГАИС «Управление» (<http://gasu.gov.ru/stratdocuments>), предназначенной для информационно-аналитического обеспечения деятельности органов госвласти по важнейшим направлениям социально-экономического развития РФ, развития ключевых отраслей экономики, реализации нацпроектов и программ, зарегистрировано тысячи документов СП разных уровней управления.

² Указ Президента РФ от 08.11.2021 N 633 «Об утверждении Основ государственной политики в сфере стратегического планирования в Российской Федерации».

связанных с этими параметрами, – количественные модели типа моделей системной динамики и другие многопараметрические модели описания ситуаций. На таком гетерогенном комплексе моделей решаются задачи целеполагания, прогнозирования динамики в различных прогнозных условиях, планирования ресурсов, мониторинга изменений и исполнений, выработки управленческих решений по коррекции целей, задач и мероприятий [9, 10].

Исходя из обозначенных проблем и актуальных задач ИП в сфере государственного СП в статье предложена модель систематизации целей – задач – индикаторов M_{G-T-IN} с использованием аппарата когнитивных карт. Когнитивная карта является разновидностью сетевой модели представления знаний, формализующей предметную область в виде множества концептов, определяющих сущности (объекты) предметной области, и отношений причинно-следственных влияний между ними. Представлены аналитические основы прослеживания, мониторинга и структурной оценки связей параметров M_{G-T-IN} на базе анализа ее структурных свойств. Задача мониторинга рассмотрена в базовой постановке с целью оценки результативности программ развития с точки зрения влияния на достижимость системы целей верхнего уровня.

2. Модель систематизации целей – задач – индикаторов в задачах индикативного планирования

2.1. ОПИСАНИЕ МОДЕЛИ

Модель систематизации целей – задач – индикаторов M_{G-T-IN} , отраженных в документах СП, определяет два вида отношений между ее параметрами:

- причинно-следственные влияния между целями $G = \{G_i\}$ и задачами $T = \{T_j\}$;
- обеспеченность целей и задач (или их групп) наборами целевых показателей и индикаторов $IN = \{In_k\}$.

Модель M_{G-T-IN} обеспечивает прослеживание структуры связей целей – задач – индикаторов различной сложности –

от типовой иерархической (линейно упорядоченной) модели до сетевой модели с обратными связями (циклами).

Причинно-следственные влияния определяются моделью связанности M_{G-T} целей $\{G_i\}$ и задач $\{T_j\}$.

Модель M_{G-T} по некоторому направлению (приоритету) D представляет собой обобщенную когнитивную карту $K^D = K^G \cup K_T^{G_1} \cup K_T^{G_2} \cup \dots \cup K_T^{G_{J_D}}$, где K^G – когнитивная карта связанности целей $G=\{G_i\}$, $K_T^{G_1}$, $K_T^{G_2}$, ..., $K_T^{G_{J_D}}$ – когнитивные карты связанности отдельных целей $G_1, G_2, \dots, G_{J_D}$ с соответствующими им задачами.

На рис. 1 приведен схематичный пример общей ациклической когнитивной карты K^D (двунаправленная пунктирная стрелка означает, что допускаются отношения причинно-следственного влияния между задачами отдельных карт).

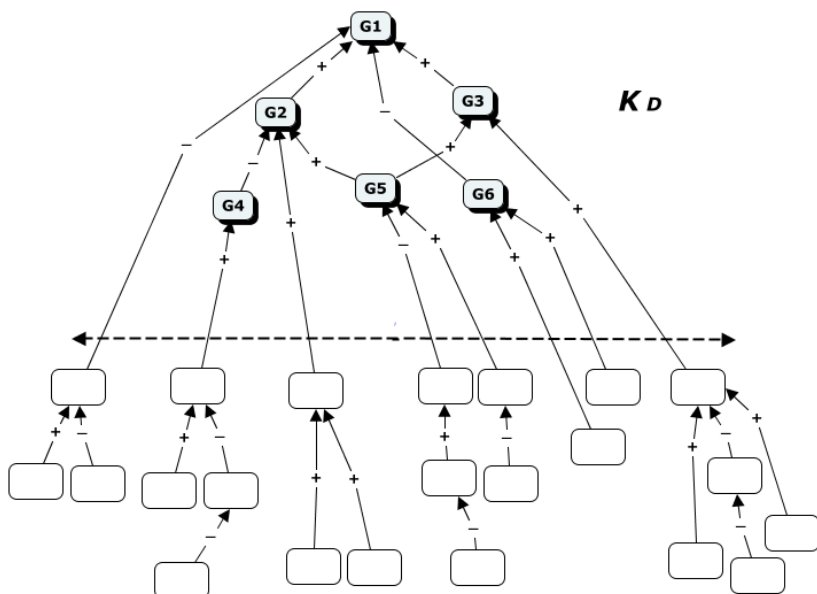


Рис. 1. Общая когнитивная карта K^D связанности целей и задач по направлению D

На рис. 2 приведены примеры нескольких отдельных когнитивных карт, выделенных из K^D : K^G , $K_T^{G_1}$, $K_T^{G_3}$.

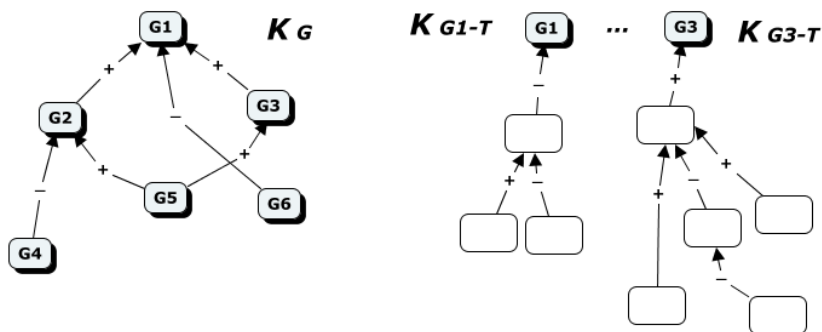


Рис. 2. Отдельные когнитивные карты: K^G – карта связанности целей, $K_T^{G_1}$ – карта связанности цели G_1 с ее задачами, $K_T^{G_3}$ – карта связанности цели G_3 с ее задачами

Формально каждая когнитивная карта $K = (P, E^{+/-})$ в модели M_{G-T} представляет собой ориентированный знаковый граф, где вершинам приписаны параметры $\{p_i\}$ – цели G или задачи T ; дугам, интерпретируемым как прямые причинно-следственные влияния $E^{+/-}$ между $\{p_i\}$, приписаны знаки («+» или «-»), характеризующие направления влияния между любой парой p_i, p_j ($p_i \xrightarrow{+/-} p_j$): «+» – усиливающее или «-» – тормозящее, ослабляющее. Ограничение: в $K = (P, E^{+/-})$ недопустимы связи $p_i \xrightarrow{+/-} p_j$, если в паре p_i – цель и p_j – задача.

Множество $E^{+/-}$ может быть представлено квадратной матрицей $[e_{ij}^{+/-}]$, где $e_{ij}^+ = 1$ – знак «+» влияния p_i на p_j , интерпретируемый как однонаправленное изменение $p_i \xrightarrow{+} p_j$: параметр-следствие p_j изменяется в том же направлении, что и параметр-причина p_i при прочих равных условиях; $e_{ij}^- = -1$ – знак «-» влияния параметра p_i на p_j , интерпретируемый как разнонаправленное изменение $p_i \xrightarrow{-} p_j$: параметр-следствие p_j изменяется

в противоположном направлении от изменения параметра-причины p_i при прочих равных условиях; если влияние между парой параметров (p_i, p_j) отсутствует, то $e_{ij} = 0$.

При построении когнитивных карт рекомендуется исходные формулировки целей (задач), взятые из документов СП, максимально сохранять для облегчения сопоставимости формулировок в исходных документах СП и модельном представлении.

На рис. 3а,б приведены примеры представления причинно-следственных влияний «+» и «-» между парой целей.

При «+» направленность изменения причины и следствия совпадает, при «-» направленность изменения причины и следствия противоположная. Интерпретацию (прочтение связи) проиллюстрируем примером на рис. 3б: сокращение ↓ (увеличение ↑) бедности приводит к повышению ↑ (снижению ↓) качества жизни населения при прочих равных условиях.

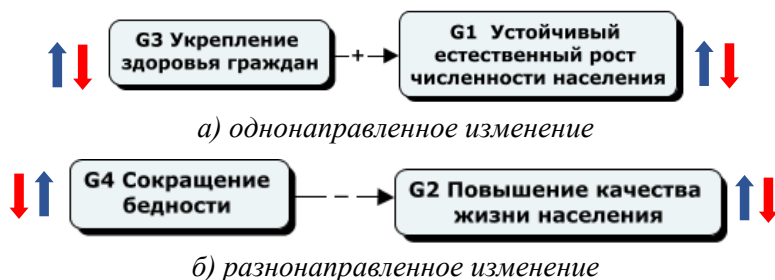


Рис. 3. Примеры представления причинно-следственного влияния между парой целей

В модели M_{G-T-IN} второй вид отношения – обеспеченность индикаторами целей и задач представляется привязкой к отдельным картам $K^G, K_T^{G_1}, K_T^{G_2}, \dots, K_T^{G_{J_D}}$ соответствующих наборов индикаторов (схематично представлено на рис. 4).

Таким образом, модель систематизации целей – задач – индикаторов включает $M_{G-T-IN} = (M_{G-T}, IN)$, где модель M_{G-T} есть объединение когнитивных карт связанности по отдельным

целям с соответствующими им задачами $M_{G-T} = K^D = K^G \cup K_T^{G_1} \cup K_T^{G_2} \cup \dots \cup K_T^{G_{JD}}$, IN – множество индикаторов, разделенных по их принадлежности к отдельным картам $IN = \{\{In_{K_T^{G_1}}\}, \{In_{K_T^{G_2}}\} \dots \{In_{K_T^{G_{JD}}}\}\}$, в том числе с соотносением индикаторов к соответствующим задачам в рамках каждой карты $K^G, K_T^{G_1}, K_T^{G_2}, \dots, K_T^{G_{JD}}$.

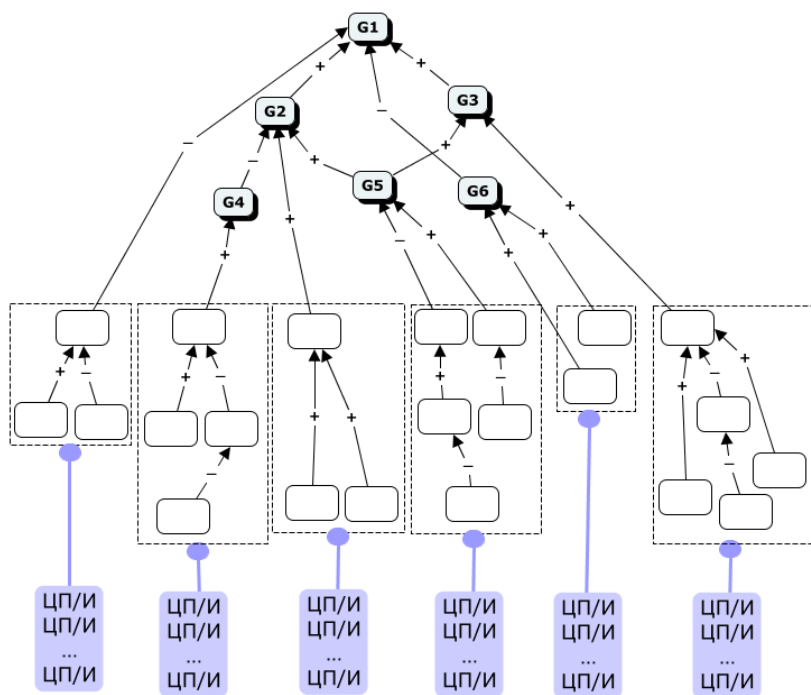


Рис. 4. Общая когнитивная карта K^D с соотносением наборов целевых показателей и\или индикаторов (ЦП/И) к соответствующим $\{K_T^{G_j}\}$, объединенных в K^D

Применение модели систематизации целей – задач – индикаторов M_{G-T-IN} в задачах ИП направлено на

- формирование общей структуры увязки параметров по связанным документам СП различных уровней;
- прослеживание преемственности целей и задач в структуре связанных документов СП различных уровней и их обеспеченности индикаторами с их целевыми (и/или пороговыми) значениями;
- формирование базы индикаторов, связанных с целями и задачами устойчивого социально-экономического развития объекта СП;
- актуализацию базы индикаторов в ходе мониторинга достижимости целей социально-экономического развития объекта СП и контроля актуальности документов СП;
- мониторинг и структурную оценку связей в модели M_{G-T-IN} с целью
 - выявления дублирования, противоречий, несогласованности параметров как в отдельных картах $K^G, K_T^{G_1}, K_T^{G_2}, ..., K_T^{G_{j_D}}$, так и в связанных картах K^D , упущенных влияний для оценки результативности программ с точки зрения влияния на систему целей верхнего уровня;
 - реакции параметров на (1) управленческие сигналы, выраженные в новых или обновляемых документах СП (например, на государственном уровне к таким документам относятся ежегодные послания Президента РФ, Указы Президента или Правительства РФ); (2) внешние сигналы ситуации вокруг объекта СП в текущей обстановке.

Аналогичную структуру прослеживания можно сформировать и по мероприятиям, направленным на достижение целей и задач документа СП верхнего уровня.

2.2. АНАЛИТИЧЕСКИЕ ОСНОВЫ ПРОСЛЕЖИВАНИЯ, МОНИТОРИНГА И СТРУКТУРНОЙ ОЦЕНКИ СВЯЗЕЙ ПАРАМЕТРОВ МОДЕЛИ M_{G-T-IN}

Конечной целью прослеживания, мониторинга и структурной оценки связей параметров M_{G-T-IN} является проверка обеспеченности целей и задач индикаторами IN с их целевыми

(и/или пороговыми) значениями для оценки достижимости целей G .

Задача прослеживания целей и задач в связанных документах СП различных уровней ($k = 1, 2, \dots$) решается с использованием критерия преемственности $Kr_{sun}(p^k, p^{k+1})$.

$Kr_{sun}(p^k, p^{k+1})$ является слабо формализованным: он представляется в виде словесного шаблона, который в логическом смысле есть предикат со свободными параметрами (p^k, p^{k+1}) . При подстановке вместо (p^k, p^{k+1}) конкретных параметров $(g_i^k, (g_i^{k+1} \vee t_j^{k+1}))$ или $(t_i^k, (g_i^{k+1} \vee t_j^{k+1}))$ формируется высказывание $Kr_{sun}((g_i^k, (g_i^{k+1} \vee t_j^{k+1}))) \vee Kr_{sun}((t_i^k, (g_i^{k+1} \vee t_j^{k+1})))$ соответственно, истинность которого означает формальное соответствие критерию $Kr_{sun}(p^k, p^{k+1})$.

Содержательно, критерий $Kr_{sun}(p^k, p^{k+1})$ выполняется, если цель g_i^{k+1} или задача t_j^{k+1} из документа уровня $k + 1$ является преемником цели g_i^k или задачи t_i^k документа уровня k . Преемственность означает, что специалист¹ подтверждает, что g_i^k или t_j^k соответствует по смыслу g_i^{k+1} или t_j^{k+1} даже если формулировки этих параметров имеют различия (различия могут быть разной степени, но важно, чтобы содержательные смыслы формулировок сопоставляемых параметров были близки (в идеале совпадали)).

Преемственность по параметрам может быть очевидная (явная), когда формулировки сопоставляемых параметров из разных документов СП практически совпадают, т.е. различия не существенные, или условная (неочевидная) преемственность. Условная преемственность означает, что формулировки сопоставляемых параметров имеют существенные различия, но исходя из опыта и знания предметной области специалист в различных формулировках сопоставляемых параметров «улавливает» сходство содержательных смыслов, в противном случае фиксируется отсутствие преемственности между сопо-

¹ Здесь и далее подразумевается специалист по СП.

ставляемыми параметрами. В сложных случаях оценки условной преемственности рекомендуется групповое согласование между специалистами.

Следует отметить, что оценка преемственности параметров в связанных документах СП различных уровней может быть трудоемкой в силу многоуровневости документов СП (что особенно характерно для документов на государственном уровне) и включением в них множества целей и задач по различным направлениям или приоритетам. Поэтому встает вопрос инструментальной поддержки этого процесса на основе применения средств интеллектуального анализа текстов; простейший вариант такой поддержки – организация направленного поиска по текстам связанных документов СП и по модели систематизации целей – задач – индикаторов M_{G-T-IN} на основе формируемых специалистом запросов (по ключевым словам). Также необходима инструментальная поддержка по визуализации, построению и коррекции M_{G-T-IN} .

Важным аспектом применимости модели M_{G-T-IN} является решение задачи полноты и достаточности декомпозиции «цели – задачи – индикаторы». Выполнение этих критериев влияет на эффективность оценки достижимости целей верхнего уровня документов СП.

По формальным признакам в предложенной модели отслеживания структурным критерием полноты является обеспеченность набором индикаторов (с их целевыми или пороговыми значениями) каждой задачи, связанной с некоторой целью. Нарушение критерия полноты приводит к невыполнению критерия достаточности из-за нарушения принятой структуры декомпозиции в модели. Примерами нарушения критерия полноты является отсутствие в модели M_{G-T-IN} задач при объявленной цели или отсутствие индикаторов для объявленных задач. При выполнении критерия полноты формальная оценка по критерию достаточности затруднительна, так как для такой оценки требуется привлечение специалистов по СП вместе с экспертами предметной области для содержательного (смыслового) анализа выявленных индикаторов. С точки зрения формализации оценки достаточности можно лишь говорить о выяв-

лении в структуре модели рискованных конструкций, по которым необходимо проверка по критерию достаточности. Простейшим случаем такой конструкции является наличие единственного индикатора, связанного с решением отдельной задачи. Выявление таких типовых конструкций является предметом отдельного исследования.

Задачи мониторинга и структурной оценки связей в модели M_{G-T-IN} ставятся с целью

- оценки результативности программ развития с точки зрения влияния на достижимость системы целей верхнего уровня (базовая задача мониторинга);
- отслеживания и оценки влияния внешних сигналов (угроз) на достижимость целей развития в текущей обстановке на основе композиции модели M_{G-T-IN} и модели внешней среды M_{Ext} (расширение базовой задачи мониторинга).

Рассмотрим базовую задачу. Пусть даны:

- модель M_{G-T-IN} ;
- вектор оценок желательной динамики параметров (ОДП) когнитивной карты M_{G-T} , $R(P)$, включающий оценки динамики целей $R(G)$ и задач $R(T)$:

$$(1) \quad r_i(p_i) = \begin{cases} +1, & \text{если благоприятен рост значения } p_i, \\ -1, & \text{если благоприятно уменьшение значения } p_i, \\ 0, & \text{если значение } p_i \text{ не оценивается} \end{cases},$$

где p_i есть $G_i \in G$ или $T_i \in T$.

Требуется оценить

- непротиворечивость между целями и согласованность между целями и связанными с ними задачами (в структуре модели M_{G-T});
- наличие отклонений от благоприятного направления изменения целей в зависимости от неблагоприятной оценки решения задач, связанных с целями в структуре M_{G-T} .

Определение непротиворечивости между целями и согласованности между целями и связанными с ними задачами состоит в сопоставлении ОДП (целей, задач) и направлений (знаков) влияний через все возможные пути (цепи), связывающие их в M_{G-T} .

Определение 1. Любая пара целей G_i и G_j непротиворечива, если для любого пути l_{ij}^k , связывающего G_i и G_j , справедливо

$$(2) \quad r_j(G_j) = \text{sign}(\text{sign}(l_{ij}^k) \times r_i(G_i)),$$

где $r_i(G_i)$ и $r_j(G_j)$ – ОДП целей G_i и G_j , l_{ij}^k – k -й путь от цели G_i к цели G_j .

Знак любого пути l_{ij}^k определяется как произведение знаков дуг, входящих в него. (Аналогично выстраивается оценка непротиворечивости между задачами).

Определение 2. Любая пара цель G_j и задача T_i согласованы, если для любого пути l_{ij}^k , связывающего T_i и G_j , справедливо

$$(3) \quad r_j(G_j) = \text{sign}(\text{sign}(l_{ij}^k) \times r_i(T_i)),$$

где $r_i(T_i)$ и $r_j(G_j)$ – ОДП T_i и G_j ; l_{ij}^k – k -й путь от задачи T_i к цели G_j .

При оценке непротиворечивости \ согласованности параметров M_{G-T} важно выделять особые случаи потенциальных противоречий, связанных с наличием путей положительного и отрицательного влияния между парой параметров (рис. 5).

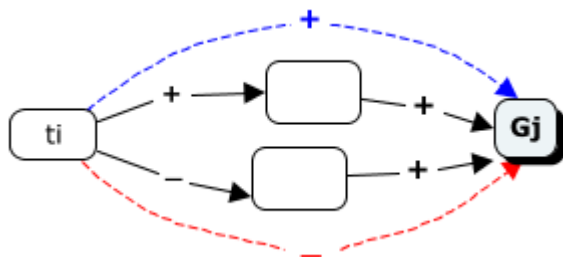


Рис. 5. Схематическое представление особого случая потенциального противоречия

Такие случаи не позволяют получить однозначную оценку непротиворечивости (согласованности) между этими параметрами.

Определение 3. Потенциальное противоречие между парой параметров p_i , p_j существует, если между этими параметрами

найдена хотя бы одна пара разнонаправленных путей $l_{p_i p_j}^+$ и $l_{p_i p_j}^-$, где $sign(l_{p_i p_j}^+) = 1$ и $sign(l_{p_i p_j}^-) = -1$.

Если в когнитивной карте M_{G-T} выполняются (1) условия непротиворечивости (согласно определениям 1 и 2) для любых пар p_i и p_j , для которых существуют пути причинно-следственных влияний в структуре карты, и (2) отсутствуют потенциальные противоречия (согласно определению 3), то формируется обобщенная структурная оценка связей – M_{G-T} согласована; в противном случае – не согласована по некоторому набору целей (задач) и/или наличию потенциальных противоречий (что требует от специалистов принятия решений по приведению к согласованности целей и задач, которое может состоять в коррекции M_{G-T} или документа СП).

Если при мониторинге исполнения документа СП формируется неблагоприятная оценка решения некоторых задач $\{T_i^{neg}\}$ ($\{-\text{ОДП}(T_i^{neg})\}$), то через структуру M_{G-T} можно определить все цели (задачи), на которые они оказывают влияние, и тем самым выделить критичные зоны достижения целей (задач), по которым требуется принятие решений специалистами.

Условиями возникновения рассогласований и противоречий между целями, задачами являются ситуации, когда принимаются решения по коррекции действующих документов СП, которые могут быть связаны с изменением состава целей и задач. Тогда требуется пересмотр структуры всей модели M_{G-T-IN} вплоть до прослеживания до уровня индикаторов.

Таким образом результаты решения базовой задачи мониторинга формируют и систематизируют необходимые данные для оценки результативности программ развития с точки зрения влияния на достижимость системы целей верхнего уровня.

В расширенной постановке решение задачи мониторинга предполагает

- структурный анализ $M_{Ext} \cup M_{G-T-IN}$ для выявления структуры проблемной ситуации, отражающей прямые и косвенные связи негативных параметров (угроз) модели M_{Ext} с целевыми параметрами M_{G-T-IN} ;

– формирование сценариев качественных изменений ситуации (в контексте достижимости целей развития) с учетом значимости и актуализации угроз в текущей обстановке для своевременного принятия решений по актуализации документа СП в изменяющейся обстановке, контроля его исполнения.

В качестве задела для постановки и решения указанной задачи предполагается использовать наработки авторов в области решения задач стратегического мониторинга слабоструктурированных проблемных ситуаций с использованием когнитивных карт [1, 2].

Актуализация документов СП должна проводиться по всем уровням СП, вплоть до документов (госпрограмм), определяющих индикаторы и мероприятия по достижению целей развития.

3. Демонстрационный пример прослеживания и мониторинга

Демонстрация предложенного подхода к формированию модели основных целей, задач и базовых индикаторов обеспечения национальной безопасности и социально-экономического развития на государственном уровне проводится на основе следующего набора документов СП РФ:

Стратегия национальной безопасности РФ, утверждена Указом Президента РФ от 2 июля 2021 г. № 400, (СНБ РФ).

Национальные цели развития РФ на период до 2030 года, утверждены Указом Президента РФ от 21.07.2020 №474, (НЦР).

Единый план по достижению национальных целей развития РФ на период до 2024 года и на плановый период до 2030 года, утвержден распоряжением Правительства РФ от 01.10.2021 N 2765-р, (ЕдПлан).

Государственная программа РФ «Развитие здравоохранения», утверждена Постановлением Правительства РФ от 26.12.2017 N 1640, (ГП Здрав).

Верхний уровень систематизации целей – задач – индикаторов формируется на основе СНБ РФ, которая является базовым документом СП, определяющим национальные интересы и стратегические национальные приоритеты (СНП) РФ, цели

и задачи государственной политики в области обеспечения национальной безопасности и устойчивого социально-экономического развития РФ на долгосрочную перспективу. При этом СНБ РФ не определяет систему целевых показателей (индикаторов), по которой можно было бы оценить достижимость целей СНБ РФ.

Цели и задачи, предусмотренные в рамках СНП в СНБ РФ, решаются путем разработки, корректировки и исполнения конкретизирующих документов СП. Поэтому прослеживание преемственности целей и задач СНБ РФ в этих документах направлено на выявление в них целевых показателей и индикаторов, связанных с прослеживаемыми целями и задачами СНБ РФ, для последующего мониторинга и контроля достижимости целей.

Демонстрация задачи мониторинга в базовой постановке ограничивается оценкой согласованности выборочного подмножества целей и задач СНБ РФ, по которым проводилось прослеживание и формирование модели M_{G-T-IN} (раздел 3.2).

3.1. ПРОСЛЕЖИВАНИЕ

Этап 1. Структуризация параметров СНБ РФ, направленная на увязку целей (задач) между собой и увязку задач с целями.

Структура связей между параметрами СНБ РФ представляется в виде когнитивной карты (см. раздел 2.1).

Когнитивные карты, отражающие причинно-следственные влияния параметров СНП1 СНБ РФ – «Сбережение народа России и развитие человеческого капитала»: целей $G1 - G6$ и выборочного подмножества задач, связанных с реализацией целей $G1$ и $G3$, $K^{G1,G3}$, представлены на рис. 6 и 7 соответственно.

На рис. 7 серым цветом выделены цели $G2, G4 - G6$, которые не включены в демонстрационный фрагмент карты (стрелки с многоточием указывают на их связи в полной карте, не представленные в данном фрагменте).

Сформированная структура связей в $K^{G1,G3}$ (см. рис. 7) позволяет прослеживать цепочки влияний параметров $K^{G1,G3}$ внутри СНБ РФ, выделять фрагменты связанных между собой параметров для последующего прослеживания их в других документах.

СНП1 Сбережение народа России и развитие человеческого потенциала
G1-G6: Цели

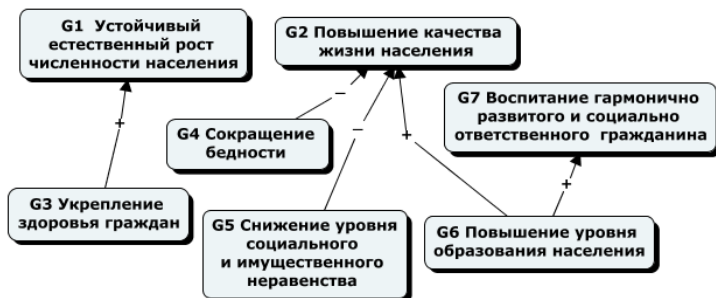


Рис. 6. Когнитивная карта целей G1–G6 СНП1

СНП1: Сбережение народа России и развитие человеческого потенциала

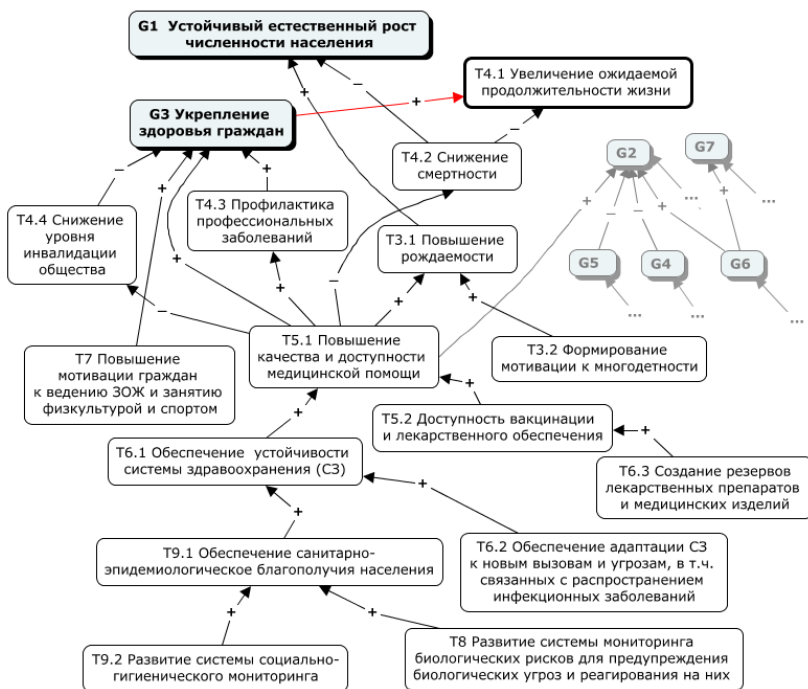


Рис. 7. Когнитивная карта $K^{G1,G3}$ увязки задач с целями G1, G3

Особенностью карты $K^{G1, G3}$ является то, что выделена сомнительная связь (красная стрелка на рис. 7) влияния цели «G3 – Укрепление здоровья граждан» на задачу «Т4.1 – Увеличение ожидаемой продолжительности жизни», несмотря на то, что в M_{G-T-IN} связи типа $G \xrightarrow{+/-} T$ недопустимы (см. раздел 2.1).

Объяснение причинно-следственного влияния G3 на T4.1 состоит в том, что укрепление здоровья граждан приводит к увеличению продолжительности жизни (при прочих равных условиях), однако параметр T4.1, по сути, сформулирован как индикатор (расчетный показатель), отражающий состояние здоровья населения. Включение T4.1 в перечень задач достижения целей СНП1 в принципе некорректно, целесообразно рассматривать его как индикатор достижения цели G3, а формулировку задачи корректировать.

Этап 2. Оценка связанности целей и задач СНБ РФ с Национальными целями развития и их целевыми показателями.

Следующим по значимости документом СП является НЦР (национальные цели развития РФ), который определяет главные цели развития и связанные с ними целевые показатели, характеризующие их достижение.

Для демонстрации выбрана НЦР1: «Сохранение населения, здоровье и благополучие людей», для которой в документе определены четыре целевых показателя (см. рис. 8).

Содержательный (смысловой) анализ формулировок СНП1 и НЦР1 показывает, что СНП1 «Сбережение народа России и развитие человеческого капитала» в СНБ РФ согласуется с НЦР1 «Сохранение население, здоровье и благополучие людей».

На рис. 9 представлена увязка целевых показателей T11–T14 НЦР1 с целями и задачами СНП1 из СНБ РФ, которая состоит в прослеживании параметров СНБ РФ в документе НЦР по критерию преемственности Kr_{sun} (раздел 2.2).

В частности, критерий $Kr_{sun}(G1, T11)$ выполняется, так как T11: «Обеспечение устойчивого роста численности населения РФ» является преемником G1: «Устойчивый естественный рост численности населения».

НЦР1 Сохранение населения, здоровье и благополучие людей
Т11–Т14: Целевые показатели

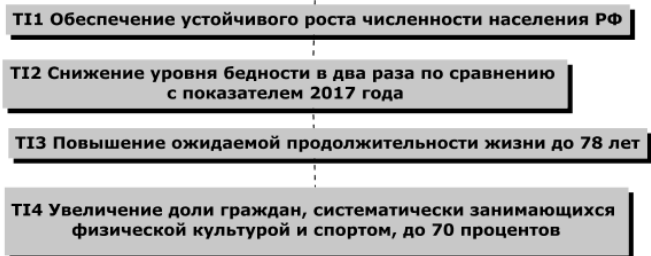


Рис. 8. Целевые показатели Т11–Т14 достижения НЦР1
 «Сохранение население, здоровье и благополучие людей»

СНП1 Сбережение народа России и развитие человеческого потенциала



Рис. 9. Увязка целевых показателей Т11–Т14 НЦР1
 с целями и задачами СНП1, представленными на рис. 7

Преимственность означает, что формулировки G1 и T11 согласуются по смыслу: устойчивый естественный рост населения (формируемый как разница между рождаемостью и смертностью) является составной частью устойчивого роста, формируемого в том числе и с учетом миграционных процессов. Нетруд-

но заметить, что $Kr_{sun}(G4, TI2)$, $Kr_{sun}(T4.1, TI3)$ и $Kr_{sun}(T7, TI4)$ также выполняются (см. рис. 9).

Примечание. Далее по тексту упоминание Kr_{sun} в описании этапов прослеживания по документам не приводится в целях снижения перегрузки описания демонстрации, хотя в основе любого прослеживания лежит выполнение этого критерия.

Этап 3. Дальнейшее прослеживание целей и задач СНБ РФ в документе ЕдПлан.

ЕдПлан является дальнейшей конкретизацией НЦР, так как определяет систему индикаторов и комплекс мероприятий по достижению НЦР и связанных с ними целевых показателей.

Выборочное прослеживание демонстрируется по $G1-TI1$ и $T4.1-TI3$ (на рис. 9 этот фрагмент выделен прямоугольником).

На рис. 10 приведена структуризация параметров документа ЕдПлан, связанных с $G1-TI1$ и $T4.1-TI3$. (Важно отметить, что в ЕдПлан $TI1$ и $TI3$ рассматриваются в качестве целей.)

В округлых четырехугольниках определены задачи, влияющие на достижение $TI1$ и $TI3$. По составу эти задачи согласуются с задачами, направленными на достижение $G1$ и $T4.1$ СНБ РФ в когнитивной карте $K^{G1, G3}$ (см. рис. 7), за исключением задач «поддержка старшего поколения» и «обеспечение качественного миграционного прироста» (выделены жирной штриховкой). Указанные задачи не объявлены в СНБ РФ в перечне задач СНП1. Исходя из этого следует вывод, что состав задач, связанных с СНП1 СНБ РФ в части увязки с целью $G1$ и задачей $T4.1$ (рис. 8) в документе ЕдПлан расширен. Включение этих задач свидетельствует о том, государственная политика, направленная на сбережение народа РФ, расширяет направления деятельности в этой сфере (относительно объявленных задач в СНБ РФ).

На рис. 10 перечисляются индикаторы (в пунктирных прямоугольниках), характеризующие результативность решения задач, влияющих на достижение $TI1$ и $TI3$. С учетом увязки $TI1$ и $TI3$ с $G1$ и $T4.1$ (см. рис. 9) эти индикаторы также характеризуют результативность решения соответствующих задач, влияющих на достижение $G1$ и $T4.1$ в СНБ РФ (см. рис. 7).

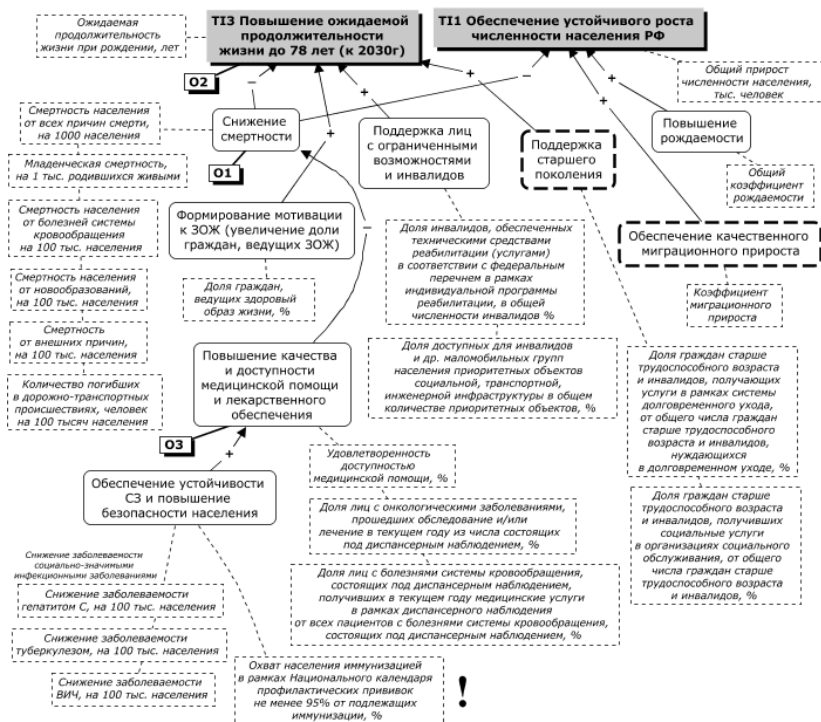


Рис. 10. Структуризация параметров документа ЕдПлан, связанных с ТП1 и ТП3

Документ ЕдПлан определяет конкретизирующие документы, связанные с задачами ЕдПлана. В частности, документ ГП Здрав определяет основные направления реализации, связанные со снижением смертности и повышением качества медицинских услуг.

Этап 4. Дальнейшее прослеживание целей и задач СНБ РФ в документе ГП Здрав.

На рис. 11 представлены связи параметров ГП Здрав, где основные цели программы 01–03 выделены в прямоугольниках с тенью, с задачами СНП1 СНБ РФ (см. рис. 7).

При этом на рис. 10 соответствующие метки 01–03 указывают на их согласованность с параметрами ЕдПлан, являющимся промежуточным между СНБ РФ и ГП Здрав.

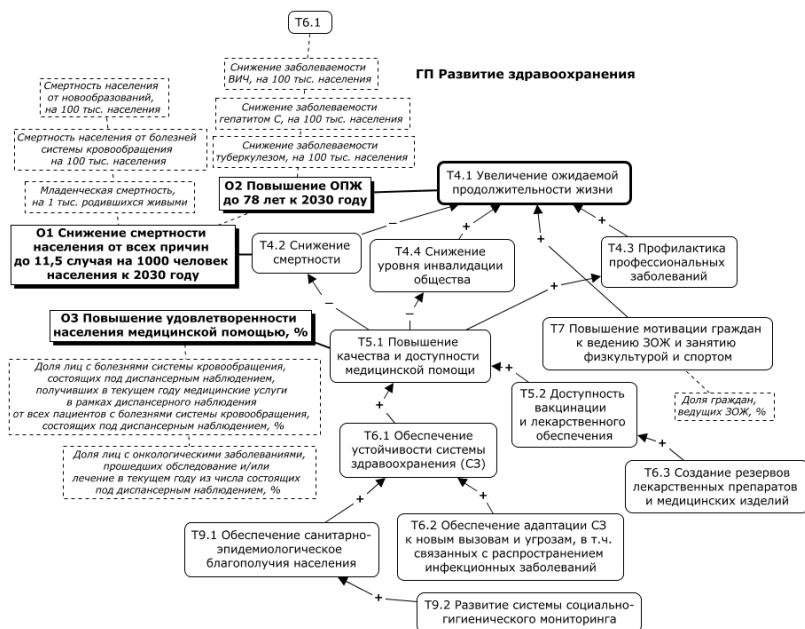


Рис. 11. Структуризация параметров из ГП Здрав

Таким образом, задачи ГП Здрав с учетом связей между ними (см. рис. 11) согласованы с задачами СНП1 и их связями (см. рис. 7) с учетом прослеживания последних в цепочке документов СНБ РФ → НЦР → ЕдПлан → ГП Здрав. (Задачи ГП Здрав направлены на достижение целей G1 и G3 в СНП1 СНБ РФ.) В пунктирных прямоугольниках перечислены индикаторы ГП Здрав, которые согласуются с индикаторами документа ЕдПлан, за исключением индикатора «Охват населения иммунизацией в рамках Национального календаря профилактических прививок не менее 95% от подлежащих иммунизации, %» (на рис. 10 выделен восклицательным знаком), который в ГП Здрав не представлен (упущен).

Выводы по прослеживанию. В рамках демонстрации проведено выборочное прослеживание преемственности целей и задач СНП1 СНБ РФ (на выделенном подмножестве) в документах СП: НЦР, ЕдПлан и ГП Здрав с целью формирования системы базовых целевых показателей и индикаторов, необходимых

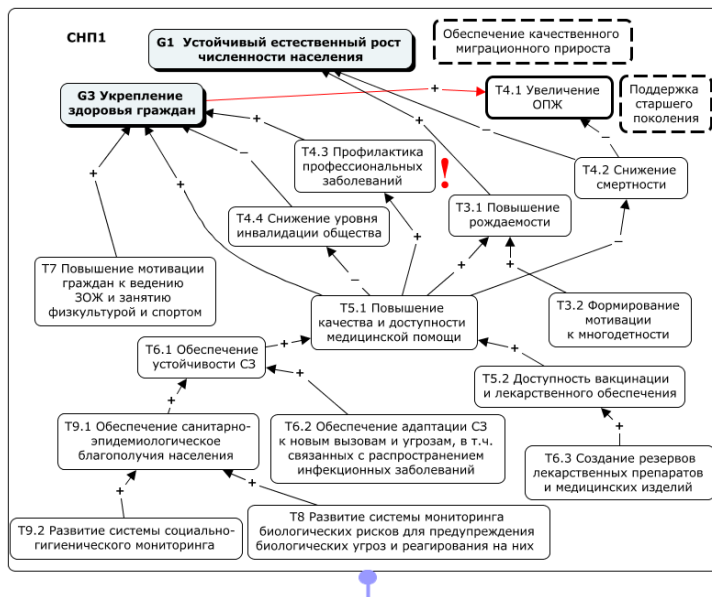
для оценки достижимости целей СНБ РФ. На рис. 12, 13 приведены сводные результаты прослеживания.

На рис. 12 сверху приведена когнитивная карта $K^{G1, G3}$ выделенных целей и задач СНП1 СНБ РФ. Снизу приведен перечень индикаторов, связанных с картой $K^{G1, G3}$, который был сформирован в ходе прослеживания. Каждому индикатору присвоен идентификатор его принадлежности цели или задаче СНП1 СНБ РФ в карте $K^{G1, G3}$. После формулировки индикатора в скобках указывается метка документа, в котором объявлен индикатор: ЕП – ЕдПлан, ГП – ГП Здрав.

В округлых четырехугольниках определены задачи, влияющие на достижение $T11$ и $T13$. По составу эти задачи согласуются с задачами, направленными на достижение $G1$ и $T4.1$ СНБ РФ в когнитивной карте $K^{G1, G3}$ (см. рис. 7), за исключением задач «поддержка старшего поколения» и «обеспечение качественного миграционного прироста» (выделены жирной штриховкой). Указанные задачи не объявлены в СНБ РФ в перечне задач СНП1. Исходя из этого следует вывод, что состав задач, связанных с СНП1 СНБ РФ в части увязки с целью $G1$ и задачей $T4.1$ (рис. 8), в документе ЕдПлан расширен. Включение этих задач свидетельствует о том, государственная политика, направленная на сбережение народа РФ, расширяет направления деятельности в этой сфере (относительно объявленных задач в СНБ РФ).

В целях снижения перегрузки описания демонстрации декомпозиция индикаторов до числовых значений не приводится; по всем выделенным индикаторам (рис. 12, 13) в соответствующих документах СП указаны их целевые значения по годам: 2017–2024 гг. и 2030 г.

На рис. 12 и 13 приведены две задачи (в пунктирных четырехугольниках) и соответствующие им индикаторы (рис. 13), выделенные при прослеживании в ЕдПлан. В составе задач СНП1 СНБ РФ они не объявлены.



- T4.1** • Ожидаемая продолжительность жизни при рождении, лет (ЕП, ГП)
- G1** • Общий прирост численности населения, тыс. чел. (ЕП)
- T3.1** • Общий коэффициент рождаемости (ЕП)
- T4.2** • Смертность населения от всех причин смерти, на 1000 населения (ЕП, ГП)
- T4.2** • Младенческая смертность, на 1 тыс. родившихся живыми (ЕП, ГП)
- T4.2** • Смертность населения от болезней системы кровообращения на 100 тыс. нас. (ЕП, ГП)
- T4.2** • Смертность населения от новообразований, на 100 тыс. населения (ЕП, ГП)
- T4.2** • Смертность от внешних причин, на 100 тыс. населения (ЕП)
- T4.2** • Кол-во погибших в дорожно-транс. происшествиях, чел. на 100 т. нас. (ЕП)
- T7** • Доля граждан, ведущих здоровый образ жизни, % (ЕП, ГП)
- T4.1** • Доля инвалидов, обеспеченных техническими средствами реабилитации (услугами) в соответствии с федеральным перечнем в рамках индивидуальной программы реабилитации, в общей численности инвалидов % (ЕП)
- T4.1** • Доля доступных для инвалидов и др. маломобильных групп населения приоритетных объектов социальной, транспортной, инженерной инфраструктуры в общем количестве приоритетных объектов, % (ЕП)
- T5.1** • Удовлетворенность доступностью медпомощи, % (ЕП, ГП)
- T5.1** • Доля лиц с онкозаболеваниями, прошедших обследование и/или лечение в текущем году из числа состоящих под диспансерным наблюдением, % (ЕП, ГП)
- T5.1** • Доля лиц с болезнями системы кровообращения, состоящих под диспансерным наблюдением, получивших в текущем году медицинские услуги в рамках диспансерного наблюдения от всех пациентов с болезнями системы кровообращения, состоящих под диспансерным наблюдением, % (ЕП, ГП)
- T6.1** • Снижение заболеваемости гепатитом С, на 100 тыс. населения (ЕП, ГП)
- T6.1** • Снижение заболеваемости туберкулезом, на 100 тыс. населения (ЕП, ГП)
- T6.1** • Снижение заболеваемости ВИЧ, на 100 тыс. населения (ЕП, ГП)
- T6.1** • Охват населения иммунизацией в рамках Нац. календаря профилактических прививок не менее 95% от подлежащих иммунизации, % (ЕП)

Рис. 12. Сводные результаты прослеживания

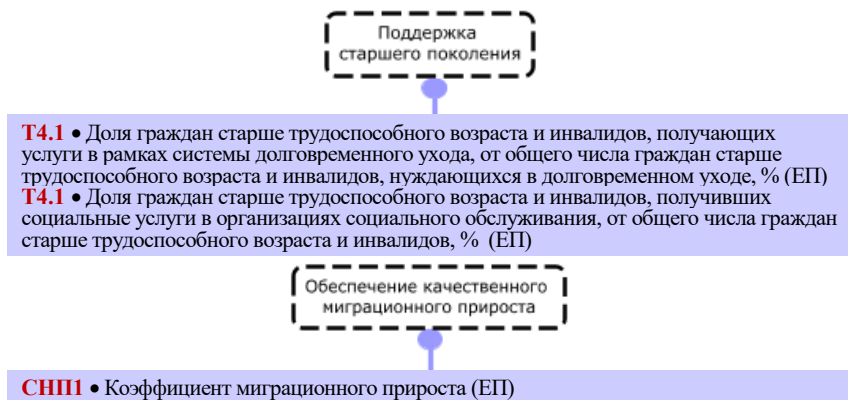


Рис. 13. Сводные результаты прослеживания (продолжение)

Общая оценка по критериям полноты и достаточности декомпозиции «цели – задачи – индикаторы». Построенная демонстрационная модель (см. рис. 12, 13) формально не удовлетворяет критерию полноты, потому что в карте $K^{G1, G3}$, отражающей структуру связей целей $G1$, $G3$ с соответствующими им задачами, ряд задач не обеспечен индикаторами (например, $T3.2$, **T4.3**, $T5.2$, $T9.1$ и др.). Соответственно, оценка достижимости целей ограничена из-за недостаточности индикаторов, по фактическому состоянию которых может проводиться оценка выполнения поставленных задач для достижения целей. По задаче $T4.3$ «Профилактика профессиональных заболеваний» (на рис. 12 помечена ! в $K^{G1, G3}$) не прослеживаются индикаторы (и мероприятия) в выделенных документах, при том, что формулировка этой задачи упоминается в ГП Здрав. По остальным задачам требуется прослеживание в других документах СП. Вызывает сомнение достаточности единственного индикатора «Коэффициент миграционного прироста» для контроля реализации задачи «Обеспечение качественного миграционного прироста» (см. рис. 13) за счет мероприятий по обеспечению приезда востребованных высококвалифицированных специалистов в страну из-за рубежа.

Выявленные некорректности, упущения и несоответствия.

1. *Некорректные связи.* Выделена сомнительная связь (красная стрелка в $K^{G1, G3}$ на рис. 12) влияния цели «G3 – Укреп-

ление здоровья граждан» на задачу «Т4.1 – Увеличение ожидаемой продолжительности жизни», несмотря на то что в модели M_{G-T-IN} связи типа $G \xrightarrow{+/-} T$ недопустимы. Параметр Т4.1 сформулирован как индикатор, отражающий состояние здоровья населения. Включение Т4.1 в перечень задач достижения целей СНП1 СНБ РФ в принципе некорректно, целесообразно рассматривать его как индикатор достижения цели G3, а формулировку задачи корректировать.

2. *Упущенные параметры при прослеживании.* Индикатор In: «Охват населения иммунизацией в рамках Национального календаря профилактических прививок не менее 95% от подлежащих иммунизации», % (выделен жирным шрифтом на рис. 12), связанный с задачей Т6.1 СНБ РФ, введен в документе ЕдПлан, но далее упущен в ГП Здрав и как следствие, не прослеживаются мероприятия, связанные с этим индикатором, в ГП Здрав.

3. *Расширение состава задач для достижения целей верхнего уровня.* На рис. 12, 13 приведены две задачи (в пунктирных четырехугольниках) и соответствующие им индикаторы, выделенные при прослеживании в ЕдПлан.

Эти задачи включены в реализацию НЦР, что свидетельствует о том, что государственная политика, направленная на сбережение народа РФ, расширяет направления деятельности в этой сфере (относительно объявленных в СНБ РФ). Иными словами, эти задачи не конкретизируют цели и задачи СНБ РФ, а дополняют (расширяют) их. Поэтому, учитывая связанность документов СНБ РФ и НЦР и их обоюдную значимость в системе СП, необходимо провести актуализацию СНБ РФ, чтобы не упустить важные показатели состояния национальной безопасности при проведении государственного мониторинга ее состояния.

3.2. МОНИТОРИНГ

Как видно на рис. 12, в когнитивной карте M_{G-T} цели G1 и G3 не связаны, поэтому анализ на их противоречивость не имеет смысла.

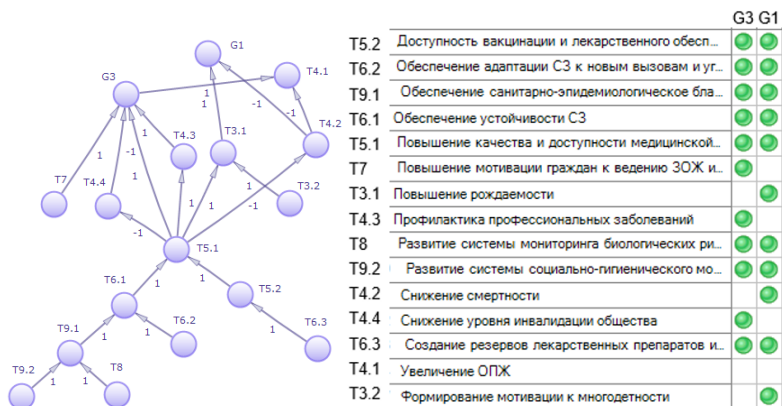


Рис. 14. Согласованность целей и задач в когнитивной карте M_{G-T}

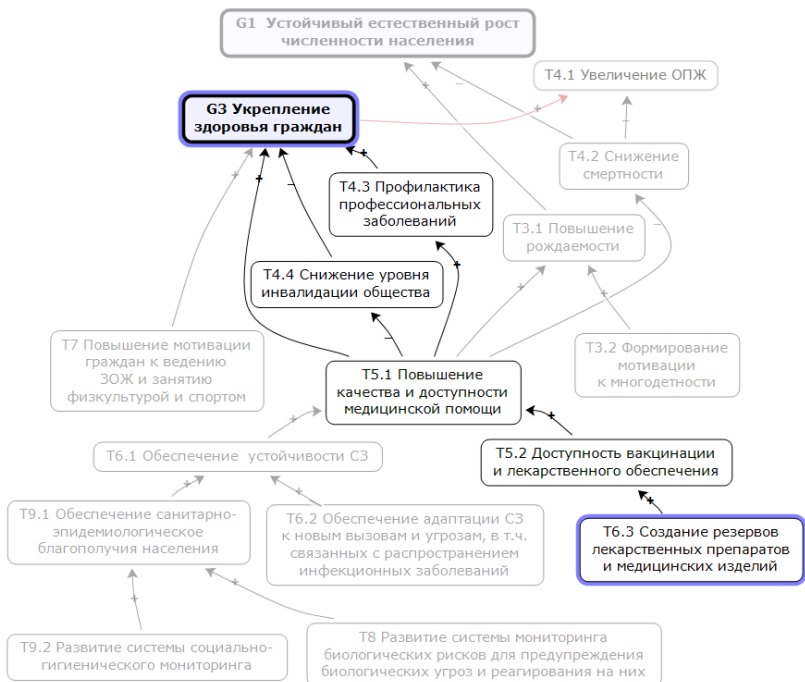


Рис. 15. Демонстрация отсутствия потенциального противоречия между целью G3 и задачей T6.3

Согласно определению 2, в структуре M_{G-T} цели $G1$ и $G3$ согласованы с задачами, т.е. с учетом благоприятной оценки решения задач обеспечивается желательное направление изменения целей (на рис. 14 зеленые шарики обозначают согласованность).

Согласно определению 3, потенциальные противоречия отсутствуют. Например, от задаче $T6.3$ и цель $G3$ связывают три пути (рис. 15) распространения положительного влияния.

4. Заключение

В рамках развития научно-методического и информационно-аналитического обеспечения ИП, относящегося к одному из основных компонентов СП на государственном уровне, предложен подход к систематизации параметров социально-экономического развития и обеспечения безопасности в сложной структуре документов СП на разных уровнях госуправления.

В основе предлагаемого подхода лежит построение модели увязки целей – задач – индикаторов M_{G-T-IN} с использованием аппарата когнитивных карт. Модель ориентирована на направленную, систематическую и комплексную оценку достижимости целей развития и обеспечения безопасности. Комплексность обеспечивается благодаря формированию структуры прослеживания указанных параметров от документов, определяющих стратегические цели развития объекта СП, до госпрограмм, задающих комплекс мероприятий и индикаторов, на основе которых фиксируется и отслеживается достижимость этих целей. Направленность и систематичность обеспечивается путем организации мониторинга и структурной оценки связей в модели M_{G-T-IN} . Аналитические основы прослеживания, мониторинга и структурной оценки связей параметров M_{G-T-IN} опираются на анализ ее структурных свойств. Также предложенный подход облегчает координацию между различными структурами и ведомствами госвласти, ответственных за реализацию документов СП.

Описание модели и демонстрация представлены только в контексте систематизации целей – задач – индикаторов,

без учета комплекса мероприятий. Аналогичную структуру прослеживания можно сформировать и до уровня мероприятий, что обеспечивает возможность оценивания выполнения поставленных задач, связанных с целями верхнего уровня, с учетом выполнения плановых мероприятий и достижимости целевых значений индикаторов.

Представленная демонстрация с использованием ряда официальных документов СП РФ свидетельствует о работоспособности предложенного подхода к формированию системы целей – задач – индикаторов для связанных документов СП. В небольшом фрагменте прослеживания выявлены некорректности и несоответствия в рассмотренных государственных документах СП. В ходе прослеживания относительно выделенных целей определен набор индикаторов по связанным документам СП, но не покрывающий весь перечень задач, связанных с этими целями. Поэтому оценка достижимости целей ограничена из-за недостаточности индикаторов, по фактическому состоянию которых может проводиться оценка выполнения поставленных задач для достижения целей.

Демонстрационный пример показывает, что прослеживание преемственности параметров в связанных документах СП различных уровней является трудоемкой задачей в силу многоуровневости документов СП (что особенно характерно для документов на государственном уровне) и включением в них множества целей и задач по различным направлениям или приоритетам. Поэтому стоит вопрос инструментальной поддержки этого процесса на основе применения средств интеллектуального анализа текстов.

Задача мониторинга рассмотрена в базовой постановке с целью формирования и систематизации необходимых данных для оценки результативности программ развития с точки зрения влияния на достижимость системы целей верхнего уровня. Дальнейшее развитие предполагает модификацию модели M_{G-T-IN} за счет наращивания ее моделью внешней среды и развитие методов мониторинга с учетом актуализации и влияния угроз внешней среды на стратегические цели развития объекта СП посредством структурного анализа и сценарного моделирования.

Еще одно направление развития предложенного подхода связано с разработкой методов визуализации и анализа, порождаемых моделью сложных структур прослеживания в многоуровневых государственных документах СП. Эти методы должны быть направлены на облегчение прочтения связей между параметрами в связанных документах на различных уровнях детализации, обеспечение свертки или развертывания структур параметров отдельных (или совокупности) документов для их сравнения. За счет применения стандартных метрик сетевых структур и развития новых в информационно-аналитических системах СП станет возможным решить следующие задачи:

- систематизировать критерии полноты, согласованности, преемственности, информационной обеспеченности, обеспеченности измеримыми индикаторами и сформировать схемы анализа сети по этим критериям;
- разработать принципы технологической и алгоритмической поддержки композиции системы цели – задачи – мероприятия – индикаторы в разных масштабах и уровнях по заданному целевому приоритету;
- разработать принципы визуальной аналитики получаемых сетей;
- разработать принципы и схемы формирования информационных массивов и лингвистического обеспечения функций поиска несоответствий.

Литература

1. АВДЕЕВА З.К., КОВРИГА С.В. *Диагностирование проблемных ситуаций в развитии сложных систем на основе когнитивных карт* // Управление большими системами. – 2013. – Вып. 42. – С. 5–28.
2. АВДЕЕВА З.К., КОВРИГА С.В. *Подход к постановке задач управления на когнитивной модели ситуации для стратегического мониторинга* // Управление большими системами. – 2016. – Вып. 59. – С. 120–146.
3. БЕЛЯЕВСКАЯ-ПЛОТНИК Л.А., БОЧАРОВА Л.К., СОРОКИНА Н.Ю. *Развитие подходов к мониторингу рисков не-*

- достижения национальных целей развития Российской Федерации и ее регионов // Региональная экономика. Юг России. – 2021. – Т. 9, №2. – С. 71–82.
4. *Индикативное планирование. Большая российская энциклопедия.* – URL: <https://bigenc.ru/c/indikativnoe-planirovanie-19f927> (дата обращения: 20.05.2024).
 5. ЛАДЫКОВА Т.И. *Теоретические аспекты индикативного планирования в современных условиях* // Oeconomia et Jus. – 2017. – №2. – С. 12–27.
 6. ЛЕНЧУК Е.Б. *Стратегическое планирование в России: проблемы и пути решения* // Инновации. – 2020. – №2(256). – С. 24–28.
 7. ПОПОВА Е.В. *Индикативное планирование как основной метод формирования долгосрочной социально-экономической стратегии России.* Инновации. – 2008. – №9(119). – С. 15–27.
 8. СМЕРНОВА О.О. *Формирование отечественной модели индикативного планирования* // МИР (Модернизация. Инновации. Развитие). – 2020. – Т. 11, №3. – С. 266–279.
 9. ПАЩЕНКО А.Ф., ИВАНЮК В.А., АБДИКЕЕВ Н.М. *Индикативное планирование в управлении процессами изменений* // Управленческие науки. – 2015. – №2. – С. 63–71.
 10. ЦВИРКУН А.Д., РЕЗЧИКОВ А.Ф., КУШНИКОВ В.А. и др. *Модели и методы проверки достижимости целей и выполнимости планов в крупномасштабных системах на примере целей и планов ликвидации последствий наводнения* // Автоматика и телемеханика. – 2023. – №12. – С. 49–63.

FUZZY-COGNITIVE-MAPS ANALYSIS AND MODELLING IN THE STATE STRATEGIC PLANNING SYSTEM

Zinaida Avdeeva, V.A. Trapeznikov Institute of Control Sciences of RAS, Moscow, Bauman State Technical University, Cand.Sc., Senior researcher (avdeeva@ipu.ru).

Svetlana Kovriga, V.A. Trapeznikov Institute of Control Sciences of RAS, Moscow, Researcher (kovriga@ipu.ru).

Abstract: One of the main tasks of scientific and methodological support of the state strategic planning is the development of methodology of indicative planning and monitoring of the implementation of strategic planning documents. Indicative planning is a tool for assessing the achievement of the goals of socio-economic development and security, it is based on a system of indicators. Indicators make it possible to measure, monitor, assess the achievability of the set development goals and, in case of deviations, to ensure the correction of public policy in the right direction of sustainable socio-economic development. The problems of system inconsistency of strategic planning documents, the lack of mechanisms (1) to check the consistency of goals, objectives and measures provided for in these documents, the provision of goals and objectives with sets of target indicators and indicators, (2) to trace the structure of links between parameters (goals – objectives – indicators) have been identified. Based on the identified problems, a model of systematisation of goals – objectives – indicators using the model of fuzzy cognitive maps is proposed. Analytical bases of tracking and implementation of the basic task of monitoring the parameters of this model on the basis of analysing its structural properties are presented. The workability of the proposed scientific-methodological and information-analytical toolkit is demonstrated using a number of strategic planning documents of the Russian Federation. Some directions for further research are outlined in terms of (1) expanding the task of monitoring the achievability of development goals with the inclusion of a model of the external environment to track the impact of external signals on the goals in the current situation, (2) the development of methods of composition and analysis of a multilevel network of goals-tasks-activities, formed by a cross-section of given strategic goals on the system of strategic planning documents.

Keywords: strategic planning, indicative planning, fuzzy cognitive map, monitoring, structure analysis.

УДК 338.26 + 303.732.4 + 004.822

ББК 65.054

DOI: 10.25728/ubs.2024.111.6

*Статья представлена к публикации
членом редакционной коллегии А.В. Горбуновой.*

Поступила в редакцию 03.07.2024.

Опубликована 30.09.2024.

ФОРМИРОВАНИЕ ГРУППОВОЙ И ИНДИВИДУАЛЬНОЙ ТРАЕКТОРИЙ УСПЕВАЕМОСТИ ПО ДАННЫМ E-LEARNING-ПЛАТФОРМ

Владова А. Ю.¹

*(ФГБУН Институт проблем управления
им. В.А. Трапезникова РАН, Москва, Финансовый
университет при Правительстве РФ, Москва)*

Поддержание высокого уровня образования является одной из основных задач управления вузом. Несмотря на постоянный мониторинг успеваемости учащихся, менеджмент образовательных учреждений недостаточно использует методы прогноза при формировании траекторий успеваемости. Предлагаемый подход отличается от известных тем, что анализируют признаки, содержащие оценки за различные работы, выполненные на e-learning-платформе, изменяют размерность пространства признаков за счет нормализации оценок по единой шкале, а также создания новых динамических признаков и признаков, содержащих решение задачи классификации учащихся, находящихся в группе академического риска и задачи прогнозирования экзаменационных оценок. Далее по дополненным данным, выделяют группы учащихся со схожими траекториями успеваемости для индивидуализации консультаций. Таким образом, формирование групповой и индивидуальной траекторий успеваемости с учетом результатов прогноза направлено на упреждающее повышение уровня академической успеваемости учащихся в университетских условиях.

Ключевые слова: академическая успеваемость, статистический анализ, линейный дискриминантный анализ, регрессия, траектории обучения, дистанционная платформа.

1. Введение

Управление образовательным учреждением учитывает успеваемость учащихся как одну из важнейших характеристик, по которой можно судить о достигнутых результатах или существующих проблемах. Успеваемость, как правило, мониторят с помощью множества аудиторных, самостоятельных, контрольных и курсовых работ, ответах на семинарах. Часть этих активностей

¹ Алла Юрьевна Владова, д.т.н., в.н.с. (avladova@ipu.ru).

часто автоматизирована с помощью e-learning-платформ [7]. Результаты мониторинга успеваемости используют для управления учебным процессом, при принятии решений о присуждении государственных академических и именных стипендий, выдаче дипломов с отличием и других задачах.

2. Анализ литературы

Оценка успеваемости является одним из важнейших компонентов электронного обучения. Авторы исследования [12] оценили успеваемость в области медицинского образования через показатели получения и восприятия знаний учащимися, уровень уверенности, удобство пользования e-learning-платформой и готовность рекомендовать электронное обучение. Изъян предлагаемого подхода заключается в качественном характере анализируемых признаков, их сильной зависимости от мнения разных экспертов. Исследователи [10] дополнили набор количественными признаками, которые позволяет собирать e-learning-платформа, а именно: набор онлайн-действий студента, время, необходимое, чтобы ответить на вопрос или отправить задание, пропущенные студентами вопросы. Кроме того, исследователи отдельно мониторили деструктивное поведение, которое включает низкую вовлеченность, чрезмерные опоздания, высокий процент отсева, списывание на заданиях и тестах, низкую эффективность обучения, уничижительные комментарии в онлайн-дискуссиях или по электронной почте. Недостатком данного подхода можно считать то, что данные брались без дополнительных преобразований, влияющих на производительность модели.

Исследование [14] предлагает модель прогнозирования эффективности онлайн-обучения, основанную на слиянии и выборе признаков. Модель использует взаимосвязь между признаками и объединяет учащихся в соответствии с категорией, которая достигла лучшей эффективности. Авторы [3, 4] в результате проведенной разведки данных для выявления скрытых взаимосвязей создали наборы синтетических признаков, позволяющих повысить эффективность моделей прогнозирования параметров организационных систем.

Авторы статьи [6] выделяют линейный дискриминантный анализ при вычислении вероятностного профиля принадлежности к целевым кластерам, на основе этого строится итоговое заключение.

Авторы [11] отмечают, что один из методов определения эффективности обучения предполагает использование линейной регрессии. Для этого они строят регрессионную модель по трем признакам (тест оценки знаний, активность в ходе курса и итоговый экзамен) и анализируют ее параметры. Перед построением регрессионной модели авторы изучили корреляционную матрицу, чтобы определить, насколько тесно связаны признаки. Минусы подхода заключаются в небольшой размерности пространства признаков и отсутствии учета динамики успеваемости.

Авторы [8] используют алгоритм кластеризации k -means, который выявляет учащихся с низкой вовлеченностью. Количество кластеров задано априори и недостаточно отражает учащихся с удовлетворительным уровнем взаимодействия. Исследователи отмечают [13], что кластеризация пользователей в группы с общими интересами очень полезна при изучении нескольких языков. Они и авторы [2] также используют алгоритм k -means из-за его простоты. В то же время известен метод DBScan [9], чья способность приспосабливаться к различным распределениям данных и идентифицировать кластеры различных форм и размеров делает его универсальным и мощным алгоритмом для выявления сложных закономерностей и структур в наборах данных.

Для выявления деструктивного поведения клиентов, которые недовольны предоставляемыми услугами, авторы статьи [1] рассмотрели различные методы классификации, включая деревья принятия решений, наивный байесовский классификатор, случайный лес, искусственные нейронные сети, линейный дискриминантный анализ, метод опорных векторов и их ансамбли. Эффективность классификаторов оценена с помощью таких показателей, как площадь под кривой AUC, точность, F1-мера и время классификации. Авторы указывают, что для четырех исследованных датасетов (после применения метода главных компонент для уменьшения пространства признаков) получены фактически идентичные оценки F1, точности и AUC.

Таким образом, установлено, что при построении моделей успеваемости авторы недостаточно учитывают динамику изменения показателей успеваемости, минимально используют преобразования имеющихся данных при повышении точности моделей, а также отмечено отставание в применении современных классификационных моделей при формировании групповой и индивидуальной траекторий успеваемости. Основное противоречие заключается в том, что при формировании индивидуальных траекторий успеваемости учащихся не учитывают результаты прогноза и групповые траектории.

В результате анализа литературы сформулированы новизна подхода и практическая применимость.

Новизна формирования групповой и индивидуальной траекторий успеваемости учащихся отличается от известных тем, что на основе значительного количества независимых от эксперта разновременных оценок генерируют динамические признаки, далее формируют дополнительные признаки решением задачи классификации с выделением группы учащихся под академическим риском и задачи регрессии с предсказанием индивидуальной траектории успеваемости, а в заключение решают задачу кластеризации по поиску учащихся со схожими траекториями успеваемости с неизвестным априори количеством кластеров.

Практическая применимость подхода сформулирована следующим образом:

- высокая предсказательная способность на основе исторических данных обучения;
- прогнозы и факторы, влияющие на эти прогнозы, должны быть понятны для преподавателей и администраторов, принимающих решения;
- результаты прогнозирования по модели не должны сильно изменяться на незначительных изменениях входных данных;
- отсутствие предвзятости и дискриминации по отношению к различным группам учащихся;
- адаптируемость к изменяющимся потребностям образовательной среды.

3. Статистический анализ данных мониторинга успеваемости

3.1. ОПИСАНИЕ ДАННЫХ

Исходный датасет содержит оценки по разным элементам дисциплины «Анализ данных», прочитанной двумя преподавателями. Баллы за тесты с автоматической проверкой сформированы в том числе по произвольным, созданным преподавателем, шкалам. Датасет с промежуточными и итоговыми баллами учащихся за первый семестр содержит 20 признаков: два текстовых признака с фамилиями и именами; номер группы; и две группы количественных признаков с оценками во время первой и второй половины семестра за домашние задания, контрольные, проекты и тесты, размещенные на платформе электронного обучения. Остальные количественные признаки содержат баллы, выставленные преподавателями за навыки программирования, активность на семинарах, а также итоговые баллы. Для визуализации количественных признаков использована диаграмма «ящик с усами» (рис. 1).

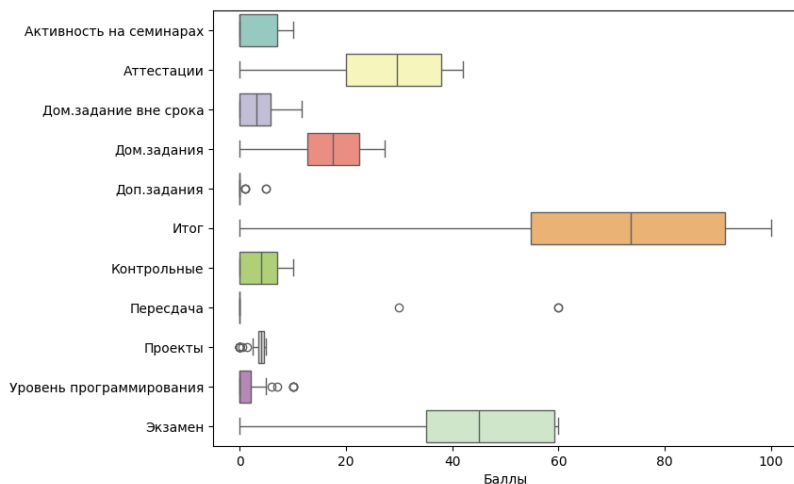


Рис. 1. Сравнение баллов по различным видам работ в семестре

Диаграмма содержит набор ящиков с усами, отображающих распределение баллов учащихся по различным видам работ. Каждый ящик представляет один вид работы и показывает медиану, верхний и нижний квартили, а также наличие выбросов в баллах. Таким образом, диаграмма позволила увидеть разные оценочные шкалы для разных видов работ, незначительное количество учащихся, выполнивших дополнительные задания или пересдавших экзамен. Медиана экзаменационных баллов равна 45, а итоговых баллов – 73,5.

3.2. ПОСТАНОВКА ЗАДАЧИ

Математическая постановка задачи может быть выражена следующим образом: пусть n – количество учащихся, а m – количество признаков успеваемости каждого учащегося. Тогда для каждого i -го учащегося имеем вектор признаков успеваемости $x_i = (x_{i1}, x_{i2}, \dots, x_{im})$, где x_{ij} представляет собой значение признака. Увеличение размерности пространства признаков происходит за счет нормализации существующих и создания новых признаков, расширяющих пространство $x_i \in R^d$, где $d > m$.

Для выделения учащихся в группу академического риска обозначим бинарную переменную принадлежности i -го студента к группе академического риска (1 – находится в группе академического риска, 0 – не находится) как y_i . Тогда средние значения признаков для каждой группы обозначим μ_0 и μ_1 . Направление оптимальной гиперплоскости, разделяющей две группы учащихся по академическому риску, задаем с помощью критерия линейного дискриминантного анализа:

- (1) $w = S_w^{-1} \Delta\mu$,
- (2) $S_w = cov(x_1) + cov(x_0)$,
- (3) $\Delta\mu = \mu_0 - \mu_1$,

где w – вектор весов, который максимизирует отношение двух классов, S_w – обобщенная ковариационная матрица.

Линейный дискриминантный анализ является более устойчивым к выбросам, чем логистическая регрессия, и достигает хорошей точности даже при малых размерах выборки. Но этот метод требует подготовки данных следующим образом: результи-

рующий признак должен быть категориальным; входные признаки должны быть распределены близко к стандартному нормальному распределению, чтобы их дисперсии были близки.

Прогнозирование итогового балла за семестр проведем с помощью линейно-регрессионной модели, где для каждого i -го студента строим модель успеваемости:

$$(4) \quad y_i = \alpha_0 + \alpha_1 x_{i1} + \alpha_2 x_{i2} + \dots + \alpha_d x_{id} + e_i,$$

где y_i – итоговый балл за семестр; $\alpha_0, \alpha_1 x_{i1}, \alpha_2 x_{i2}, \dots, \alpha_d$ – коэффициенты регрессии; e_i – случайная ошибка.

Для поиска учащихся со схожими траекториями успеваемости при тех же вводных (n – количество учащихся, m – количество признаков для каждого студента, тогда X – матрица оценок учащихся размерности $n \times m$, где каждая строка представляет собой набор оценок для одного студента) найти k групп учащихся, объединенных вместе на основании схожести их оценок, с целью формирования групповых траекторий успеваемости.

Таким образом, математическая постановка задачи включает в себя увеличение размерности пространства признаков, построение модели для прогнозирования успеваемости и последующую индивидуализацию консультаций на основе схожести траекторий успеваемости учащихся.

4. Реализация

4.1. ОБЪЕДИНЕНИЕ ДАННЫХ ПО ГРУППАМ УЧАЩИХСЯ

Набор оценок на обучающей платформе формируется для каждой группы учащихся. Поэтому реализовано слияние наборов оценок разных групп учащихся в единый датасет. После слияния отсутствующие значения заменены на ноль. Этот шаг обеспечил согласованность данных и позволил проводить дальнейший анализ без сбоев, вызванных отсутствием данных.

Процесс слияния и очистки данных закладывает основу для проведения всестороннего анализа и получения значимой информации путем ее консолидации из нескольких источников.

4.2 АНАЛИЗ ПРОПУСКОВ И НУЛЕЙ

Пропуски в данных связаны с отсутствием у студента баллов за активность. Причиной отсутствия баллов может быть пропуск

контрольной по уважительной или неуважительной причине, отсутствие ответов на семинарах, невыполнение курсовой работы, а также непроведение какого-то типа активности преподавателем в данной группе. Визуализация нулевых значений признаков представлена на рис. 2.

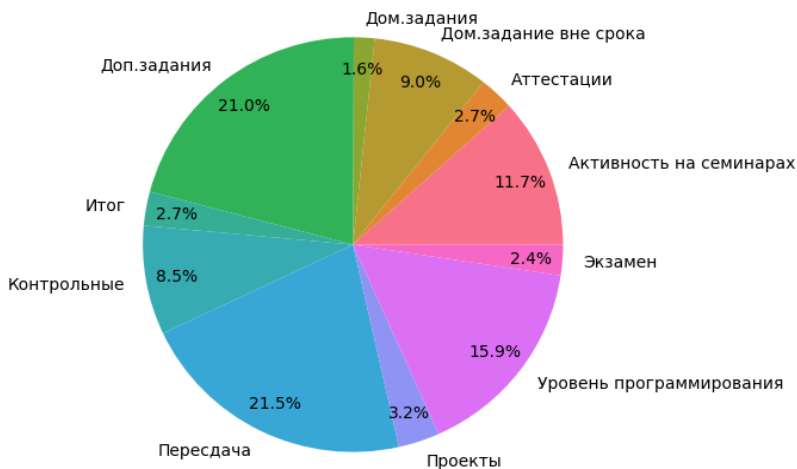


Рис. 2. Визуализация нулевых значений

Признаками со значительным количеством нулей являются выполнение дополнительных заданий, пересдача экзамена, уровень программирования и активность на семинарах.

4.3 ИЗМЕНЕНИЕ ПРОСТРАНСТВА ПРИЗНАКОВ

Разница в производительности методов машинного обучения напрямую зависит от предварительной обработки входных данных. Поэтому для повышения производительности модели используют различные методы проектирования признаков [4]. В работе на основе имен и окончаний фамилий добавлен бинарный атрибут Пол. Для анонимизации учащихся введен признак под названием Index, состоящий из первых букв фамилии, имени, пола и номера группы учащегося. Установлено, что баллы за тестовые, проектные и домашние задания варьируются от 0 до 5,

баллы за промежуточные оценки варьируются от 0 до 22, экзаменационные баллы – от 0 до 60, итоговые баллы – от 0 до 100. Таким образом, все баллы нормированы в диапазон от 0 до 1 делением на соответствующие максимальные значения. Это преобразование приводит к появлению нормированных признаков, которые характеризуют академическую успеваемость по шкале от 0 до 1 (рис. 3).

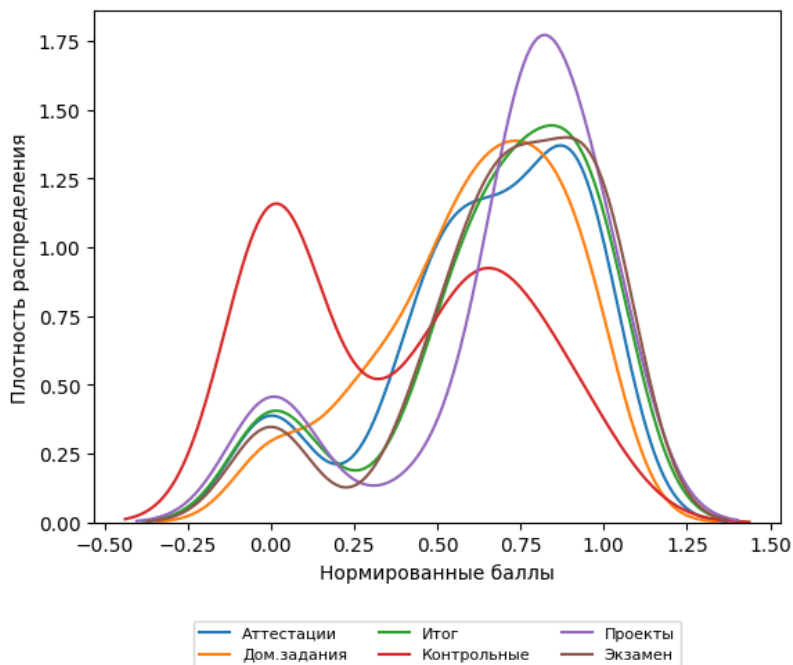


Рис. 3. Распределения нормированных признаков

Более низкая (по сравнению с левой) правая вершина распределения баллов за контрольные работы говорит о том, что успеваемость большинства учащихся по этой активности недостаточно хороша.

Для учета динамики успеваемости по одноименным активностям вычислены разницы между нормированными баллами

во второй и первой половине семестра. Распределения значений полученных признаков представлены на рис. 4.

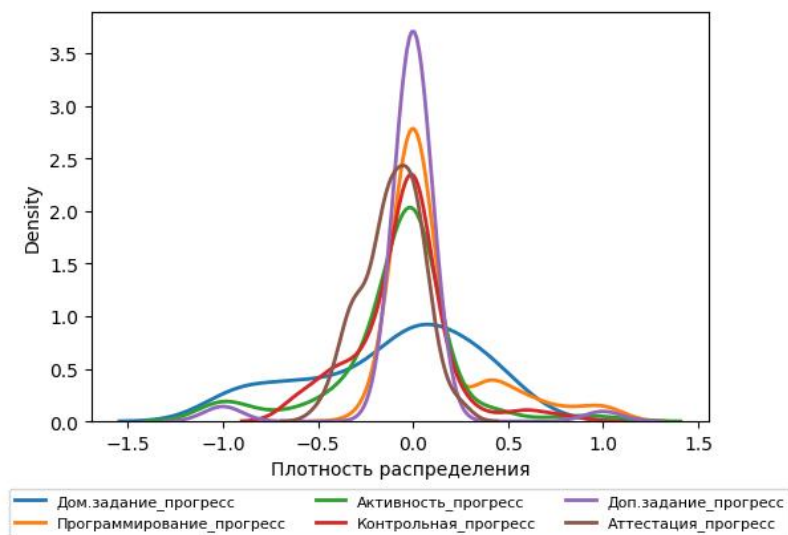


Рис. 4. Распределения признаков, учитывающих динамику успеваемости

Характер распределений новых признаков стал ближе к нормальному с незначительным смещением математического ожидания в отрицательную область у всех признаков кроме уровня программирования. Наличие отрицательных значений свидетельствует о снижении баллов у ряда учащихся во второй половине семестра. Нормализация распределений признаков улучшает сходимость алгоритмов, особенно когда диапазоны значений исходных признаков значительно отличаются.

Использование динамических признаков улучшило интерпретируемость модели за счет уменьшения числа входных признаков и более точном определении значимости каждого признака для классификации. Введение зашумленных признаков улучшило обобщающую способность модели.

4.4. ВЫДЕЛЕНИЕ ГРУППЫ УЧАЩИХСЯ ПОД АКАДЕМИЧЕСКИМ РИСКОМ ПОД АКАДЕМИЧЕСКИМ РИСКОМ

Для выделения группы учащихся под академическим риском использован линейный дискриминантный анализ. Метод создает линейную комбинацию признаков, которая максимально разделяет два класса учащихся, позволяя классифицировать новых студентов как тех, кто скорее всего сдаст экзамен, и тех, кто, возможно, не сдаст его. При подготовке данных целевая переменная преобразована в бинарную.

Полученный набор данных разделен в пропорции 80/20 на обучающий и тестовый. Причем в обучающий и в тестовый наборы попали студенты, не сдавшие экзамен. Для избежания переобучения и улучшения обобщающей способности модели с помощью библиотеки FunctionTransformer значения признаков обучающего набора данных изменены добавлением нормально распределенного случайного шума [5]. Диаграмма рис. 5 показывает желтыми треугольниками обученные точки и сравнивает местоположение имеющихся тестовых точек в виде серых прямоугольников и прогнозируемых точек в виде зеленых треугольников.

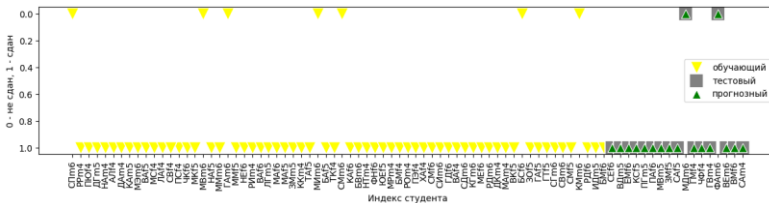


Рис. 5. Классификация учащихся на сдавших и не сдавших экзамен

Результаты классификации экзаменационных оценок выглядят следующим образом:

- константа дискриминации $C = 9,241$;
- уравнение дискриминационной функции: $D(x) = -5,908x_1 - 2,163x_2 - 4,470x_3 + 0,893x_4 + 1,912x_5 - 0,237x_6 + 0,174x_7 + C$, где $x_1, \dots, x_7$ – признаки, расположенные по возрастанию влияния, такие как группа учащегося (что говорит о неоднородности групп,

а также влиянии экспертных оценок), приращение оценок по домашним заданиям, пол, приращение активности на семинарах и дополнительных заданиях и др.;

- среднее гармоническое значение точности и полноты, представленное оценкой $F1 = 92,86\%$, означает, что модель способна хорошо разделять классы, дает низкую долю ложноположительных предсказаний и может правильно классифицировать большинство положительных и отрицательных случаев;

- оценка среднего гармонического значения точности и полноты для исходных данных (без динамических признаков) составляет $83,61\%$.

Эти результаты отражают потенциальную эффективность модели при двухклассовой классификации успеваемости.

4.5. ПРЕДСКАЗАНИЕ ИНДИВИДУАЛЬНОЙ ТРАЕКТОРИИ УСПЕВАЕМОСТИ

Нормализованные признаки разбиты на обучающую (80%) и тестовую (20%) части. Экземпляр модели линейной регрессии обучается на большей выборке и делает прогнозы для тестовой части. Результат прогнозирования части экзаменационных оценок визуализирован с помощью точечной диаграммы (рис. 5).

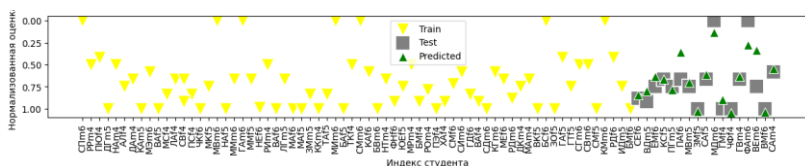


Рис. 5. Прогноз результатов экзамена

Для оценки работоспособности модели линейной регрессии на тестовых данных рассчитаны метрики ошибок: $MAE = 0,108$, $MSE = 0,024$, $RMSE = 0,156$, $R^2 = 0,712$. Модель гребневой регрессии снизила значения ошибок: $MAE = 0,099$, $MSE = 0,023$, $RMSE = 0,102$, $R^2 = 0,727$. Для имеющегося набора данных модель показывает высокую точность, поскольку известные тестовые и прогнозируемые значения достаточно близки, а ошибки не превышают 10%.

Индивидуальные траектории учащихся с прогнозом следующей оценки представлены на рис. 6.

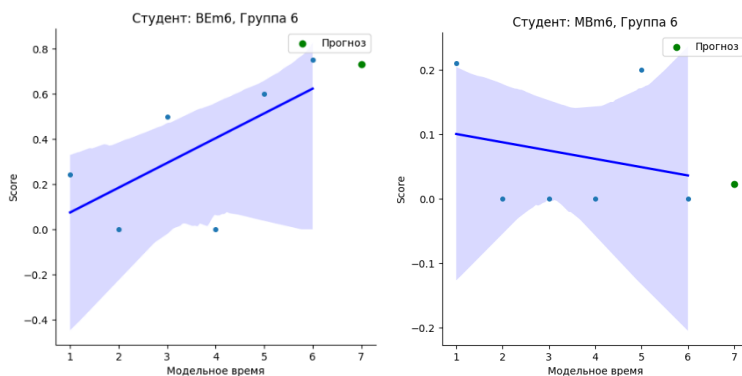


Рис. 6. Прогноз индивидуальных траекторий успеваемости:
а) растущий тренд; б) снижающийся

Графики на рис. 5 демонстрируют разный подход к обучению у двух учащихся одной группы. Второй учащийся нуждается в дополнительных консультациях.

4.6. ФОРМИРОВАНИЕ ГРУППОВЫХ ТРАЕКТОРИЙ УСПЕВАЕМОСТИ

При создании персонализированных учебных планов, выявлении успешных/неуспешных стратегий обучения удобно выделить группы учащихся со схожими траекториями успеваемости. Для этого использован метод DBScan на том же размеченном наборе данных. Метод кластеризует данные по различной плотности и отличает плотные регионы от разреженных областей, обрабатывая выбросы и шум. Результаты кластеризации представлены на рис. 7.

DBScan делит данные на три категории: основные точки, которые имеют достаточное количество соседних точек данных в пределах заданного радиуса эpsilon (ϵ); граничные точки, которые находятся в пределах ϵ -радиуса от основной точки, но не имеют достаточного количества соседей, чтобы считаться основными точками; и шумовые точки, которые не принадлежат

ни к одному кластеру из-за их изолированности и недостаточной близости к другим точкам.

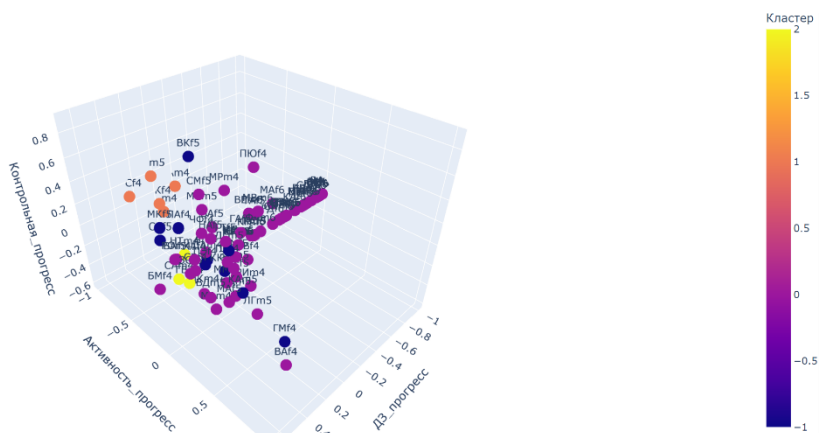


Рис. 7. Кластеризация учащихся по траекториям успеваемости

После применения DBSCAN для кластеризации данных, получены следующие результаты по количеству точек в каждом кластере: кластер –1: 10 точек, кластер 0: 66 точек, кластер 1: 5 точек, кластер 2: 3 точки. Эти результаты могут быть объяснены следующим образом: кластер –1 содержит шумовые точки. Значимые кластеры 0, 1 и 2 вместе содержат 74 точки. Точки в этих кластерах характеризуются общим поведением.

Таким образом, кластеры определены на основе локальной плотности точек и пространственных отношений, что преодолевает традиционные ограничения кластеризации.

5. Выводы

Обзор литературы посвящен прогнозированию траекторий успеваемости учащихся с использованием методов машинного обучения. Недостатками практической реализации предлагаемых методов являются значительное количество качественных признаков при небольшой размерности пространства признаков, сильное влияние экспертов, недостаточный учет динамики успеваемости.

Проведен статистический анализ признаков e-learning-платформы, включающий оценку дисбаланса данных, выявление пропусков, построение плотности распределений и изучение корреляционной матрицы.

Предлагаемый подход включает 4 этапа: вычисление приращений разновременных одноименных признаков и нормализация набора данных, формирование рискованной группы учащихся, которые могут не сдать предмет, прогноз оценки по предмету для каждого учащегося, формирование множества групп учащихся с близкими компетенциями.

Для повышения точности прогнозных моделей изменена размерность пространства признаков за счет приведения баллов за разные активности к общей шкале и создания новых признаков, таких как индексы учащихся и приращение баллов по разным активностям в процессе обучения.

При решении задачи кластеризации для поиска учащихся со схожими траекториями успеваемости, в которой количество кластеров заранее неизвестно, в качестве дополнительных признаков использованы результаты решения задачи классификации – выделения учащихся под академическим риском – и решения задачи прогнозирования – расчет итоговой оценки студента.

Результаты прогнозирования по модели на зашумленных данных и с использованием дополнительно сгенерированных данных изменились в пределах 7%.

Вопрос адаптируемости предложенного четырехэтапного подхода к изменяющимся потребностям образовательной среды требует дальнейшего исследования.

Упреждающее выявление и поддержка учащихся с риском неуспеваемости может помочь учебным заведениям адаптировать мероприятия и поддержку для учащихся, которые больше всего в этом нуждаются.

Признавая разнообразные потребности студенческих групп и соответствующим образом адаптируя поддержку, исследование направлено на создание более персонализированной и эффективной учебной среды.

Примеры данных и кода представлены на странице <https://github.com/avladova/Student-performance-prediction>.

Литература

1. АРЗАМАСЦЕВ С. А., БГАТОВ М., СЕМЕНЧИКОВ Д. *Предсказание оттока абонентов: сравнение методов машинного обучения* // Computer Tools in Education. – 2018. – №5. – С. 5–23.
2. ВЛАДОВА А.Ю. *Кластерный анализ изменений пространственного положения трубных секций магистрального нефтепровода по данным внутритрубных обследований* // Безопасность труда в промышленности. – 2018. – №1. – С. 22–25.
3. ВЛАДОВА А.Ю. *Формирование пространства признаков и авторегрессионных моделей для прогноза отступлений железнодорожного полотна* // Проблемы управления. – 2023. – Вып. 2. – С. 54–64.
4. ВЛАДОВА А.Ю., ШЕК Е.Д. *Подготовка данных для машинного анализа ключевых показателей эффективности территориальных менеджеров* // Бизнес-информатика. – 2021. – №3. – С. 48–59.
5. ЗАДАДАЕВ С.А., ОРЛОВА И.В. *Опыт применения эконометрического инструментария для прогнозирования показателей национальных целей развития РФ* // Фундаментальные исследования. – 2022. – №10-1. – С. 54–59. – DOI: 10.17513/ft.43343.
6. КУРАВСКИЙ Л.С., ЮРЬЕВ Г.А., ЗЛАТОМРЕЖЕВ В.И. и др. *Оценка действий экипажа воздушного судна на основе модели рисков человеческого фактора* // Экспериментальная психология. – 2020. – Т. 13, №2. – С. 153–181. – DOI: 10.17759/exppsy.2020130211.
7. BAKHTADZE N., ZAIKIN O., ZYLAWSKI A. и др. *Incentive Model of a Project Learning Process* // 7th Int. Conf. on Frontiers of Industrial Engineering, ICFIE-2020: Virtual, Singapore, 2020. – P. 73–81. – DOI: 10.1109/ICFIE50845.2020.9266738.
8. BAYAZIT A., ILGAZ H., ERDEN S. *Profiling students via clustering in a flipped clinical skills course using learning analytics* // Medical Teacher. – 2022. – №7(45). – С. 724–731.

9. KIM J., CHO J. *DBSCAN-Based Tracklet Association Annealer for Advanced Multi-Object Tracking* // *Sensors*. – 2021. – Vol. 21. – P. 5715.
10. LIU M., YU D. *Towards intelligent E-learning systems* // *Education and Information Technologies*. – 2023. – No. 7(28). – P. 7845–7876.
11. NAGY V., DUMA L. *Measuring efficiency and effectiveness of knowledge transfer in e-learning* // *Heliyon*. – 2023. – No. 7(9).
12. OLUWADELE D., SINGH Y., ADELIYI T. *E-Learning Performance Evaluation in Medical Education – A Bibliometric and Visualization Analysis* // *Healthcare*. – 2023. – No. 11. – P. 232.
13. TROUSSAS C., VIRVOU M., ALEPIS E. *Comulang: towards a collaborative e-learning system that supports student group modeling* // *SpringerPlus*. – 2013. – No. 1(2). – P. 387.
14. YE M., SHENG X. *SA-FEM: Combined Feature Selection and Feature Fusion for Students' Performance Prediction* // *Sensors*. – 2022. – No. 22(22). – P. 8838.

DEVELOPING GROUP AND INDIVIDUAL PERFORMANCE PATHS BASED ON E-LEARNING PLATFORM DATA

Alla Vladova, V.A. Trapeznikov Institute of Control Sciences of RAS, Moscow, Financial University under the Government of Russian Federation, Dr.Sc. (avladova@ipu.ru).

Abstract: Maintaining a high level of education is a key task in university management. Despite continuous monitoring of student performance, educational institution management fails to adequately utilize performance forecasting methods when shaping student learning paths. The proposed approach differs from existing ones in several aspects. Firstly, it analyzes features containing grades for various assignments completed by students on the e-learning platform, expanding the feature space by normalizing grades on a single scale and creating new features: an index and changes in performance for different types of assignments. Secondly, it identifies students at academic risk. Thirdly, it predicts exam scores for each student using a linear regression model. Fourthly, it groups students with similar learning trajectories for personalized consultations. The approach to predicting exam results for individual students demonstrates a commitment to providing comprehensive support beyond simple assessment. Through analysis, modeling, and personalized consultations, the research aims to proactively enhance academic performance in university settings.

Keywords: academic performance, statistical analysis, linear discriminant analysis, regression, learning trajectories, distant learning system.

УДК 65.012.4 + 681

ББК 65.9

DOI: 10.25728/ubs.2024.111.7

*Статья представлена к публикации
членом редакционной коллегии Я.И. Квинто.*

*Поступила в редакцию 27.03.2024.
Опубликована 30.09.2024.*

**ОРГАНИЗАЦИОННО-ЭКОНОМИЧЕСКИЙ
МЕХАНИЗМ ВНЕДРЕНИЯ ТЕХНОЛОГИЙ
АВТОМАТИЗАЦИИ УПРАВЛЕНИЯ
ВОЗДУШНЫМИ СУДАМИ
В РАМКАХ АВИАТРАНСПОРТНЫХ СИСТЕМ**

Клочков В. В.^{1а,б}, Егшин С. Ф.^{2б}

(^а ФГБУН Институт проблем управления

им. В.А. Трапезникова РАН, Москва,

^б ФГБУ НИЦ «Институт им. Н.Е. Жуковского»,

Жуковский

В современных условиях исследование крупномасштабных систем позволяет существенно повысить эффективность отраслевого управления. Рассматриваемый на примере авиационной отрасли подход позволяет выработать обоснование для организационного механизма, который отличается от традиционных бизнес-моделей применения авиации возможностью достигнуть максимального экономического эффекта с учетом соблюдения современных требований к авиационным работам и услугам в части безопасности полетов. В основу механизма положено априорное принятие авиапроизводителем ответственности за все отказы, возникающие в процессе эксплуатации новой авиационной техники на начальном этапе ее жизненного цикла. Методология исследования базируется на математическом моделировании и методах системного анализа. В качестве примера рассмотрена отрасль перспективного применения беспилотных авиационных систем «доставка последней мили». Предложенный подход позволяет выполнять более эффективные планирование развития и выработку управленческих решений (стратегий управления) при внедрении новой авиационной техники, обладающей потенциалом широкого применения, в том числе при формировании государственных программ Российской Федерации, стратегий развития промышленных организаций авиастроения.

Ключевые слова: беспилотные авиационные системы, беспилотные летательные аппараты, конструктивно-производственные недостатки, доставка «последней мили», управление развитием отрасли.

¹ Владислав Валерьевич Клочков, д.э.н. (klochkovvv@nrczh.ru).

² Сергей Федорович Егшин, специалист (egoshinsf@nrczh.ru).

1. Введение

При современном уровне развития авиации, широкомасштабное внедрение новой авиационной техники (АТ) невозможно вне сложных организационно-технических систем [5] и ведет к изменению не одного, а сразу нескольких, как правило, взаимосвязанных показателей функционирования этих систем.

С одной стороны, в современной авиации подавляющее большинство авиационных происшествий и инцидентов происходит по причине «человеческого фактора» [13, 14], при всей неоднозначности этого термина (поскольку в авиации как рукотворной области всё происходит по вине и воле человека). Поэтому в качестве магистрального направления развития технологий с целью повышения уровня безопасности полетов (БП) рассматривается автоматизация управления воздушных судов (ВС) – как полная (переход к автономному управлению), так и частичная (снижение нагрузки на экипаж).

С другой стороны, при внедрении высокоавтоматизированных или даже полностью автономных ВС по определению возрастает доля авиационных происшествий (АП) и инцидентов, вызванных отказами средств автоматизации управления, сбоями, некорректной работой алгоритмов автоматизации управления. И неверный алгоритм управления или некорректно обученная интеллектуальная система управления уже будут относиться к «отказам» АТ, вызванным ее конструктивно-производственными недостатками (КПН). И вина за эти «отказы», разумеется, лежит на разработчиках и производителях АТ. Тогда как ошибки управления ВС, допускаемые их экипажами в настоящее время, в основном относятся к сфере ответственности эксплуатантов АТ.

При этом новизна технологий глубокой автоматизации управления ВС, непрозрачность интеллектуальных методов и систем управления вызывают справедливые опасения общества и авиационных властей [17, 18], что по крайней мере на начальных стадиях внедрения таких технологий частота вызванных этими «отказами» авиационных происшествий и инцидентов

будет высокой. Это один из главных барьеров на пути внедрения соответствующих технологий.

В качестве еще одного противоречия может быть указано, что внедрение технологий автоматизации управления ВС несет не только высокие риски для ухудшения уровня БП, но и ведет к изменению стоимостных характеристик АТ. И поскольку каждое АП так или иначе ведет к убыткам для непосредственных участников авиационной деятельности (производителей и эксплуатантов АТ и пр.) и третьих лиц, данное противоречие также становится существенным сдерживающим фактором развития отрасли: участники авиационной деятельности в стремлении сократить собственные убытки придерживаются выжидательной стратегии вместо активного внедрения инноваций.

Поиск организационных механизмов, которые способствовали бы ускорению развития предприятий соответствующих отраслей, не теряет своей актуальности [8, 10, 12, 15, 16]. В настоящей работе рассматриваются беспилотные авиационные системы (БАС) как одна из активно развивающихся отраслей, затрагивается проблема устранения КПН и демонстрируется эффективность соответствующего механизма управления. Суть механизма заключается в том, что один из участников отрасли выступает в роли единого оператора, максимизирующего общую выгоду, тем самым способствуя не только ускоренному росту отрасли, но и ускоренному снижению количества АП в отрасли вследствие более быстрого накопления необходимого опыта, т.е. устранения КПН и причин эксплуатационных ошибок.

2. Моделирование отрасли внедрения БАС

Для обоснования действенности рассматриваемого механизма и раскрытия искомого эффекта использован подход, аналогичный предложенному в работе [7], где он был использован для анализа эффективности стратегии конкретного разработчика авиационных двигателей в части устранения КПН и в части организации технического обслуживания и ремонта (ТОиР) с использованием конкретной выборки исходных данных.

В настоящей работе указанный подход применяется в более широком смысле: для анализа границ применимости организационного механизма внедрения радикальных инноваций, связанных с БАС, в условиях высоких рисков для БП, вызванных переходом к дистанционному или автономному управлению движением входящих в состав БАС беспилотных воздушных судов (БВС), а также субъективным восприятием этих рисков в обществе.

Отрасль применения БАС, как и в [7], моделируется в упрощенной форме как состоящая из производителя АТ и эксплуатанта АТ.

Производитель АТ в рассматриваемый год t выпускает $q(t)$ штук БАС при продажной цене p . Количество выпускаемых БАС зависит от доли прибыли, выделяемой эксплуатантом АТ для закупки новых БАС (о чем подробнее будет сказано ниже). При этом прибыль производителя $\Pi^{\text{произв}}(t)$, рассматриваемая накопительным итогом к году t , есть разница между выручкой от продажи суммарного количества БАС $Q(t)$, выпущенных начиная с $t = 1$, и понесенными суммарными расходами на проектирование и производство БАС:

$$(1) \quad \Pi^{\text{произв}}(t) = p \times Q(t) - TC(Q(t)) - c_{\text{отн}}^{\text{произв}}(t),$$

где $TC(Q)$ – суммарные затраты на производство Q штук БАС; $c_{\text{отн}}^{\text{произв}}(t)$ – накопленные к концу года t затраты производителя, связанные с отказами.

Суммарные затраты TC включают три составных части:

- фиксированные затраты $FC_{\text{произв}}$ на разработку нового изделия и технологическую подготовку производства (ТПП);
- стоимость материалов для изготовления Q штук БАС, рассчитываемая пропорционально удельной стоимости материалов $c_{\text{мат}}$ для изготовления одного экземпляра БАС;
- стоимость трудозатрат $c_{\text{тр}}(Q)$ для выпуска Q штук БАС, при расчете которой учитывается накопление опыта благодаря эффекту обучения персонала производителя:

$$(2) \quad c_{\text{тр}}(Q) = c_{\text{тр}}^1 \sum_{q=1}^Q \left(1 - \lambda_{\text{произв}}\right)^{\log_2 Q},$$

где $c_{тр}^1$ – стоимость трудозатрат на сборку первого серийного экземпляра; $\lambda_{произв}$ – темп обучения в процессе производства.

Что касается затрат производителя, которые были накоплены к концу года t в связи с отказами, то в данном случае возможны два варианта, о чем будет сказано ниже.

Эксплуатант АТ в рассматриваемый год t эксплуатирует все БАС, выпущенные к этому году (в гипотетическом предположении, что повреждения БВС в авариях и катастрофах могут быть устранены и БВС возвращено к эксплуатации) в количестве $m(t)$:

$$(3) \quad m(t) \equiv Q(t-1) = m(1) + \sum_{s=1}^{t-1} q(s),$$

где $m(1)$ – объем стартового заказа к началу массовой эксплуатации.

Прибыль эксплуатанта $\Pi_{\text{экс}}(t)$, рассматриваемая накопительным итогом к году t , есть разница между выручкой, пропорциональной годовому налету η одного БВС и продажной цене r летного часа БВС для потребителя услуги, и расходами, которые включают себестоимость летного $c_{\text{час}}$ и могут включать затраты $c_{\text{отн}}^{\text{экс}}$ на устранение последствий отказов:

$$(4) \quad \Pi^{\text{экс}}(t) = \sum_{s=1}^t (r - c_{\text{час}} - c_{\text{отн}}^{\text{экс}}) \times \eta \times m(s).$$

Отсюда количество $q(t)$ новых БАС, закупаемых в год t , составит:

$$(5) \quad q(t) = \left\lfloor \frac{\gamma \times (r - c_{\text{час}} - c_{\text{отн}}^{\text{экс}}) \times \eta \times m(t)}{p} \right\rfloor,$$

где γ – доля прибыли, выделяемой эксплуатантом АТ для закупки новых БАС; $\lfloor \cdot \rfloor$ – целая часть числа.

При эксплуатации БАС рассматривается возможность возникновения двух типов отказов: вследствие ошибок эксплуатации и вследствие КПН. В первом случае предполагается, что поток отказов однороден и равномерен, а потому может быть задан через фиксированную среднюю наработку на отказы $\tau_{\text{др}}$.

Во втором случае, средняя наработка на отказ из-за КПН непостоянна и возрастает по закону:

$$(6) \quad \tau_{\text{КПН}}(t) = \tau_{\text{КПН}}(1) \times (1 - \lambda_{\text{надежн}})^{-\log_2 N_{\text{КПН}}(t)},$$

где $\tau_{\text{КПН}}(1)$ – средняя наработка на отказ из-за КПН в первый год эксплуатации; $\lambda_{\text{надежн}}$ – темп обучения в процессе устранения КПН; $N_{\text{КПН}}(t)$ – накопленное к году t суммарное количество отказов из-за КПН.

Введением зависимости (6) учитывается накопление опыта разработки и эксплуатации беспилотной АТ с постепенным повышением уровня безопасности эксплуатации.

В рамках модели производится сравнение двух возможных сценариев развития некоторой отрасли применения БВС:

– «сценарий 1», при котором производитель и эксплуатант несут самостоятельную ответственность за инциденты в соответствии со степенью их виновности;

– «сценарий 2», при котором производитель априорно принимает на себя всю ответственность за инциденты.

В «сценарии 1» затраты производителя и эксплуатанта, связанные с отказами, в рассматриваемый год t в расчете на летный час есть:

$$(7) \quad c_{\text{отн}}^{\text{произв}}(t) = \frac{c_{\text{восст}} + c_{\text{деф}} + c_{\text{отв}}}{\tau_{\text{КПН}}(t)},$$

$$(8) \quad c_{\text{отн}}^{\text{эксп}}(t) = \frac{c_{\text{восст}} + c_{\text{деф}} + c_{\text{отв}}}{\tau_{\text{др}}},$$

где $c_{\text{восст}}$, $c_{\text{деф}}$ и $c_{\text{отв}}$ – стоимости восстановления БВС, дефектации БВС и установления ответственности соответственно.

При таком расчете затрат производитель несет ответственность с периодичностью, зависящей только от $\tau_{\text{КПН}}$, а эксплуатант – от $\tau_{\text{др}}$, однако они оба несут дополнительные издержки $c_{\text{отв}}$, связанные с необходимостью определения степени собственной вины в каждом отдельном случае отказа.

В «сценарии 2» эти же затраты производителя и эксплуатанта задаются как

$$(9) \quad c_{\text{отн}}^{\text{произв}}(t) = (c_{\text{восст}} + c_{\text{деф}}) \times \left(\frac{1}{\tau_{\text{КПН}}(t)} + \frac{1}{\tau_{\text{др}}} \right),$$

$$(10) \quad c_{\text{отн}}^{\text{эсп}}(t) = 0.$$

Выражения (9)–(10) означают, что эксплуатант не несет каких-либо издержек, связанных с устранением отказов, поскольку все они ложатся на производителя, и при этом исчезают издержки $c_{\text{отв}}$, связанные с определением степени ответственности за каждый случай отказа.

Таким образом, целью сравнительных расчетов с учетом формул (7)–(8) или (9)–(10) является выработка условий, при которых априорное принятие на себя ответственности ведет к повышению суммарной прибыли в отрасли и уровня БП, и наоборот.

3. Исходные данные и допущения

В практическом плане, рассматривается такая отрасль возможного широкого применения БАС, как доставка «последней мили».

Вводится предположение, что БАС создаются на основе однотипного парка БВС. Тип применяемого БВС выбирается как «мультиротор легкий» из перечня планируемых к массовому применению [11]. Моделируемый период развития отрасли задается равным 10 годам.

Оценочные технико-экономические характеристики (ТЭХ) БАС представлены в таблице 1.

Расчетные значения параметров, указанные в таблице 1, за исключением взятых из [11], введены исходя из следующих допущений.

1. Полеты БВС выполняются в светлое время суток и с ограничением 8-часового рабочего дня оператора, поэтому, с учетом необходимых перерывов на техобслуживание и пр., годовой налет БВС 2000 л.ч. сопоставим с налетом самолетов малой авиации при их достаточно интенсивной эксплуатации.

Таблица 1. ТЭХ БАС

Характеристика	Обозначение	Значение
Объем стартового заказа к началу массовой эксплуатации, шт.	$m(1)$	100
Годовой налет, л.ч.	η	2000
Продажная цена БАС (включая БВС), млн руб.	p	0,5
Доля прибыли эксплуатанта, тратящаяся на закупку новых БАС, %	γ	50
Наработка на отказ из-за КПП в течение первого года эксплуатации, л.ч.	$\tau_{\text{КПП}}(1)$	548
Темп обучения в процессе устранения КПП	$\lambda_{\text{надежн}}$	0,15
Наработка на другие отказы, л.ч.	$\tau_{\text{др}}$	2000
Стоимость восстановления БВС, млн руб.	$c_{\text{восст}}$	0,3
Стоимость дефектации БВС, млн руб.	$c_{\text{деф}}$	0,1
Стоимость установления ответственности, млн руб.	$c_{\text{отв}}$	0,1
Продажная цена л.ч. для потребителя услуги, руб.	r	3500
Себестоимость л.ч., руб.	$c_{\text{час}}$	3000
Затраты на проектирование и ТПП, млн руб.	$FC^{\text{произв}}$	150
Стоимость материалов для изготовления одного экземпляра БАС, млн руб.	$c_{\text{мат}}$	0,133
Стоимость трудозатрат на сборку первого серийного экземпляра, млн руб.	$c_{\text{тр}}^1$	0,534
Темп обучения в процессе производства	$\lambda_{\text{произв}}$	0,15

2. Наработка на отказ из-за КПП в течение первого года эксплуатации БАС вводится по формуле

$$(11) \quad \tau_{\text{КПП}}(1) = \frac{\eta}{365} \times m(1).$$

Данная взаимосвязь означает, что при эксплуатации парка БВС в границах обслуживаемой территории (например, городской агломерации) в течение одного дня эксплуатации допустимо не более одного отказа по причине КПП. В этом случае, благодаря возможности подмены неисправного БВС (при значи-

тельной численности парка), не происходит срыва графика доставки грузов. При этом сотрудники ремонтного подразделения производителя успевают прибыть на место инцидента и устранить неисправность в течение суток, так что уже на следующий день БВС снова будет выполнять полеты, а подразделение – уже реагировать на новый возможный инцидент.

3. Нарботка на другие отказы задается как одно событие в год.

4. Сумма стоимости дефектации и стоимости восстановления не должна превышать 80% от продажной цены (остальные 20% – норма прибыли от продажи БАС): в противном случае производителю выгоднее изготовить и поставить новый комплект БАС, а не заниматься восстановлением неисправного/поврежденного. При этом стоимость дефектации составляет 20–25% от стоимости трудозатрат на сборку первого серийного экземпляра как условная стоимость разборки элементов БАС.

5. Стоимость установления ответственности выбирается из условия, что ее сумма со стоимостью дефектации и стоимостью восстановления не должна превысить продажную цену. В противном случае сложится ситуация, что в «сценарии 1» потери производителя вследствие отказа изделий будут всегда превышать его выручку от продажи изделий (хотя в модели и не затрагивается вопрос функционирования предприятий производителя в условиях постоянной убыточности на протяжении нескольких лет).

6. Считается, что заявленная стоимость летного часа БВС [11] включает в себя прибыль, норма прибыли при этом составляет 15–20%.

7. Затраты на проектирование и ТПП выбраны исходя из информации, что ценовой предел для сертификации БАС устанавливается в размере 30–70 млн руб. [9], отсюда стоимость $FC_{\text{произв}}$ будет кратно больше (например, в 3 раза).

8. Стоимость материалов для изготовления одного экземпляра БАС выбрана из условия, что при крупносерийном производстве стоимость материалов составит треть от себестоимости производства экземпляра, т.е. 26,67% от продажной цены (при норме прибыли 20%).

9. Стоимость трудозатрат на сборку первого серийного экземпляра БАС выбрана в продолжение предыдущих рассуждений. Оставшаяся часть продажной цены (53,33%) соответствует стоимости трудозатрат на сборку одного экземпляра при крупносерийном производстве. Тогда стоимость трудозатрат на сборку первого серийного экземпляра вследствие эффекта обучения будет как минимум вдвое больше, т.е. 106,67% от продажной цены.

4. Анализ результатов расчетов

Результаты расчетов представлены на рис. 1–6.

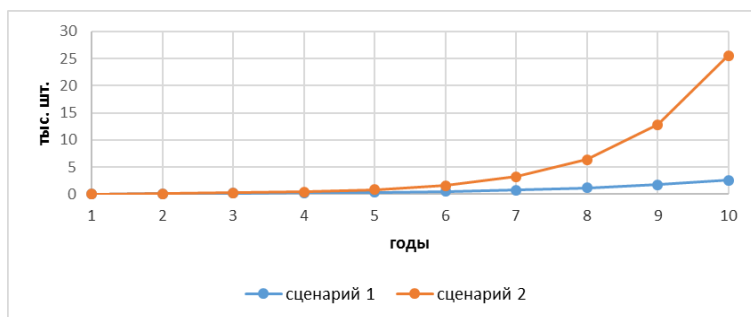


Рис. 1. Парк БВС

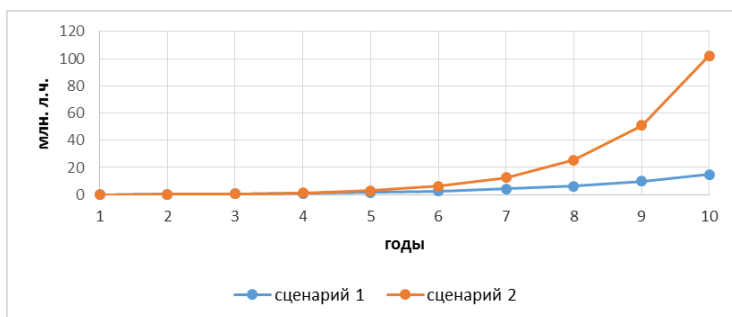


Рис. 2. Суммарный налет парка БВС к году эксплуатации

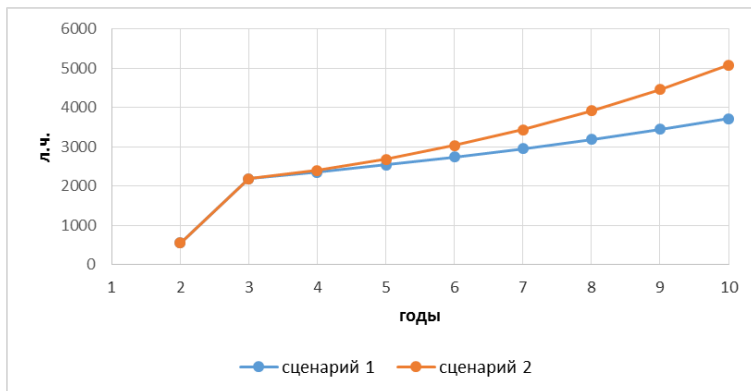


Рис. 3. Средний налет БВС на отказ вследствие КПН

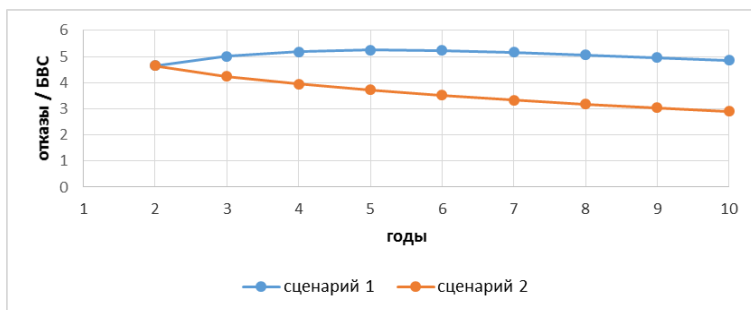


Рис. 4. Среднее количество накопленных отказов на одно БВС к году эксплуатации

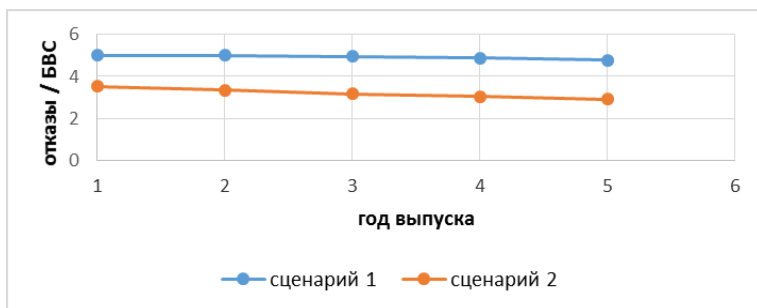


Рис. 5. Среднее количество суммарных отказов на одно БВС за жизненный цикл

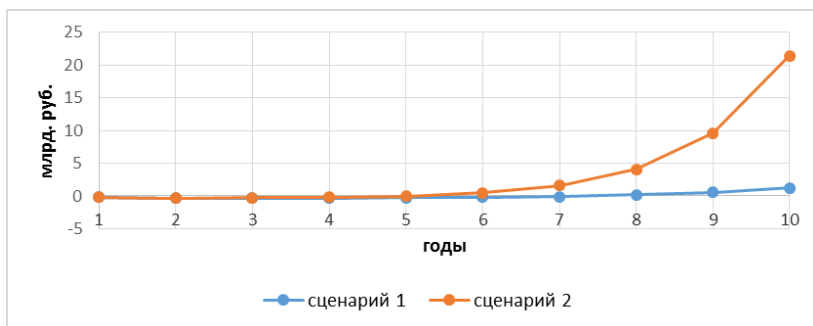


Рис. 6. Суммарно накопленная прибыль отрасли

Как можно видеть, характер роста парка БВС отрасли доставки «последней мили» (рис. 1) и суммарного налета этого парка (рис. 2) одинаков: при «сценарии 2» отрасль развивается гораздо быстрее, чем при «сценарии 1». Это связано с тем, что при заданных значениях параметров в «сценарии 2» отсутствуют дополнительные затраты, связанные с определением ответственности.

Средний налет на отказ из-за КПН (рис. 3) растет быстрее для «сценария 2», хотя итоговая разница за период моделирования невелика и не превышает 25%.

Среднее количество накопленных к году эксплуатации отказов на одно БВС в общем случае изменяется по закону, соответствующему «сценарию 1» (рис. 4): начальный рост сменяется дальнейшим падением. Однако для «сценария 2» эта закономерность не отображается на графике, поскольку период роста соответствует кратковременному промежутку между годами 2 и 3. В целом же к окончанию периода моделирования (10 лет) среднее количество накопленных отказов на одно БВС для «сценария 2» оказывается на 40% меньше, чем в «сценарии 1», что обусловлено более быстрыми темпами обучения в части устранения КПН при большем парке эксплуатируемых БВС.

Средние количества суммарных отказов на одно БВС за жизненный цикл (как период эксплуатации одного БВС) могут быть рассчитаны для БВС, выводимых из эксплуатации начиная с 11 года (в силу допущения, что назначенный ресурс БВС равен

20000 л.ч., что соответствует 10 годам эксплуатации при заданном годовом налете 2000 л.ч.). Соответственно, при увеличении периода моделирования до 15 лет данный показатель может быть рассчитан только для БВС, выпущенных в течение первых пяти лет (рис. 5). Увеличение же периода моделирования более чем на 15 лет ставит вопрос корректности применения модели (например, в силу насыщения рынка, т.е. наличия ограничения на спрос услуг БАС). Впрочем, качественные выводы могут быть сделаны и без этого, поскольку закономерности на рис. 5 повторяют закономерности с рис. 4 в части роста/падения, при этом в «сценарии 2» количество отказов за период эксплуатации на 30–40% меньше, чем в «сценарии 1».

Суммарно накопленная прибыль производителя и эксплуатанта для обеих отраслей (рис. 6) оказывается на порядок больше в случае «сценария 2».

Таким образом, по результатам моделирования показано, что механизм априорного признания ответственности производителями за возможные инциденты с участием производимой ими беспилотной АТ позволяет, при большем росте отрасли применения данной АТ, быстрее накапливать необходимый опыт, устранять угрозы БП и повышать надежность эксплуатируемой АТ, что на долгосрочных временных интервалах обеспечивает выигрыш в накопленной прибыли за жизненный цикл изделий АТ. Фундаментальной причиной данного эффекта является исключение в рамках объединенной отрасли такой статьи затрат, как «стоимость установления ответственности», величина которой может быть весьма существенной [2]. Как результат, при априорном принятии производителем ответственности на себя суммарная прибыль и рост уровня БП в отрасли, независимо от выбора значений параметров, будет всегда больше.

5. Анализ чувствительности системы к варьируемым параметрам

Суммарно накопленная прибыль производителя и эксплуатанта может быть выбрана в качестве основного критерия дея-

тельности, для которого может быть выполнен анализ чувствительности к варьируемым управляющим параметрам системы.

Анализ чувствительности строится на основе метода градиентов, широко применяемого при разработке новой авиационной техники (авиационных комплексов, авиационных систем) на этапе уточнения проектных решений [4].

Объект АТ рассматривается как система, и для критерия S_{Σ} этой системы проводится анализ, какой из параметров a_j , $1 \leq j \leq J$, оказывает наибольшее влияние на S_{Σ} , т.е. дает наибольшее приращение S_{Σ} :

$$(12) \Delta S_{\Sigma, j} = \frac{\partial S_{\Sigma}}{\partial a_j} \Delta a_j.$$

В ситуации с проектированием АТ по выявлении искомого наиболее значимого параметра конструктор должен направить усилия в первую очередь на улучшение значения именно этого параметра. Аналогичным образом любая прикладная наука может вырабатывать приоритетные управленческие решения, обеспечивающие наилучшее развитие исследуемых систем с точки зрения рассматриваемых критериев.

В то же время применение метода имеет некоторые особенности.

Во-первых, анализ выражений (12) фактически означает замену нелинейной многомерной поверхности (которая может быть немонотонной, периодической, разрывной) гиперплоскостью в пространстве размерности J .

В данной ситуации, во избежание значительных погрешностей, приращения Δa_j должны быть невелики по сравнению с абсолютным значением параметра a_j , а поверхность, соответствующая S_{Σ} , быть близка к монотонной и не иметь разрывов.

Во-вторых, приращения переменных Δa_j должны соответствующим образом соотноситься между собой. Это связано в том числе с тем, что в большинстве случаев параметры a_j не являются независимыми и вариация одного из них неизменно влечет изменение других.

Корректным шагом в данной ситуации становится введение функциональных зависимостей варьируемых параметров a_j от новых независимых переменных b_k так, что

$$(13) \Delta a_j = \sum_k \frac{\partial a_j}{\partial b_k} \Delta b_k,$$

причем задаваемые вариации Δb_k будут одинаковы для всех a_j . Однако применение формул (13) требует разработки достаточно сложной математической модели исследуемой системы, которая в большинстве случаев даже в упрощенной форме будет так или иначе привязана к обликам существующих систем, что не позволит варьировать a_j в достаточно широком диапазоне без нарушения законов, определяющих облик этих систем.

Поэтому на практике прибегают к следующему способу исследования, позволяющему получать приемлемый результат. Вводятся так называемые нормализованные переменные $\bar{a}_j$ [3]:

$$(14) \bar{a}_j = \frac{a_j - a_{j\min}}{a_{j\max} - a_{j\min}}, \quad 0 \leq \bar{a}_j \leq 1, \quad j=1, \dots, J,$$

где $a_{j\min}$, $a_{j\max}$ – нижний и верхний пределы изменения каждого параметра a_j соответственно. Целевую функцию S_Σ при этом также нормализуют следующим образом:

$$(15) \bar{S}_\Sigma = \frac{S_\Sigma}{S_\Sigma^0},$$

где S_Σ^0 – значение S_Σ , рассчитанное в исходной, базовой точке.

Наконец, в-третьих, нахождение производных в (12) может быть сильно затруднено в силу отсутствия/невозможности получения аналитических зависимостей S_Σ от варьируемых параметров.

Тогда при условии одинаковости и малости $\Delta \bar{a}_j$ (например, $\Delta \bar{a}_j \sim 0,01$ [4]) и с заменой производных на приращения:

$$(16) \Delta \bar{S}_{\Sigma,j} = \frac{1}{S_\Sigma^0} (S_{\Sigma,j} - S_\Sigma^0).$$

Корректность применения данного метода трактуется следующим образом: равенство приращений нормализованных переменных означает равенство «усилий» для достижения границ пространства.

Применительно к рассматриваемой отрасли корректность применения вышеуказанной методологии для суммарно накопленной прибыли производителя и эксплуатанта как критерия следует из рассмотрения характера функций (рис. 6). Если пренебречь целочисленностью некоторых модельных параметров, а также незначительным падением прибыли в первые годы в сравнении с итоговым значением последнего года, то можно считать суммарную прибыль непрерывной и монотонной функцией. Отсюда в качестве управляющих переменных могут быть выбраны параметры, приведенные в таблице 2 (также указаны границы их достижимых значений), а в качестве критерия – суммарно накопленная прибыль производителя и эксплуатанта в последний год периода моделирования.

Таблица 2. Управляющие параметры и ограничения их значений

Характеристика	Модельное обозначение	$a_{j \min}$	$a_{j \max}$
Годовой налет, л.ч.	η	2000	4000
Темп обучения в процессе устранения КПН	$\lambda_{\text{надежн}}$	0,15	0,25
Доля прибыли эксплуатанта, тратящаяся на закупку новых БАС, %	γ	50	100
Стоимость установления ответственности, млн руб.	$c_{\text{отв}}$	0,1	0
Начальный объем заказа к началу массовой эксплуатации, шт.*	$m(1)$	100	200 / 1000

* – в числителе – условное минимальное (двукратное) увеличение начального объема заказа, в знаменателе – условное максимальное (десятикратное) увеличение начального объема заказа.

Касательно выбора границ достижимых значений, указанных в таблице 2: для всех параметров в качестве $a_{j \min}$ взяты значения, при которых были проведены вышеприведенные расчеты, соответствующие рис. 1–6. Значения же $a_{j \max}$ для параметров выбирались следующим образом:

1. Годовой налет в 4000 л.ч. примерно соответствует полусуточному ежедневному налету: полеты совершаются все светлое время суток, прочие сопутствующие процедуры (ТОиР и др.) проводятся уже в темное время суток, когда безопасное пилотирование в условиях городской застройки становится невозможным. Списанием БВС вследствие выработки ресурса пренебрегается, так как количество БВС в парке, которые были произведены в первые годы (годы 1–5, рис. 1), мало по сравнению количеством БВС, выпущенных позже.

2. Темп обучения в процессе устранения КПН увеличен умозрительно до уровня «удвоение опыта приводит к снижению потока отказов вследствие КПН на 25%» (о существенности влияния именно такого выбора будет сказано ниже).

3. Эксплуатант тратит 100% прибыли на расширение парка.

4. Установление ответственности для «мультиротора легкого» не представляет сложностей, например, в силу простоты конструкции и гипотетического наличия системы автоматического определения причины отказов, а потому в идеале беззатратно.

5. Начальный объем заказа – формально неограниченная величина. Поэтому в целях проведения анализа рассматривается два возможных варианта: двукратное и десятикратное увеличение начального объема заказа (как и в случае темпа обучения в процессе устранения КПН, о влиянии выбора данных значений будет сказано ниже).

Результаты проведенных расчетов изменения нормированной суммарно накопленной прибыли производителя и эксплуатанта $\Delta \bar{S}_{\Sigma, j}$ представлены на рис. 7–8.

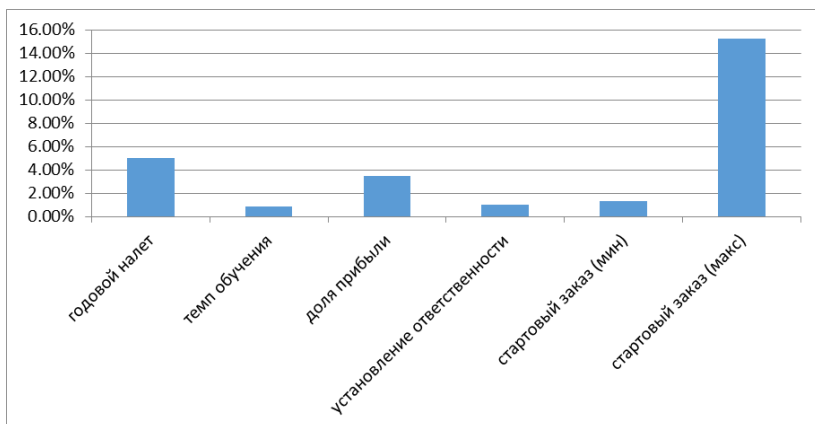


Рис. 7. Изменение нормированной суммарно накопленной прибыли производителя и эксплуатанта, «сценарий 1»

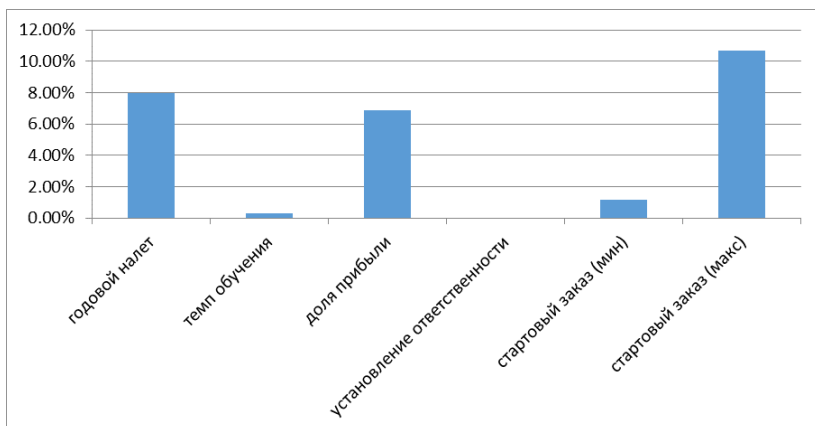


Рис. 8. Изменение нормированной суммарно накопленной прибыли производителя и эксплуатанта, «сценарий 2»

Анализ рис. 7–8 показывает, что варьируемые параметры можно разделить на две группы. К группе «слабо влияющие» относятся темп обучения и стоимость установления ответственности, к группе «сильно влияющие» – годовой налет и доля прибыли на покупку новых БАС. При этом:

- влияние стоимости установления ответственности рассчитано только для «сценария 1», поскольку в «сценарии 2» этот параметр – нулевая константа;

- влияние темпа обучения настолько мало, что для получения итоговых выводов выбор адекватного значения $a_{j\max}$ для этого параметра (к вопросу выбора вышеуказанного значения равным 0,25) несущественен.

Что касается объема стартового заказа, рассчитанного для двух возможных значений $a_{j\max}$, то его также можно отнести к группе «сильно влияющие» при условии, что $a_{j\max} \geq 3 \times a_{j\min}$ (что соответствовало бы столбцу с высотой, промежуточной между двумя правыми столбцами диаграмм рис. 7–8).

Такое разделение параметров на группы означает, что наибольшее влияние на суммарную прибыль оказывают те из них, которые способствуют максимально быстрому росту парка БАС и суммарной годовой выручки.

6. Стратегия управления производителя с целью максимизации его прибыли

В приведенных выше расчетах наибольший рост рассматриваемых отраслей происходит, если существует некий организационный механизм, при наличии которого показатели деятельности участников отрасли рассматриваются объединенно, а не обособленно для каждого участника. В практическом плане это могут быть варианты:

- государство выступает как надсистемный координатор развития отрасли, действующий во благо общего роста;

- производитель АТ берет на себя функции координатора и добровольно принимает на себя все издержки, связанные с развитием отрасли (в модели – это параметр «стоимость установления ответственности»).

По факту, конечно же, в условиях капиталистического общества производитель и эксплуатант действуют как самостоятельные субъекты, заинтересованные только в максимальном увеличении собственной прибыли. Как показывает мировая практика, отзыв продукции вследствие производственного бра-

ка может быть выгоден производителю [1, 6]. Отсюда возникает задача поиска производителем оптимальной стратегии управления: в каком случае ему выгодно действовать как координатору отрасли, а в каком – обособленно при справедливом распределении ответственности за отказы.

Суть искомой стратегии: если первоначально производитель действует как координатор, в каком году t^* ему следует изменить политику на самостоятельное развитие, чтобы его прибыль, суммарно накопленная за период моделирования, стала максимальной.

В рамках введенной модели развития отрасли были выполнены соответствующие расчеты (рис. 9–10) с учетом того, что, согласно анализу предыдущего раздела, в целях максимального роста отрасли объем стартового заказа был увеличен до 1000 БВС.

Закономерность изменения суммарно накопленной прибыли производителя (прибыли в понимании сокращения/роста убытков) может быть пояснена на примере рис. 9, $t^* = 8$ год: до года изменения стратегии $t < 9$ год кривая этой прибыли очевидным образом соответствует общему участку всех кривых, построенных для $t \geq 9$ год. В год изменения стратегии $t = t^*$ соответствует первому излому кривой (точка А): производитель, получив долю прибыли эксплуатанта как оплату за новую партию БАС, вследствие введения справедливой ответственности начинает работать «в прибыль», и последующий рост его прибыли становится «индивидуальным» (кривая более не совпадает с кривыми, соответствующими другим стратегиям). В следующий год $t = t^0 + 1$ происходит второй излом кривой (точка Б): эксплуатант, понеся дополнительные расходы, заказывает меньшее количество БАС, чем мог бы, если бы этих расходов не было (рис. 10).

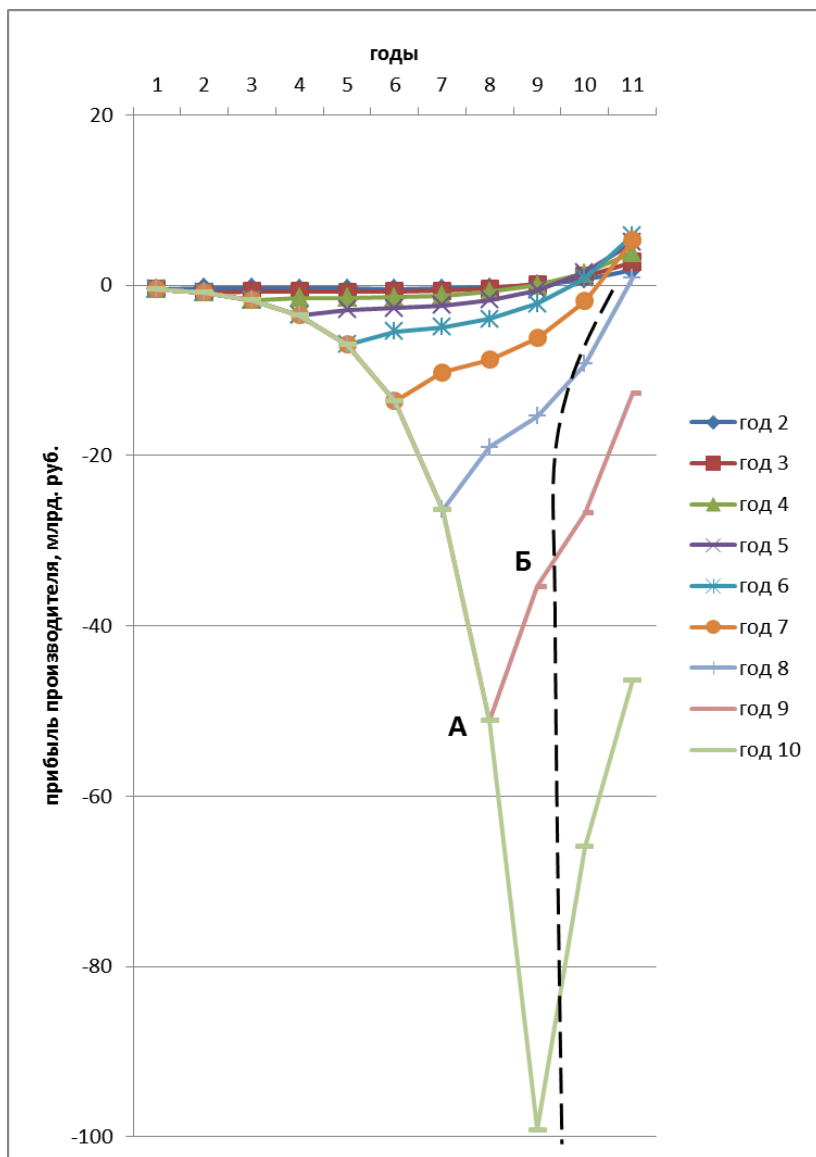


Рис. 9. Изменение суммарно накопленной прибыли производителя «мультироторов легких» для различных вариантов t^*

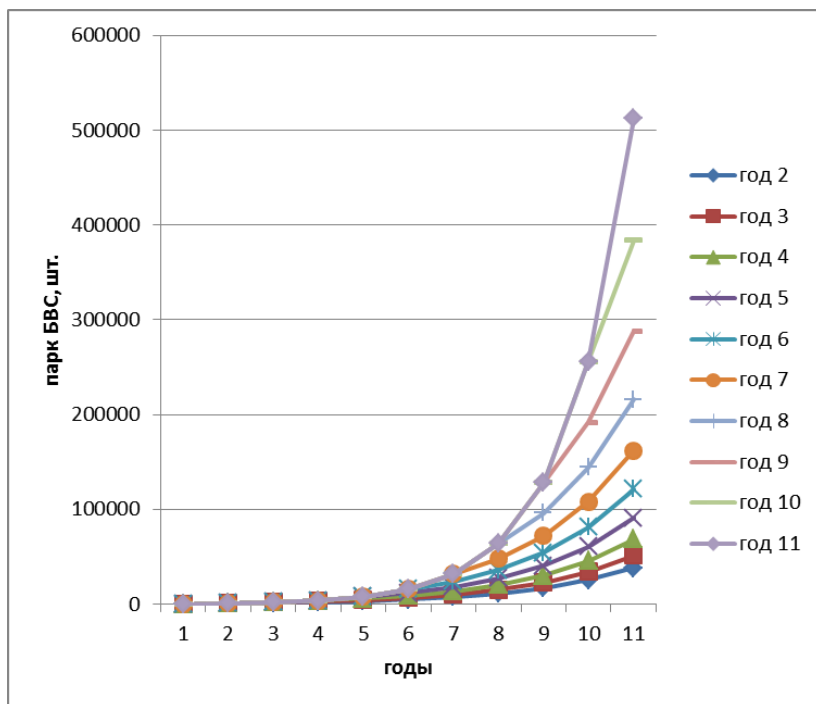


Рис. 10. Изменение парка «мультироторов легких» для различных вариантов t^*

При этом наблюдается следующая ситуация: чем позже производитель отказывается от априорного принятия ответственности, тем более интенсивен последующий рост его суммарной прибыли. И при гипотетическом моделировании неограниченного рыночного спроса на БАС (дальнейшее производство при $t > 10$ лет) наибольший рост прибыли был бы достигнут для кривой, соответствующей $t^* = 10$ год. В реальности при ограниченном спросе на БАС возможна ситуация, что, затягивая с изменением стратегии, производитель может не достигнуть максимально возможной суммарно накопленной прибыли. Например, если рынок БАС ограничен 200 тыс. экземпляров (условная пунктирная кривая рис. 9), то для $t^* = 9$ год и $t^* = 10$ год производитель не достигнет положительной прибыли

как таковой вообще (значения в области правее пунктирной кривой невозможны); максимальная же суммарно накопленная прибыль соответствует $t^* = 6$ год.

Моделирование развития отрасли с целью определения оптимальной стратегии производителя показало следующее. При долгосрочном производстве БАС ($t > 10$ лет, что, с учетом прогнозируемых масштабов рассматриваемой отрасли, соответствовало бы росту парка БВС до 1 млн штук и более) производителю выгодно изменить свою стратегию за 3 года до насыщения рынка БАС, т.е. до завершения массового производства БАС (рис. 11). Год изменения стратегии при этом должен соответствовать моменту, когда парк эксплуатируемых БВС составит ~10% от прогнозируемого спроса (рис. 12).

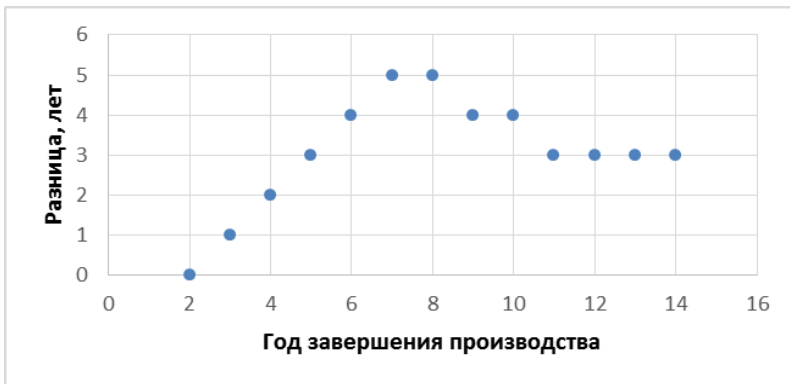


Рис. 11. Разница между годом завершения производства БАС и годом t^* изменения стратегии производителя в зависимости от года завершения производства БАС

Выявленный эффект по сути означает, что производителю выгодно максимально быстро преодолеть период обучения и устранить КПН, чтобы в дальнейшем произвести основную долю БАС (в рамках рассматриваемой отрасли – приходящуюся на последние 3 года) с высоким уровнем безотказности АТ и не нести при этом затраты, связанные с прочими отказами.



Рис. 12. Доля парка БАС, соответствующая году t^* изменения стратегии производителя, в зависимости от ограниченности рынка БАС

При этом, акцентируя внимание, следует отметить, что суммарно накопленная прибыль в отрасли существенно уменьшится, что дополнительно подчеркивает важность совместного координирования всех участников авиационной деятельности при внедрении новой АТ с возможным широким применением.

7. Выводы

На современном этапе развития авиации значительно выросли количество и уровень требований, предъявляемых к выполняемым авиацией работам и услугам. В первую очередь эти требования касаются общей стоимости эксплуатации изделий АТ и безопасности полетов. Указанные требования являются существенным препятствием при широкомасштабном внедрении новой АТ, особенно если это касается противоречия интересов нескольких участников отрасли.

Наиболее комплексный подход, позволяющий выработать в таких условиях необходимое управленческое решение, – это объединение в рамках одной оптимизационной задачи варьируемых параметров, относящиеся, в частности, к производителям АТ и ее эксплуатантам. Суть подхода означает централизован-

ное формирование единого оператора, управляющего развитием отрасли применения новой АТ. Альтернативой является традиционный сценарий развития, при котором участники отрасли применения авиации действуют как самостоятельные субъекты, преследующие исключительно собственные интересы.

Сравнение двух указанных сценариев развития отрасли показывает, что при формировании единого координатора отрасли и выработке соответствующей стратегии управления, при которой авиапроизводитель на начальном этапе внедрения новой АТ априорно принимает на себя ответственность за все ее отказы, суммарно накопленная прибыль в отрасли будет существенно больше, чем в отсутствие данной стратегии управления. Эффект достигается за счет того, что при априорном принятии ответственности из затрат отрасли исключаются «лишние» затраты, связанные с установлением ответственности за причину отказов. Это приводит к росту выручки и более быстрому накоплению опыта эксплуатации, что позволяет быстрее создавать и внедрять более безотказные изделия АТ. Основными факторами, способствующими ускорению накопления опыта, являются величина стартового заказа серийных изделий АТ (как наиболее значимый параметр), возможный налет/наработка изделий АТ в единицу и доля прибыли, выделяемой эксплуатантом на закупку новых изделий.

Возможными способами реализации подобной стратегии управления являются создание единого оператора отрасли (которым может быть, например, государство) или добровольное принятие на себя производителем соответствующих функций. В последнем случае производителем может быть поставлена задача максимизации собственной прибыли за счет выработки более гибкой стратегии: на начальном этапе эксплуатации новой АТ – брать всю ответственность на себя, на последующем этапе – только за отказы вследствие КРН. При этом, безусловно, суммарно накопленная прибыль в отрасли будет меньше, однако данная стратегия также позволит быстро преодолеть период «детских болезней» новой АТ, связанных с ее конструктивными недостатками, тем самым уменьшить количество отказов при переходе к широкомасштабной эксплуатации этой АТ.

Таким образом, априорное принятие производителем ответственности на себя позволяет вырабатывать более эффективные управленческие решения, направленные в том числе на ускорение роста перспективных рынков реализации как авиационных работ и услуг, так и продукции авиастроения. Централизованное формирование единого координатора развития отрасли становится целесообразным шагом при переходе именно к новым технологическим укладам как в авиации, так и в областях ее целевого применения.

Литература

1. АРЖАНОВА Я. *Массовый отзыв продукции: как обернуть ситуацию на пользу бренду* // Бизнес-журнал. – 2013. – №2. – С. 20-22.
2. ВАСИЛЬЕВ Ю.А. *Оборот некачественной продукции: причины и последствия*. – URL: <https://www.audit-it.ru/articles/account/court/a52/921772.html> (дата обращения: 12.12.2023).
3. ЕГЕР С.М., ЛИСЕЙЦЕВ Н.К., САМОЙЛОВИЧ О.С. *Основы автоматизированного проектирования самолетов*. – М.: Машиностроение, 1986. – 232 с.
4. ЕГЕР С.М., МИШИН В.Ф., ЛИСЕЙЦЕВ Н.К. *Проектирование самолетов. Издание третье, переработанное и дополненное* / Под. ред. С.М. Егера. – М.: Машиностроение, 1983. – 616 с.
5. КИЗИМ Н.А. *Организация крупномасштабных экономико-производственных систем: монография*. – Х.: Бизнес Информ, 2000. – 108 с.
6. КЛЕКОВКИН Н. *Цена ошибки. Как бизнесу сократить потери при отзыве продукции*. – URL: <https://www.forbes.ru/biznes/356177-cena-oshibki-kak-biznesu-sokratit-poteri-pri-otzyve-produkcii> (дата обращения: 15.12.2023).
7. КЛОЧКОВ В.В., ВАРЮХИНА Е.В. *Экономические аспекты определения ответственности производителей за отказы техники* // Вестник УГТУ-УПИ. Серия: экономика и управление. – 2010. – №2. – С. 141–151.

8. КЛОЧКОВ В.В., КРИТСКАЯ С.С. *Метод оценки стартовой скидки на новые типы гражданских воздушных судов* // Финансовая аналитика: проблемы и решения. – 2014. – №45(231). – С. 2–11.
9. Минтранс РФ установил ценовой предел для сертификации беспилотников до 30-70 млн руб. Сайт информационного агентства «Интерфакс». – URL: <https://www.interfax.ru/russia/898017/> (дата обращения: 12.12.2023).
10. МУХАМЕТЗЯНОВ К.М. *Инновационная товарная стратегия авиастроительного предприятия* // Наука, теория, практика авиационно-промышленного кластера современной России. Материалы V-й Международной научно-производственной конференции, приуроченной ко Дню Российской науки. Ульяновск, 2020. – Изд.: УГТУ (Ульяновск). – С. 219–223.
11. *О формировании параметров государственного гражданского заказа в части потребностей субъектов Российской Федерации в БАС и услугах. Презентация ГТЛК, февраль 2023.* – URL: <https://gtlk.ru/> (дата обращения: 15.12.2023)
12. ПЛАТОНОВА А.В. *Формирование инновационно-ориентированной интегрированной среды в авиастроении* // Вестник Южно-Уральского государственного университета. Серия: экономика и менеджмент. – 2021. – Т. 13, №3. – С. 132–138.
13. ПОПОВ Ю.В. *Количественные показатели безопасности полетов* // Проблемы безопасности полетов. – 2015. – №2. – С. 10–19.
14. РУХЛИНСКИЙ В.М., КУМИНОВА А.П. *Проблема безопасности полетов, связанная со снижением роли экипажа в пилотировании ВС, и переход к автоматизированным полетам* // Научный вестник ГосНИИ ГА. – 2018. – №22. – С. 91–101.
15. СТЕПАНОВ Е.А., ПЛЕТНЁВ Д.А. *Траектории экспортных стратегий предприятий российского транспортного машиностроения* // П-ECONOMY. – 2022. – Т. 15, №1. – С. 19–34.
16. ТИХОНОВ А.И. *Организационно-экономические механизмы выбора стратегии развития предприятия авиационной*

промышленности // Вестник академии знаний. – 2020. – №37(2). – С. 325–330.

17. BOUCHER P. *You wouldn't have your granny using them: Drawing boundaries between acceptable and unacceptable applications of civil drones* // Science and Engineering Ethics. – 2015. – No. 22(5). – P. 1391–1418.
18. CLOTHIER R.A., GREER D.A., GREER D.G. et al. *Risk Perception and the Public Acceptance of Drones* // Risk Analysis Journal. 2015. – URL: <https://doi.org/ezproxy.libproxy.db.erau.edu/10.1111/risa.12330> (дата обращения: 14.11.2023).

ORGANIZATIONAL MECHANIZM FOR THE INTRODUCTION OF AVIATION TECHNOLOGIES TO ENSURE AN ACCEPTABLE LEVEL OF FLIGHT SAFETY AND THE TOTAL COST OF OWNERSHIP OF NEW AIRCRAFT PRODUCTS

Vladislav Klochkov, V.A. Trapeznikov Institute of Control Sciences of RAS, Moscow, National Research Center “Zhukovskiy Institute”, Zhukovskiy, Doctor of Science (klochkovvv@nrczh.ru).

Sergey Yegoshin, National Research Center “Zhukovskiy Institute”, Zhukovskiy, (sergey4791@yandex.ru).

Abstract: In modern conditions, the study of large-scale systems can significantly improve the efficiency of industry management. The approach considered makes it possible, using the example of the aviation industry, to develop a justification for an organizational mechanism that differs from traditional business models of aviation application by the ability to achieve maximum economic effect, taking into account compliance with modern requirements for aviation operations and services in terms of flight safety. The mechanism is based on the a priori acceptance by the aircraft manufacturer of responsibility for all failures that occur during the operation of a new aircraft technology at the initial stage of its lifetime. The research methodology is based on mathematical modeling and methods of system analysis. As an example, the industry of promising application of unmanned aircraft systems "last mile delivery" is considered. The proposed approach makes it possible to carry out more effective development planning and development of management decisions when introducing new aviation equipment with the potential for wide application, including in the formation of state programs of the Russian Federation, strategies for the development of industrial organizations in the aircraft industry.

Keywords: unmanned aircraft systems, unmanned aerial vehicles, design and production shortcomings, delivery of the "last mile", management of industry development.

УДК 338.47, 338.534

ББК 39.54

DOI: 10.25728/ubs.2024.111.8

*Статья представлена к публикации
членом редакционной коллегии Р.М. Нижегородцевым.*

Поступила в редакцию 10.04.2024.

Опубликована 30.09.2024.

МЕТОД ПОИСКА РАЗРЕЗОВ ГРАФА ДЛЯ ЗАДАЧИ УПРАВЛЕНИЯ ИНЖЕНЕРНОЙ ИНФРАСТРУКТУРОЙ

Вандиловская П. А.¹, Крыгин А. А.², Лукинова О. В.³,
Рошин А. А.⁴

(ФГБУН Институт проблем управления
им. В.А. Трапезникова РАН, Москва)

Целью функционирования инженерных сетей является обеспечение поставок того или иного ресурса потребителю, при этом, в идеальном случае, подача должна быть непрерывной, что напрямую зависит от целостности инфраструктуры сети. Однако различные факторы: атаки злоумышленников, природные катаклизмы, наконец, естественные технологические причины (различные аварии), приводят к отключению некоторых участков сети, что влечет нарушение режима поставки ресурса. Тогда возникает задача поиска наиболее уязвимых (критических), с точки зрения возможного ущерба, участков инженерной сети. Ее решение позволяет принять соответствующие меры по защите сети от влияния негативных факторов и обеспечить максимально бесперебойную подачу ресурсов. Инженерную сеть принято моделировать графовыми структурами, поэтому одним из методов решения данной задачи является нахождение разрезов графа сети. Такие методы существуют, но все они обладают рядом ограничений. В данной работе предлагается новый метод нахождения всех разрезов графа инженерной сети, вообще говоря, произвольной размерности; описывается алгоритм метода, а также его теоретическое обоснование. Концепция метода основана на формировании на каждой итерации особых конструкций графа (мультиразрезов) таким образом, что в результате отработки алгоритма метода осуществляется поиск всех разрезов. Примерами инженерных сетей, где данный метод может быть использован в качестве одного из инструментов принятия рациональных решений при эксплуатации сетевых объектов, являются электросети, сети водоснабжения и канализации, а также сети связи и телекоммуникаций.

Ключевые слова: граф, разрез графа, мультиразрез, инженерная сеть, свободный путь графа.

¹ Полина Александровна Вандиловская, инженер (polinavandi@yandex.ru).

² Андрей Александрович Крыгин, к.т.н. (andreyakr14@gmail.com).

³ Ольга Васильевна Лукинова, д.т.н., доцент (lobars@mail.ru).

⁴ Александр Алексеевич Рошин, к.т.н., доцент (rochinaa@ipu.ru).

1. Введение

Под инженерными инфраструктурами (ИИ) понимается комплекс технических средств, устройств, сооружений, предназначенных для жизнеобеспечения зданий, включающий водо-, электро-, тепло-, газоснабжение, канализацию, слаботочные сети и т.п. Целью функционирования таких инфраструктур является обеспечение бесперебойной подачи того или иного ресурса от источников его производства потребителям, что напрямую зависит от целостности инфраструктуры сети. Однако различные факторы: атаки злоумышленников, природные катаклизмы, наконец, естественные технологические причины (различные аварии), приводят к отключению некоторых участков сети. Тогда возникает задача поиска уязвимых участков инженерной сети, выход из строя которых нарушит режим поставки ресурса.

Для формального описания ИИ будем использовать граф, включающий вершины-источники (производители ресурса), вершины-стоки (потребители ресурса), внутренние вершины (различные технические средства инженерной сети: распределители ресурса, подстанции, накопители и т.д.). Ребрам графа соответствуют протяженные участки сети, например, линии электропередач для электросетей, трубопроводы для тепловых сетей и т.п. При этом граф представляет собой модель сети со всеми возможными резервными линиями, даже если они не задействованы в нормальном режиме функционирования.

Очевидно, что наибольшее количество аварий и повреждений происходит именно на участках сетей, что нарушает режим поставки ресурса потребителям, при этом в самом худшем случае возникает отключение всех потребителей полностью, так как переконфигурировать сеть для возобновления снабжения не представляется возможным, поскольку в модель включаются все резервные линии. Такие ситуации являются критическими и именно они являются целью моделирования в данной задаче. Например, злоумышленник ставит целью отключить некоторый фрагмент железной дороги, для чего ему необходимо обесточить все тяговые подстанции данного фрагмента. Тогда необходимо выявить все линии электросети, которые он может атаковать.

На уровне модели (графа) поврежденные участки сети ассоциируются с разрезами графа, поэтому решение исходной задачи предполагает поиск всех разрезов графа, которые прерывают все возможные пути от источников к потребителям (изолируют источники от потребителей), а также не содержат «избыточные» ребра, т.е. ребра, не лежащие на путях, ведущих от источников к потребителям.

На сегодняшний день существует достаточное количество методов построения разрезов графа [1–16], которые используют в своей основе различные формализмы. В частности, в [3, 10] описаны методы, которые основаны на алгоритмах полного перебора. Такой подход гарантированно определяет необходимые объекты, однако время вычисления растет экспоненциально. В [7, 12, 13, 15] предлагается подход, основанный на генетических алгоритмах, который сокращает время вычислений, но, в отличие от комбинаторных алгоритмов, в силу случайности поиска не дает гарантий, что хотя бы один разрез будет найден. В [2] показан способ нахождения ограниченного количества разрезов: только одно-, двух- и трехэлементных разрезов графа, что существенно сужает область применения описанного алгоритма.

Наиболее известными способами нахождения разрезов графа являются алгоритмы Каргера – Штейна [11] и Штор – Вагнера [14, 16]. Хотя оба алгоритма имеют достаточно приемлемые характеристики эффективности, у них есть свои ограничения. Алгоритм Каргера – Штейна – это рандомизированный алгоритм с достаточно высокой, но не абсолютной вероятностью нахождения разреза графа. Алгоритм Штор – Вагнера – это детерминированный алгоритм, который выполняется за $O(n^3)$ времени. К тому же алгоритм Каргера – Штейна неприменим для решения задач, где имеется источник и приемник, а именно это условие является существенным для поиска уязвимых фрагментов инженерных сетей; также оба этих метода нацелены на нахождение всего лишь одного разреза, причем минимального.

Кроме того, исследования литературы выявили следующую особенность: многие источники предлагают способы поиска разрезов графа, но для решения задач в конкретных прикладных областях, и этот факт диктует те или иные ограничивающие условия на постановку, алгоритм поиска и результат [5, 6]. Например,

в [5] при логическом проектировании дискретных устройств ставится задача поиска максимального разреза взвешенного графа. Алгоритм решения задачи строится с учетом некоторой функции, оценивающей качество разделения (как правило, на основе суммарного минимального/максимального веса ребер разреза). Для решения рассматриваемой в данной статье задачи строить алгоритм с учетом оптимизации по весам ребер просто неактуально.

В данной работе представлен метод нахождения всех таких разрезов графа, моделирующего инженерную инфраструктуру, которые изолируют источники ресурса от потребителей при наступлении аварийной ситуации. Метод основан на конструировании разрезов таким образом, что на каждом шаге генерируется новый набор ребер, который всегда является разрезом. В этом заключается существенное отличие от полного перебора, характеризующегося тем, что найденный набор является разрезом далеко не всегда, за счет чего, собственно, и возникает необходимость в полном переборе, ведущая к росту времени вычислений.

2. Описание метода

Инженерная сеть традиционно моделируется в виде ориентированного графа $G = \{V, E\}$, где V – множество вершин графа, E – множество ребер графа, $|V| = n$, $|E| = m$.

В литературе элементы множества E ориентированного графа называются ребрами, или дугами [18], мы будем пользоваться термином «ребро».

В решаемой задаче вершины графа соответствуют точечным объектам сети (подстанциям, потребителям, электростанциям и т.п.), а ребра – протяженным (линиям электропередач, трубопроводам и т.п.). Ориентация ребра выбирается согласно направлению передачи ресурса.

Обозначим ребро $e_E = (v_i, v_j)$, где $v_i, v_j \in V$ (здесь и далее у символа « e_E » нижний индекс в виде заглавной буквы обозначает принадлежность объекта, в частности, ребра, к тому или иному множеству).

Особенностью данной задачи является тот факт, что множество вершин V описывается тройкой $\{S, T, U\}$, где S – множество

источников, T – множество приемников и U – множество остальных вершин: $V = S \cup T \cup U$. Ребра, инцидентные источникам, ориентированы от вершин-источников, а ребра, инцидентные приемникам, – к вершинам-приемникам. Сделаем также допущение о том, что не существует ребер, соединяющих пары источников или пары потребителей.

Поскольку граф G имеет несколько источников $S = (s_1, s_2, \dots, s_r)$ и стоков (приемников) $T = (t_1, t_2, \dots, t_l)$, построим их фиктивными вершинами S^* и T^* (следует заметить, что введение фиктивных вершин является важным фактом с точки зрения алгоритмической реализации предлагаемого метода, который начинает свою работу именно с фиктивной начальной вершины S^* и заканчивает фиктивной конечной вершиной T^*). Также будем считать, что граф G связан.

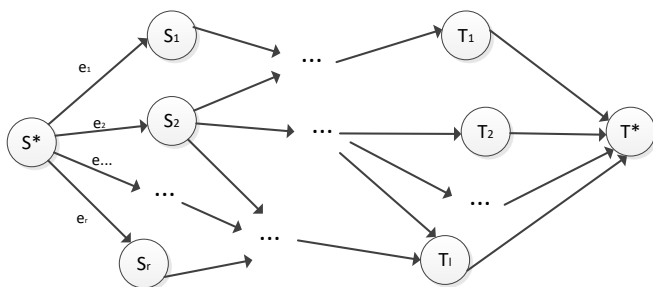


Рис. 1. Исходный граф $G = \{V, E\}$ с фиктивными вершинами S^*, T^*

Определение 1. Под разрезом связного графа с вершинами $V = V_1 \cup V_2$, где V_1 и V_2 – два непересекающихся подмножества V , принято понимать множество P всех тех ребер, которые имеют одну вершину в V_1 , а другую – в V_2 [6].

Оrientированный разрез в ориентированном графе – это разрез $\langle V_1, V_2 \rangle$, все ребра которого или исходят из V_1 , или заходят в V_2 [6].

Специфика данной задачи требует, чтобы определение разреза учитывало следующее условие.

Условие 1. V_1 должно обязательно содержать все источники S , а V_2 – все стоки T . Поэтому задача заключается в нахождении только таких разрезов графа G , что $S \subseteq V_1$ и $T \subseteq V_2$ (в дальнейшем под термином «все разрезы графа» понимается подмножество всех разрезов, обладающих указанным свойством).

Введем понятие «мультиразрез».

Определение 2. Мультиразрез – это подмножество ребер, содержащих один или несколько разрезов.

Для дальнейших рассуждений удобнее пользоваться определениями разреза и мультиразреза, сформулированными через понятие «путь».

Определение 3. Ориентированным маршрутом в ориентированном графе называется конечная последовательность вершин $v_0, v_1, \dots, v_{k-1}, v_k$ такая, что (v_{i-1}, v_i) , $1 \leq i \leq k$, является ребром графа. При этом вершины v_0, v_k называются концевыми [6].

Определение 4. Ориентированный маршрут называется ориентированной цепью, если все его ребра различны [6].

Определение 5. Открытая ориентированная цепь называется ориентированным путем, если различны все ее вершины [6].

Определение 6. Мультиразрез – такое подмножество ребер, что любой путь из S^* в T^* содержит хотя бы одно ребро из этого подмножества.

Определение 7. Разрез – это мультиразрез, не содержащий в себе другого мультиразреза.

Справедливо следующее утверждение, которое позволит в дальнейшем рассуждать в терминах путей в графе.

Утверждение 1. Определения 1 и 7 эквивалентны.

Доказательство. Пусть имеется разрез P , разделяющий вершины на два подмножества V_1 и V_2 , $S^* \in V_1, T^* \in V_2$. Покажем, что 1) любой путь из S^* в T^* содержит хотя бы одно ребро из P , и 2) не существует подмножества $\tilde{P}$, для которого выполняется первый пункт и такого, что $\tilde{P} \subset P$. Предположим, что существует путь из s^* в t^* , не содержащий ребер из P . Так как S^* и T^* принадлежат разным подмножествам, то на этом пути встретится ребро (v_{i-1}, v_i) такое, что $v_{i-1} \in V_1, v_i \in V_2$, но такое ребро должно принадлежать P . Пусть подмножество $\tilde{P}$ существует и ребро (v_{i-1}, v_i) принадлежит P , но не принадлежит $\tilde{P}$. Так как

граф связный, то существуют пути $s^* \rightarrow v_{i-1}$ и $v_i \rightarrow t^*$, и тогда для $\tilde{P}$ не выполняется первый пункт.

Пусть имеется множество ребер P такое, что любой путь из S^* в T^* содержит хотя бы одно ребро из P и любое подмножество $\tilde{P} \subset P$ этим свойством не обладает. Покажем, что P – разрез. Очевидно, что P – это несокращаемое множество ребер, при удалении которого из графа граф $\tilde{G} = G(V, E \setminus P)$ перестает быть связным, так как вершины S^* и T^* относятся к разным компонентам связности. То есть по определению P – разрезающее множество. Если бы компонент связности было больше двух, то была бы возможность построить подмножество $\tilde{P}$, исключив из P ребро, связывающее компоненту, содержащую s^* и компоненту, не содержащую ни S^* , ни T^* . По теореме о связи разрезающего множества с разрезом [6], P – разрез.

Утверждение доказано.

Постановка задачи. Найти все разрезы исходного графа G , удовлетворяющие условию 1.

Будем обозначать полное множество таких разрезов D .

Рассматриваемые далее подмножества ребер будем обозначать Q (с индексом или без), если подмножество является мультиразрезом, и P , если оно является разрезом.

Для решения исходной задачи в работе предлагается метод, позволяющий конструировать множество D . Ниже описана его идея.

Рассмотрим следующую операцию, которую, по аналогии с методом поиска в ширину, назовем *операцией ветвления*. Пусть имеются некоторый разрез P и ребро этого разреза e_0 с конечной вершиной v_0 . Обозначим подмножество ребер, для которых вершина v_0 является начальной N . Определим подмножество ребер Q следующим образом: $Q = N \cup P \setminus e_0$.

Ниже (в теореме п.3) будет показано, что возможны два варианта: полученное после операции ветвления над разрезом подмножество ребер Q обязательно будет являться: а) либо мультиразрезом, содержащим единственный разрез, б) либо собственно разрезом. Несложно показать, что мультиразрез Q является разрезом тогда и только тогда, когда $\forall e_k \in Q$ подмножество ребер $Q \setminus e_k$ не является разрезом. Соответственно (по утверждению 1),

если Q – разрез, то для любого подмножества $Q \setminus e_k$ найдется путь из S^* в T^* , не содержащий ребер этого подмножества.

Ниже представлена *процедура анализа* мультиразреза Q , в ходе которой определяется тот или иной вариант:

1. Из множества ребер E исключаются элементы, входящие в мультиразрез Q .

2. Фиксируется ребро $e_k \in Q$, где $k = 1, 2, \dots, |Q|$; оно возвращается в граф, после чего проверяется достижимость вершины T^* из S^* с учетом данного ребра (проверка на достижимость осуществляется с помощью любого стандартного метода, например, обхода в ширину).

3. Если из S^* в T^* не образуется пути(-ей), то это означает, что найденная конструкция сама по себе не является разрезом, но содержит его внутри. Для его нахождения ребро e_k исключается из Q , так как является избыточным. Назовем такое действие *операцией очистки*.

4. Если из S^* в T^* пути образовались, ребро не является избыточным и принадлежит к разрезу.

5. Пункты 2–4 повторяются для всех ребер мультиразреза Q .

После отработки процедуры анализа и операции очистки окончательный набор ребер устраняет достижимость T^* из S^* и не содержит избыточных ребер, т.е. является разрезом. Таким образом, если в процессе выполнения процедуры анализа хотя бы один раз была выполнена операция очистки, то имеем вариант 1, в противном случае – вариант 2.

Для выявления всех разрезов операции ветвления и очистки проводятся для каждого ребра текущего разреза, т.е. количество новых разрезов (без учета повторов) будет равно количеству ребер в разрезе P .

Для пояснения процедуры анализа рассмотрим пример графа на рис. 2.

Здесь вершиной-источником S^* является вершина под номером 0, а вершиной-приемником T^* – 6. Рассмотрим разрезы $P_1 = \{e_1, e_6\}$ и $P_2 = \{e_3, e_6\}$ и применим к каждому из них операцию ветвления по отношению к ребру e_6 . В результате получим множества $Q_1 = \{e_1, e_7\}$ и $Q_2 = \{e_3, e_7\}$ (в данном случае индексы 1 и 2 служат для понимания того, что перед нами разные наборы

ребер Q , которые были сконструированы от соответствующих разрезов P_1 и P_2). Далее подвергнем их операции очистки, вследствие которой получим разрезы $P_3 = \{e_7\}$, $P_4 = \{e_7\}$, являющиеся дубликатами по отношению друг к другу. Очевидно, что все последующие преобразования P_3 будут совершенно справедливыми и для P_4 , поэтому будет достаточно продолжить работу только с одним из них.

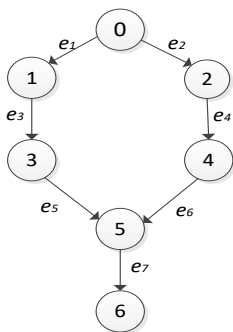


Рис.2. Ситуация дублирования разреза

Отсюда можно сформулировать правило проверки на дубликаты: если новый разрез P_j совпадает с уже содержащимся во множестве искомых разрезов D разрезом P_i ($P_j = P_i$ ($i \neq j$)), то P_j является дубликатом и далее не рассматривается; если дубликата не нашлось, то P_j добавляется в D .

Алгоритм представленного в работе метода реализуется двумя этапами. Первый этап, начальный, предназначен для обработки мультиразреза, построенного на ребрах, выходящих из фиктивной вершины S^* . Второй – основной этап работает для всех остальных, «внутренних» мультиразрезов.

Начальный этап:

Шаг 1. Из вершины S^* конструируется начальный набор ребер Q^* , который, очевидно, является мультиразрезом. Например, для графа, изображенного на рис. 1, начальным набором будет являться $Q^* = \{e_1, e_2, \dots, e_r\}$.

Шаг 2. Осуществляется процедура анализа, т.е. определяется, что найденный мультиразрез Q^* является разрезом ($Q^* = P_1$) или

содержит таковой ($Q^* \supset P_1$). Во втором случае для Q^* проводится операция очистки.

Шаг 3. Найденный P_1 включается во множество D .

Основной этап:

Шаг 1. Для каждого из ребер очередного добавленного в D разреза P_j с помощью операции ветвления строится новый мультиразрез Q .

Шаг 2. Для сконструированного Q выполняется процедура анализа. Выявляется разрез P_{j+1} .

Шаг 3. Для P_{j+1} осуществляется проверка на дубликат. Если P_{j+1} уникален, то он добавляется во множество D и для него начинают выполняться шаги 1–3.

Рассмотрим работу метода на примере графа, изображенного на рис. 3. Для удобства вершины обозначены номерами.

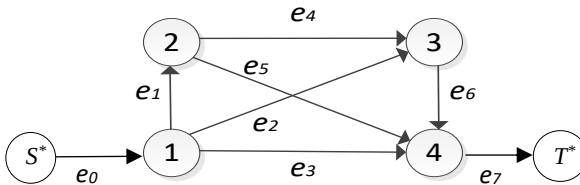


Рис. 3. Граф для демонстрации работы метода

Согласно шагу 1 начального этапа конструируем начальный набор ребер $Q^* = \{e_0\}$. На шаге 2 осуществляется процедура анализа, которая показывает, что набор Q^* является разрезом, т.е. $Q^* = P_1$, который на шаге 3 включается в искомое множество D .

Переходим к основному этапу. Так как P_1 включает одно ребро, то должен быть создан один новый мультиразрез. Применим операцию ветвления для e_0 , получим $Q = \{e_1, e_2, e_3\}$. На шаге 2 совершаем процедуру анализа, которая показывает, что Q является разрезом $P_2 = \{e_1, e_2, e_3\}$. На шаге 3 осуществляется проверка на дубликат, которая разрешает включить P_2 в D .

Так как P_2 включает три ребра, то должно быть создано три новых мультиразреза. Применим операцию ветвления для e_1 , по-

лучим $Q = \{e_4, e_5, e_2, e_3\}$. На шаге 2 совершаем процедуру анализа, которая показывает, что Q является разрезом $P_3 = \{e_4, e_5, e_2, e_3\}$. На шаге 3 осуществляется проверка на дубликат, которая разрешает включить P_3 в D .

Для найденного P_3 сконструируем новые мультиразрезы (к мультиразрезам из P_2 вернемся, когда будут найдены все разрезы в результате операции ветвления ребра e_1). Применим операцию ветвления к ребру e_4 и получим набор ребер $Q = \{e_6, e_5, e_2, e_3\}$. Применим процедуру анализа: ребро e_2 является избыточным, исключаем его и в результате очистки получаем набор ребер $P_4 = \{e_6, e_5, e_3\}$, который становится разрезом и после проверки на дубликат включается в D .

В найденном разрезе P_4 три ребра, поэтому должно быть сконструировано три новых мультиразреза. Применим операцию ветвления к ребру e_6 и получим набор $Q = \{e_7\}$, который является разрезом P_5 и после проверки на дубликат включается в D . Поскольку ребро e_7 входит в приемник T^* , дальнейшая работа с P_5 заканчивается.

Дальнейшая работа по конструированию разрезов проводится аналогичным образом.

Таким образом, результатом работы метода для графа, изображенного на рис. 3, будет следующее множество разрезов: $D = \{P_1 = \{e_0\}, P_2 = \{e_1, e_2, e_3\}, P_3 = \{e_4, e_5, e_2, e_3\}, P_4 = \{e_5, e_6, e_3\}, P_5 = \{e_7\}, P_6 = \{e_1, e_6, e_3\}\}$.

Как правило, в литературе оценка вычислительной сложности алгоритма на графах зависит от количества вершин (n) и ребер (m) графа. Однако для данной задачи следует учитывать характеристики мультиразрезов: зафиксируем некоторое ребро, тогда мультиразрез, содержащий ребро (и, в частности, зафиксированное ребро), будет проходить процедуру анализа $p_1 \cdot p_2 \cdot \dots \cdot p_k$ раз, где k – количество ребер в мультиразрезе, а p_i – количество путей от вершин источников к i -му ребру мультиразреза. Перейдем к усредненным величинам; пусть p – среднее количество путей от вершин источников к ребру, k – среднее количество ребер в мультиразрезе, полученном при ветвлении разреза, r – среднее количество таких мультиразрезов, в которых присутствует одно и то же ребро. Тогда количество операций проверки для одного

ребра можно оценить как $r \cdot p^k$. При этом не удалось найти работ, посвященных оценке величин p, k, r через n и m (а также максимальных значений количества путей, количества ребер в мультиразрезе и количества мультиразрезов, содержащих одно и то же ребро). Таким образом, теоретическую оценку вычислительной сложности провести можно, но она будет неинформативной.

3. Доказательство корректности метода

Вопрос теоретического обоснования корректности предложенного метода заключается в доказательстве следующих фактов:

1. Каждый новый мультиразрез Q , полученный с помощью алгоритма метода, содержит только один разрез, т.е. в Q не содержится других разрезов, не выявляемых процедурой анализа.

2. Множество D является полным.

Обоснование этих фактов следует из приведенной ниже теоремы о свободных путях, а также следствия из нее. Чтобы сформулировать теорему, введем следующее определение.

Определение 8. Пусть заданы некоторый мультиразрез и ребро этого мультиразреза. Пути, проходящие от источников до потребителей и содержащие это ребро, будем называть *свободными* (для заданного мультиразреза и ребра), если они не содержат других ребер мультиразреза.

Интерпретация данного определения заключается в следующем. Очевидно, что у каждого ребра e_P разреза P имеется хотя бы один свободный путь. В противном случае (когда на любом пути, проходящем через e_P , также имеется другое ребро разреза P) подмножество $P \setminus e_P$ является мультиразрезом, т.е. P – не разрез. Если мультиразрез Q не является разрезом, то любое ребро e_Q , имеющее в Q свободный путь, будет входить в любой разрез $P \subset Q$, так как в противном случае P не будет являться разрезом.

Теорема (о свободных путях). Пусть R_{old} – разрез, над которым проводится ветвление; e_0 – ребро, из которого производилось ветвление; v_0 – конечная вершина, инцидентная e_0 , из которой осуществляется ветвление; N – множество добавляемых

ребер; $R_{new} = N \cup R_{old} \setminus e_0$ – множество ребер, получившееся в результате операции ветвления; $\tilde{R} \subset R_{new}$ – произвольный разрез, содержащийся в множестве R_{new} .

Тогда набор ребер R_{new} , построенный из разреза R_{old} с помощью операции ветвления над ребром e_0 , является мультиразрезом и содержит в себе единственный разрез.

Доказательство. Логика доказательства теоремы основана на том, что будут обоснованы следующие тезисы:

1. Любая конструкция, полученная с помощью операции ветвления, является мультиразрезом Q .
2. Мультиразрез Q содержит хотя бы один разрез P .
3. Разрез $P \in Q$ является единственным.

Доказательство *первого тезиса* заключается в том, что любой путь, проходящий через e_0 , также проходит через одно из ребер N . Любой путь, не проходящий через e_0 , также проходит через одно из ребер $R_{old} \setminus e_0$, поэтому R_{new} – мультиразрез.

Второй тезис, т.е. тот факт, что любой мультиразрез Q содержит в себе хотя бы один разрез, следует из процедуры анализа (см. п. 2).

Доказательство единственности существования.

Логика доказательства единственности, т.е. того, что в мультиразрезе R_{new} содержится единственный разрез, следующая. Будет показано, что R_{new} однозначно разбивается на непересекающиеся подмножества ребер. И для каждого подмножества будет показано, что либо все его ребра должны принадлежать любому разрезу, содержащемуся в R_{new} , либо ни одно ребро подмножества не принадлежит ни одному разрезу. Из такой однозначности будет следовать, что разрез единственный.

Ниже будет показано, что $N \cap R_{old} \setminus e_0 = \emptyset$, так как R_{old} – разрез. Поэтому любое ребро из R_{new} принадлежит либо N , либо $R_{old} \setminus e_0$. Для разреза R_{old} у ребра e_0 есть свободный путь. По построению этот путь проходит через ребро из N (обозначим его e_{n1}). Тогда в мультиразрезе R_{new} у ребра $e_{n1} \in N$ есть свободный путь. Поэтому в любом разрезе $\tilde{R} \subset R_{new}$, присутствует e_{n1} . В общем случае в мультиразрезе R_{old} у ребра e_0 есть несколько свободных путей; обозначим $E_N^+ \subset N$ – подмножество ребер N , через

которое проходит хотя бы один свободный путь для мультиразреза R_{old} и ребра e_0 . Как было показано, все подмножество E_N^+ входит в любой разрез $\tilde{R} \subset R_{new}$ и это подмножество не пусто. Обозначим $E_N^- = N \setminus E_N^+$. Также из существования свободного пути для разреза R_{old} у ребра e_0 следует, что существует путь от вершины источника до e_0 , не содержащий ребер из R_{old} .

Разделим множество ребер $R_{old} \setminus e_0$ на два подмножества E_{old}^+ и E_{old}^- . Ребро $e_{old} \in E_{old}^+$, если в разрезе R_{old} у e_{old} есть свободный путь, не проходящий через вершину v_0 , в противном случае $e_{old} \in E_{old}^-$. Для $e_{old} \in E_{old}^+$ указанный свободный путь также будет свободным и в мультиразрезе R_{new} . И, следовательно, $E_{old}^+ \in \tilde{R}$ для любого разреза $\tilde{R} \subset R_{new}$.

Будем называть ребро $e_{old} \in E_{old}^-$ *лежащим выше* v_0 , если для мультиразреза R_{old} и ребра e_{old} существует свободный путь следующей последовательности вершин и ребер: $s \in S \rightarrow \dots \rightarrow v_0 \dots \rightarrow e_{old} \rightarrow \dots \rightarrow t \in T$, в противном случае ребро $e_{old} \in E_{old}^-$ *лежит ниже* v_0 . Покажем, что все ребра $e_{old} \in E_{old}^-$ лежат ниже v_0 . Для этого рассмотрим рис. 4.

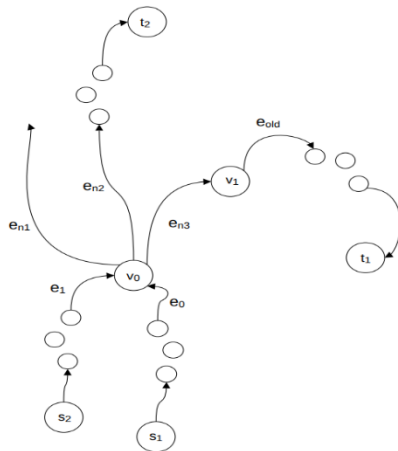


Рис. 4. Пример графа №1

В разрез R_{old} входят ребра e_{old} и e_0 , при этом у каждого из них есть свободные пути, проходящие через v_0 . Свободный путь

для e_{old} состоит из $(s_2 \rightarrow \dots e_1 \rightarrow v_0 \rightarrow e_{n3} \rightarrow v_1 \rightarrow e_{old} \rightarrow \dots \rightarrow t_1)$, свободный путь для e_0 состоит из $(s_1 \rightarrow \dots e_0 \rightarrow v_0 \rightarrow e_{n2} \rightarrow \dots \rightarrow t_2)$. На рисунке показана ситуация, когда e_{old} лежит выше v_0 и при этом есть наличие свободных путей. Тогда путь $(s_2 \rightarrow \dots e_1 \rightarrow v_0 \rightarrow e_{n2} \rightarrow \dots \rightarrow t_2)$ вообще не проходит через ребра разреза, то есть R_{old} – не разрез и не мультиразрез.

Вернемся к множеству E_N^- – подмножеству N , через которое не проходят свободные пути для разреза R_{old} и ребра e_0 . Выше было показано, что существует путь от вершины источника до e_0 , не содержащий ребер из R_{old} , обозначим его w_0 . Рассмотрим путь, состоящий из w_0 , e_0 , $e_n^- \in E_N^-$ и дальнейшего отрезка пути w_1 до вершины потребителя. Покажем, что на пути w_1 лежит ребро из E_{old}^+ . Для разреза R_{old} и ребра e_0 этот путь не свободен, поэтому на пути w_1 лежит ребро (или несколько ребер) $e_* \in R_{old}$. Предположим (от противного), что $e_* \in E_{old}^-$; если на этом пути лежит несколько ребер из E_{old}^- , то будем рассматривать последнее ребро e_* , после которого на пути w_1 нет ребер из R_{old} (рис. 5).

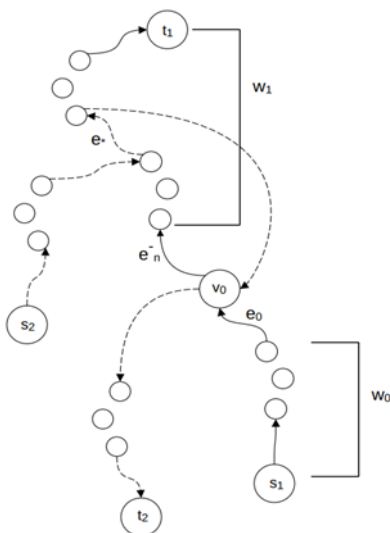


Рис. 5. Пример графа №2

На рисунке сплошной линией показан рассматриваемый путь. У ребра e_* в разрезе R_{old} есть свободный путь и, как было показано, e_* лежит ниже v_0 ; этот путь изображен пунктирной линией. Тогда у e_* есть свободный путь, не содержащий v_0 : $s_2 \rightarrow \dots e_* \rightarrow \dots \rightarrow t_1$, поэтому $e_* \in E_{old}^+$ – противоречие.

Соответственно, для мультиразреза R_{new} и любого ребра $e_N^- \in E_N^-$ любой путь, проходящий через e_N^- , также проходит через ребро $e_* \in E_{old}^+$. Все подмножество E_{old}^+ включено в каждый разрез $\tilde{R} \subset R_{new}$ и поэтому e_N^- (и все подмножество E_N^-) не включено в $\tilde{R}$, так как у него не будет свободных путей в $\tilde{R}$.

Аналогично и для любого ребра $e_{old} \in E_{old}^-$: все свободные в R_{old} пути, проходящие через e_{old} , подходят снизу к v_0 и далее проходят либо через ребро из E_N^- и ребро из E_{old}^+ , либо через ребро из E_N^+ , т.е. E_{old}^- не включено ни в один разрез $\tilde{R} \subset R_{new}$.

Таким образом, мультиразрез R_{new} однозначно разбивается на 4 подмножества: E_{old}^- , E_{old}^+ , E_N^- , E_N^+ ; подмножества E_{old}^- , E_N^- не входят ни в один разрез $\tilde{R} \subset R_{new}$, подмножества E_{old}^+ , E_N^+ целиком входят в каждый $\tilde{R} \subset R_{new}$, откуда следует единственность разреза $\tilde{R} \subset R_{new}$.

Теорема доказана.

Следствие. Множество D , формируемое методом, содержит полное множество разрезов.

Доказательство. Процедура анализа и проверки на дубликаты позволяют сделать вывод о том, что каждый элемент множества D является уникальным разрезом. Покажем от противного, что D является полным множеством, предположив, что существует разрез $P \notin D$. Такая ситуация может возникнуть в двух случаях:

1. Когда операция проверки выполнялась над мультиразрезом, содержащим несколько разрезов, среди которых присутствует P .

2. Когда P не содержался ни в одном из мультиразрезов, над которыми проводилась операция проверки.

По теореме о свободных путях первого случая быть не может, так как операция ветвления всегда выполнялась над разрезами. Покажем, что и второй случай невозможен.

Так как P – разрез, то у каждого его ребра есть хотя бы один свободный путь. Будем рассматривать лес, состоящий из отрезков свободных путей от вершин источников до ребер P . Если свободные пути двух ребер пересекаются, то отбросим один из участков от вершины источника до первой (считая от мультиразреза) вершины пересечения (рис. 6).

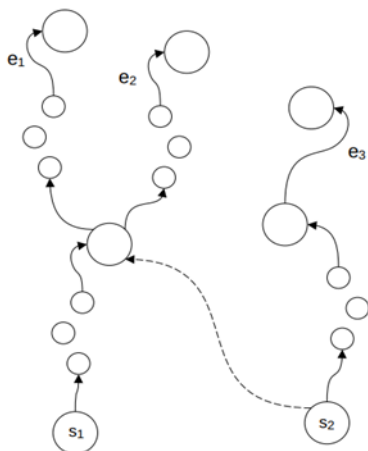


Рис. 6. Объединение свободных путей

На рис. 6 $e_1, e_2, e_3 \in P$, а пунктирной линией показан отброшенный участок.

Ввиду полноты перебора различных наборов операций {ветвления – проверки} для достижения противоречия достаточно показать, что существует последовательность этих операций, приводящая к разрезу P .

Эту последовательность будем строить следующим образом: проводить ветвление только тех ребер, которые принадлежат построенному лесу. Если после операций ветвления и проверки ребро разреза не будет принадлежать лесу, то дальнейшее ветвление из него не проводится. Дальнейшее ветвление также не проводится и из ребра, принадлежащего P . Поэтому (так как P – разрез) указанная последовательность операций приведет к P ,

и противоречие не возникнет только в том случае, когда в ходе операции проверки будет удалено ребро из леса.

Рассмотрим любой мультиразрез после операции ветвления разреза, полученный в описанной последовательности. Так как лес состоит из свободных путей, которые должны быть перекрыты мультиразрезом и по построению на этих путях есть только одно ребро мультиразреза, то никакое ребро мультиразреза, принадлежащее лесу, не может быть удалено в ходе операции проверки, так как оно принадлежит свободному пути.

Следствие доказано.

4. Заключение

В статье представлен метод поиска всех разрезов ориентированного графа с учетом специфики предметной области. Концепция метода основана на алгоритме поиска разрезов таким образом, что на каждом шаге генерируется новый набор ребер (мультиразрез), который всегда либо является разрезом, либо содержит единственный разрез. Описан алгоритм метода, сформулированы правила, положенные в его основу, представлена теорема, обосновывающая положения метода, а также следствие, доказывающее тот факт, что найденное множество разрезов полное.

Выявление таких компонент графа позволяет принимать более рациональные решения при управлении инженерными инфраструктурами, позволяющие свети к минимуму количество отключений потребителей от ресурсов при наступлении деструктивных факторов.

Вопросы вычислительной сложности алгоритма метода, в силу невозможности показать теоретические оценки, авторы планируют исследовать экспериментально на различных вычислителях и представить в будущих работах. Хотя следует заметить, что исходный граф, а также любой найденный разрез состоят из конечного числа ребер, кроме того, метод включает проверку на дубликаты, благодаря которой исключаются циклические операции. Эти факты говорят о том, что предлагаемый метод гарантирует обход графа за конечное время.

Литература

1. ГРИШКЕВИЧ А.А., РІАТЕК Л., БУРМУТАЕВ А. *Нахождение одно-, двух- и трехэлементных разрезов графа* // Вестник ЮрГУ, серия «Математическое моделирование и программирование». – 2008. – №15(115). – Вып. 1. – С. 12–22.
2. ДОРРИ М.Х., РОЩИН А.А., СЕРЕДА Л.А. *Применение программного комплекса РДС для расчетов и визуализации последствий выхода из строя инженерных сооружений* // Автоматизация в промышленности. – 2017. – № 11. – С. 11–14.
3. РЯБИНИН И.А. *Надежность и безопасность структурно-сложных систем*. – СПб.: Изд-во СПб. гос. ун-та, 2007. – 276 с.
4. ПОТТОСИН Ю.В., ПОТТОСИНА С.А. *Поиск разреза графа в решении некоторых задач логического проектирования* // Vesci Nacyânal'naj akademiï navuk Belarusi. Seryâ fizika-matematychnyh navuk. – 2016. – №3. – С. 111–118.
5. ПИРОВА А.Ю. *Параллельные алгоритмы разделения графов: учебное пособие*. – Нижний Новгород: Нижегородский госуниверситет, 2019. – 20 с.
6. СВАМИ М., ТХУЛАСИРАМАН К. *Графы, сети и алгоритмы*. – М.: Мир, 1984. – 454 с.
7. AGUDELO L., MUNOZ N., LÓPEZ-LEZAMA J.M. *Vulnerability assessment of power systems to intentional attacks using a specialized genetic algorithm* // Dyna (Medellin, Colombia). – 2015. – Vol. 82, Iss. 192. – P. 78–84.
8. CAGNO E., GRANDE O., TRUCCO P. *Towards an integrated vulnerability and resilience analysis for underground infrastructures* // Reliability Engineering & System Safety. – 2011. – Vol. 96, Iss. 1. – P. 139–148.
9. GREBENYUK G.G., NIKISHOV S.M. *Blocking of Energy and Resource Supply of Target Objects in Network Infrastructures* // Automation and Remote Control. – 2018. – Vol. 79(3). – P. 535–544.
10. HAENNI R. *Generating Diagnoses from Conflict Sets* // Proc. of the 11th Int. Conf. FLAIRS. – 1998 – URL: www.aaai.org/Papers/FLAIRS/1998/FLAIRS98-081.pdf (дата обращения: 23.03.2023).

11. KARGER D.R. *Global Min-cuts in RNC, and Other Ramifications of a Simple Min-Cut Algorithm* // SODA: Journal. – 1993. – Vol. 93. – P. 21–30.
12. KARIMI E., MADANI S.M., EBRAHIMI A. *Power transmission system vulnerability assessment using genetic algorithm* // Intelligent Systems in Electrical Engineering Fall. – 2012. – Vol. 3, No. 3. – P. 1–10.
13. KIM T., WRIGHT S.J., BIENSTOCK D. et al. *Vulnerability Analysis of Power Systems* // IEEE Trans. on Network Science and Engineering. – 2016. – Vol. 3, Iss. 3. – P. 132–146.
14. MAY R.P. *Genetic Algorithms for Agent-Based Infrastructure Interdependency Modeling and Analysis*. – URL: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.455.354&rep=rep1&type=pdf> (дата обращения: 23.03.2023).
15. ROSELYNA J.P., DEVARAJB D., DASH S.S. *Multi-Objective Genetic Algorithm for voltage stability enhancement using re-scheduling and FACTS devices* // Ain Shams Engineering Journal. – 2014. – Vol. 5, Iss. 3. – P. 789–801.
16. STOER M., WAGNER F. *A simple min-cut algorithm* // Journal of the ACM. – 1997. – Vol. 44(4). – P. 585–591.
17. VALENCIA V.V., MAJ P.E. *Network Interdependency Modeling for Risk Assessment on Built Infrastructure Systems*. – 2013. – URL: <https://pdfs.semanticscholar.org/95ba/f36ae65157638a83f82084c39884b3f0fcb2.pdf?ga=2.74081425.2003714931.1570034586-1895369083.1570034586> (дата обращения: 23.03.2023).

METHOD FOR FINDING CUTS FOR ENGINEERING INFRASTRUCTURE MANAGEMENT TASKS

Polina Vandilovskaya, V.A. Trapeznikov Institute of Control Sciences of RAS, Moscow, Engineer (polinavandi@yandex.ru).

Andrey Krygin, V.A. Trapeznikov Institute of Control Sciences of RAS, Moscow, Cand.Sc., Senior Researcher (andreyakr14@gmail.com).

Olga Lukinova, V.A. Trapeznikov Institute of Control Sciences of RAS, Moscow, Doctor of Engineering Sciences, Head Scientist Researcher (lobars@mail.ru)

Alexander Roschin, V.A. Trapeznikov Institute of Control Sciences of RAS, Moscow, Cand.Sc., Senior Researcher (rochinaa@ipu.ru).

Abstract: The goal of the functioning of engineering networks is to ensure the delivery of a particular resource to a consumer, ideally with uninterrupted supply which directly depends on the integrity of the network infrastructure. However, various factors such as attacks by malicious actors, natural disasters, and technological accidents can lead to the disconnection of certain parts of the network, resulting in the disruption of resource delivery. This necessitates the task of identifying the most vulnerable parts of the engineering network in terms of potential damage, which allows appropriate measures to be taken to protect the network from negative factors and ensure maximum uninterrupted resource delivery. Engineering networks are commonly modeled as graph structures, therefore one method of solving this task is to find the cuts of the network graph. While such methods exist, they all have certain limitations. This study proposes a new method for finding all cuts of a directed graph of arbitrary dimension, which eliminates these limitations, describes the algorithm for this method, and provides theoretical justification. The concept of the method is based on constructing graph cuts (multicuts) in such a way that all cuts are identified, and can be done in a reasonable amount of time. Examples of engineering networks where this method can be used as a tool for making rational decisions in operating network objects include power grids, transportation networks, water supply and sewerage networks, as well as communication and telecommunications networks.

Keywords: graph, cut, multicuts, engineering network, free path in graph.

УДК 519.178, 658.26

ББК 22.176

DOI: 10.25728/ubs.2024.111.9

*Статья представлена к публикации
членом редакционной коллегии Н.Н. Непейводой.*

Поступила в редакцию 04.12.2023.

Опубликована 30.09.2024.

МОДЕЛИРОВАНИЕ ВЗАИМОВЛИЯНИЯ СКВАЖИН ДЛЯ АНАЛИЗА ЭФФЕКТИВНОСТИ СИСТЕМ ЗАВОДНЕНИЯ НА МАЛЫХ ВЫБОРКАХ ДАННЫХ

Тырсин А. Н.¹

*(ФГБУН Научно-инженерный центр «Надежность и ресурсы
больших систем и машин» УрО РАН, Екатеринбург;
Уральский федеральный университет, Екатеринбург)*

Кащеев С. Е.²

*(Южно-Уральский государственный университет,
Челябинск)*

На завершающей стадии разработки нефтяных месторождений актуальной проблемой является поддержание приемлемых уровней добычи нефти путем оперативного управления заводнением. Сложность усугубляется все более растущим количеством функционирующих на месторождении скважин и изменчивостью процесса их взаимодействия. Это требует новых подходов, учитывающих указанные тенденции в нефтедобыче. Популярным подходом для анализа эффективности систем заводнения нефтяных месторождений в последние годы стало использование прокси-моделей семейства CRM (capacitance-resistive models, емкостно-резистивные модели), представляющих собой математические модели материального баланса. При этом решают обратную задачу для определения параметров модели. Однако малые размеры выборки данных и большое число функционирующих скважин в системе заводнения ограничивает эффективное практическое применение этого подхода. Цель статьи – повышение оперативности мониторинга систем заводнения за счет снижения размера обучающей выборки данных и расширение масштаба анализируемых систем от нескольких десятков до сотен скважин. Предложены два алгоритма, ориентированных на большие размерности и малые выборки данных. Они апробированы на модельных данных, в которых 60 нагнетательных, 160 добывающих скважин и 17 наблюдений. Преимости нагнетательных скважин – это фактические данные с реальной системы заводнения, дебиты добывающих скважин – это модельные значения с учетом случайных ошибок, присутствующих на практике. Данные алгоритмы продемонстрировали приемлемые характеристики как по точности и быстродействию, так и по возможности их применения для прогноза нефтедобычи.

¹ Александр Николаевич Тырсин, д.т.н., профессор (at2001@yandex.ru).

² Станислав Евгеньевич Кащеев, аспирант (kashchevs@susu.ru).

Ключевые слова: модель материального баланса, система заводнения, маска взаимовлияния скважин, энтропия, прогноз, емкостно-резистивная модель.

1. Введение

В последние годы наблюдается тенденция перехода большинства разрабатываемых нефтяных месторождений на завершающую стадию разработки. Особенностью данной стадии является увеличение воды в добываемой продукции до 60–70% и выше [12]. Это делает актуальной проблему поддержания приемлемых уровней добычи нефти путем оперативного управления заводнением. Оперативная оптимизация систем заводнения становится приоритетной задачей на зрелых, зачастую низкоэффективных месторождениях. Огромные неопределенности гидродинамических моделей, связанные с отсутствием знаний о строении межскважинного пространства, с проявлением масштабных эффектов при переносе керновых данных на ячейки гидродинамической модели, а также с особенностями численного решения описывающих пластовые процессы дифференциальных уравнений, приводят к необходимости использования более простых математических моделей [1, 6, 8]. Существует ряд разновидностей систем заводнения (блоковые, рядные, площадные и др.). В частности, на рис. 1 схематично приведена блоковая система заводнения.

Одним из перспективных направлений сопровождения разработки нефтяных месторождений является использование емкостно-резистивных аналитических моделей CRM, представляющих собой аналитическое решение уравнения динамического материального баланса в виде временного ряда [2, 4, 7, 11, 15, 21]. Модель CRM впервые была описана в 2005 г. в форме дифференциального уравнения, описывающую динамику дебита жидкости [21]. Аббревиатура CRM (Capacitance Resistive Model) отражает известную аналогию между гидродинамическими законами фильтрации в пласте и законами электрического тока [13]. Так, понятие «электрическая емкость» (capacitance) аналогично понятию «запас упругой пласто-

вой энергии», а понятие «электрическое сопротивление» (resistive) аналогично понятию «фильтрационное сопротивление».

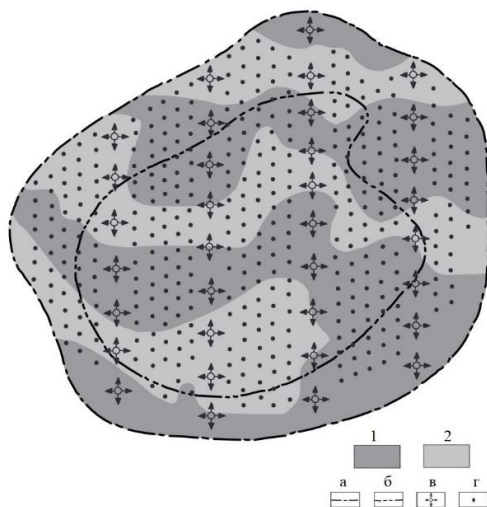


Рис. 1. Блоковая система заводнения.

Зоны с толщиной и коллекторскими свойствами пласта:

- 1) высокими; 2) низкими. Контуры нефтеносности: а) внешний;
б) внутренний; в) нагнетательные; г) добывающие*

Существует множество видов моделей CRM: CRMT, CRMP, CRMIP и др. [7]. Рассмотрим общий вид CRM-модели в варианте CRMP при наличии $L > 1$ добывающих и $M > 1$ нагнетательных скважин. Дебит жидкости для каждой добывающей скважины в k -й момент времени описывается как [10]

$$(1) \quad q_l(k) = q_l(k-1)e^{-1/\tau_l} + (1 - e^{-1/\tau_l}) \sum_{j=1}^M f_{jl} I_j(k) + \varepsilon(k),$$

где $q_l(k)$, $I_j(k)$ – дебит l -й добывающей скважины и приемистость j -й нагнетательной скважины в k -й период времени соответственно; $\tau_l > 0$ – постоянная времени (связана с запасом упругой энергии в пласте и коэффициентом продуктивности скважины и определяет скорость распространения возмущений в пласте);

$f_{jl} \geq 0$ – коэффициент взаимовлияния нагнетательной скважины I_j и добывающей скважины P_l , k – номер периода (например, месяца), $l = 1, \dots, L$, $j = 1, \dots, M$, $k = 1, \dots, N$. Слагаемое $\varepsilon(k)$ в (1) отражает присутствующие в реальных данных случайные ошибки, которые вызваны погрешностями измерений дебитов и приемистостей скважин, неоднородностью пластов, неполнотой соответствия модели рассматриваемому процессу и т.д. В (1), так же как и в [10], сделано допущение о постоянстве забойного давления в нагнетательных скважинах. Это является упрощением и обусловлено следующими причинами: 1) формальный учет изменения забойного давления не увеличит числа неизвестных параметров модели (1), а лишь несколько усложнит вид модели (присутствующие в качестве параметров коэффициенты продуктивности определяются по результатам гидродинамических исследований и эксплуатации скважин и считаются в данном случае известными); 2) целью исследования является в упрощенном варианте постоянства забойного давления изучение возможности определения параметров CRM-модели в условиях малых выборок данных и числа скважин, сопоставимого с реальными системами заводнения; 3) в дальнейшем, безусловно, изменение забойного давления будет включено в модель.

Популярность CRM-моделей обусловлена главным образом тем, что они не требуют задания свойств межскважинного пространства, которые на практике полностью никогда не известны. Как отмечено в [7, с. 14], параметры τ_l и f_{jl} модели (1) позволяют учесть характеристики системы «пласт – скважины» и на практике вполне достаточны для получения довольно качественного прогноза динамики технологических показателей по CRM-модели.

Сопровождение разработки нефтяных месторождений включает в себя решение задач мониторинга системы заводнения и прогнозирования нефтедобычи. Для этого необходимо по имеющимся данным о приемистости нагнетательных и дебите добывающих скважин определить коэффициенты взаимовлияния и постоянные времени, и затем по найденным параметрам и планируемым значениям приемистости нагнетательных сква-

жин сделать прогноз дебита жидкости добывающих скважин. Наличие большого числа скважин, невыпуклый характер функциональной зависимости и наличие ряда ограничений на допустимую область значений приводит к неединственности решения и трудоемкости его нахождения [10].

В [10] предложен новый подход для CRM-моделирования, основанный на определении маски взаимовлияния, что значительно сокращает число определяемых коэффициентов взаимовлияния. Под маской взаимовлияния понимается множество нагнетательных скважин, влияющих на формирование дебита, и соответствующие значения коэффициентов взаимовлияния f_{ji} .

Однако при апробации этого подхода выявлены недостатки.

Во-первых, здесь требуется относительно большое количество обучающих данных, порядка 30 наблюдений для каждой скважины. Такое количество наблюдений не всегда практически реализуемо, поскольку в течение этого периода времени могут измениться параметры модели.

Во-вторых, не решена проблема масштаба реальных систем заводнения, количество скважин на которых может достигать нескольких сотен.

Целью работы является повышение оперативности мониторинга систем заводнения за счет снижения размера обучающей выборки данных и расширение масштаба анализируемых систем от нескольких десятков до сотен скважин.

2. Методы исследования

Одной из проблем оценивания параметров CRM-модели является наличие большого количества добывающих и нагнетательных скважин. На практике их количество может приближаться к сотням и тысячам соответственно. При определении коэффициентов взаимовлияния это приводит к увеличению количества возможных пар взаимовлияния, и как следствие, к вычислительной неустойчивости. Фактически на дебит $q_i(k)$ каждой из добывающих скважин P_l , $l = 1, \dots, L$, влияют не все нагнетательные скважины, а некоторая часть из них. Рассмотрен-

ние только существенных нагнетательных скважин позволит избежать снижения точности оценок коэффициентов взаимовлияния [7]. Множество таких переменных, влияющих на формирование дебита, будем называть маской взаимовлияния.

В рамках CRM-модели (1) будем рассматривать систему заводнения как сетевую структуру, в которой нагнетательные и добывающие скважины стохастически связаны между собой через коэффициенты взаимовлияния. Под сетевой структурой понимается совокупность устойчивых связей между элементами системы, обеспечивающих воспроизводимость при изменяющихся условиях [17]. Сетевая структура помимо элементов системы может содержать подсистемы (подсети), что тоже должно отражаться как взаимосвязи на уровне подсистем [5, 16].

Разобьем задачу оценивания CRM-модели на 3 этапа: 1) формирование предварительной маски взаимовлияния скважин; 2) уточнение маски взаимовлияния и формирование начального приближения для CRM-модели; 3) определение параметров CRM-модели для всей системы заводнения.

2.1. ФОРМИРОВАНИЕ ПРЕДВАРИТЕЛЬНОЙ МАСКИ ВЗАИМОВЛИЯНИЯ

Недостаток информации о свойствах пласта и наличие большого количества добывающих и нагнетательных скважин приводит к необходимости использования статистических методов для предварительного определения маски взаимовлияния.

Под маской взаимовлияния скважин понимается матрица значимых коэффициентов взаимовлияния добывающих и нагнетательных скважин. Минимальным набором входных данных для формирования маски являются координаты скважин, значения приемистостей и дебитов. Дополнительной информацией для формирования матрицы может являться информация о строении пластов, системе разработки.

Сократить размерность задачи можно за счет: 1) обнуления коэффициентов взаимовлияния f_{ji} для нагнетательных скважин, расположенных достаточно далеко от добывающей скважины, для которых отсутствует какая-либо объективная информация о влиянии приемистости скважины I_j на дебит скважины P_i ;

2) обнуления коэффициентов взаимовлияния f_{jl} для нагнетательных скважин, значения которых оказались ниже заданного порогового уровня, формируемого на основе априорной информации о системе скважин [10]. В результате имеем предварительную маску взаимовлияния $Q = \{(j(l), l)\}$, $l = 1, \dots, L$, $j(l) \in \{1, \dots, M\}$ – множество пар номеров, соответствующих значимых нагнетательным скважинам для каждой из добывающих скважин P_l с ненулевыми коэффициентами взаимовлияния, т.е. $\forall (j, l) \notin Q \ f_{jl} = 0$.

Отметим, что количество значимых нагнетательных скважин для каждой добывающей скважины, формирующих коэффициенты взаимовлияния, ограничено числом наблюдений обучающей выборки данных. В частности, для рассматриваемого далее примера из 17 наблюдений для каждой добывающей скважины из-за использования в (1) дебита $q_l(k)$ в текущий k -й и предыдущий $(k-1)$ -й моменты времени число значимых нагнетательных скважин формально ограничено пятнадцатью, так как число фактически имеющихся соотношений (равно 16) не может быть меньше количества неизвестных параметров – 15 коэффициентов взаимовлияния f_{jl} и одного параметра a_l .

2.2. УТОЧНЕНИЕ МАСКИ ВЗАИМОВЛИЯНИЯ И ФОРМИРОВАНИЕ НАЧАЛЬНОГО ПРИБЛИЖЕНИЯ ДЛЯ CRM-МОДЕЛИ

Для каждой добывающей скважины оцениваем параметры CRM-модели с помощью минимизации суммы квадратов невязок между фактическими и модельными значениями с физическими ограничениями на параметры модели:

$$(2) \quad \sum_{k=2}^N (q_l(k) - q_l^*(k))^2 \rightarrow \min_{a_l, b_{j(l),l}}, \quad l = 1, \dots, L, \quad j(l) \in \{1, \dots, M\},$$

$$(3) \quad q_l^*(k) = a_l q_l(k-1) + \sum_{j(l)} b_{j(l),l} I_{j(l)}(k), \quad q_l^*(1) = q_l(1),$$

$$(4) \quad b_{j(l),l} \geq 0, \quad 0 < a_l < 1.$$

где $a_l = e^{-1/\tau_l}$, $b_{jl} = (1 - e^{-1/\tau_l}) f_{jl}$, $\tau_l = -1 / \ln a_l$, $f_{jl} = b_{jl} / (1 - a_l)$.

В результате решения L задач минимизации (2)–(4) получим вектор параметров $a_1^*, \dots, a_L^*$ и множество коэффициентов

взаимовлияния $f_{j(l),l}^*$, $l = 1, \dots, L$, $j(l) \in \{1, \dots, M\}$. Слишком малые значения коэффициентов взаимовлияния, значения которых близки к нулю (например, если $f_{jl} < 10^{-4}$), удаляем из рассмотрения. Пороговый уровень зависит от многих причин, в частности от средних значений дебитов и приемистостей, от величины ошибки, от точности измерений дебитов и приемистостей. Для модельных данных порог задавался примерно равным дроби, в числителе которой минимально допустимая величина по точности коэффициента взаимовлияния ($\sim 0,001$), а в знаменателе – отношение стандартного отклонения случайных ошибок ($\sim 0,2$) к среднему (по скважинам) дебиту нагнетательных скважин (~ 40). Средние значения дебитов и приемистостей имели сопоставимые уровни. Для реальных систем заводнения коэффициенты взаимовлияния точнее, чем 0,01 устойчиво установить вряд ли возможно, поэтому пороговый уровень, очевидно, будет выше, например, не менее чем 10^{-3} .

Далее для каждой добывающей скважины рассмотрим три подсистемы:

1) влияние интерференции Δq_l нагнетательных скважин на дебит скважины P_l формируем как $\Delta q_l(k) = q_l(k) - a_l^* q_l(k-1)$, где a_l^* – оценка параметра a_l в результате решения задачи (2)–(4);

2) совокупность нагнетательных скважин $\mathbf{I}_l = (I_{1(l),l}, \dots, I_{M(l),l})$;

3) совокупность скважин $\mathbf{I}_l$, из которой удалена произвольная j -я скважина $\mathbf{I}_l^j = \mathbf{I}_l \setminus I_{j,l} = (I_{1(l),l}, \dots, I_{j-1(l),l}, I_{j+1(l),l}, \dots, I_{M(l),l})$.

В качестве индикатора используем энтропийный показатель взаимосвязи [9, 10], выражаемый через определители корреляционных матриц как

$$(5) \quad \delta_{jl} = \frac{G(\Delta q_l \cap \tilde{\mathbf{I}}_l) - G(\Delta q_l \cap \tilde{\mathbf{I}}_l^j)}{G(\Delta q_l \cap \tilde{\mathbf{I}}_l^j)} \cdot 100\%,$$

$$\text{где } G(\Delta q_l \cap \mathbf{I}_l) = -\frac{1}{2} \ln \frac{|\mathbf{R}_{\Delta q_l \cup \mathbf{I}_l}|}{|\mathbf{R}_{\mathbf{I}_l}|}, \quad G(\Delta q_l \cap \mathbf{I}_l^j) = -\frac{1}{2} \ln \frac{|\mathbf{R}_{\Delta q_l \cup \mathbf{I}_l^j}|}{|\mathbf{R}_{\mathbf{I}_l^j}|}.$$

Коэффициент δ_{jl} характеризует уменьшение в процентах энтропийного показателя взаимосвязи между системами I_l и Δq_l при удалении из системы I_l произвольной скважины I_j . Из маски взаимовлияния исключаем коэффициенты взаимовлияния f_{jl} нагнетательных скважин, для которых δ_{jl} меньше порогового уровня, который задаем на основе априорной информации о системе скважин и требуемой точности оценивания параметров CRM-модели, как показано выше. Далее еще раз для каждой добывающей скважины решаем задачи (2)–(4) с уточненным набором значимых для нее нагнетательных скважин и определяем оценки коэффициентов взаимовлияния $f_{j(l),l}^0$ и параметров a_l^0 .

Решение L задач (2)–(4) позволяет избежать процедуры вычисления приведенных преимуществ нагнетательных скважин, учитывающих эффект запаздывания реакции добывающих скважин на изменение режимов нагнетательных скважин, кото-

рые определены в [10] как
$$\tilde{I}_{jl}(k) = \frac{\sum_{t=0}^{k-1} (a_l^*)^t I_j(k-t)}{\sum_{t=0}^{k-1} (a_l^*)^t},$$
 где пара-

метры a_l^* определяются в результате решения L задач минимизации (2)–(4). Это значительно сокращает требуемый объем выборки для определения параметров системы заводнения, который зависит от максимального значения постоянной времени τ_l добывающих скважин.

Чем больше постоянная времени τ_l , тем больше параметр a_l^* и тем медленнее сходится конечная сумма $\sum_{t=0}^{k-1} (a_l^*)^t$ к сумме бесконечно убывающей геометрической прогрессии, т.е. растет число наблюдений для расчета значений $\tilde{I}_{jl}(k)$. Например, если $\max_l \tau_l = 6$, то, как показано в [10], вычисление приведенных преимуществ нагнетательных скважин с приемлемой точностью потребует не менее 15 дополнительных наблюдений (при этом $a_l = 0,85$ и $(a_l)^{15} = 0,087 < 0,1$). В результате первых 15 наблюдений выборки потребуются для «стабилизации» рассчитываемых значений приведенных преимуществ и мини-

мально допустимый объем выборки данных достигает порядка 30 наблюдений, что не всегда возможно на практике.

2.3. ОПРЕДЕЛЕНИЕ ПАРАМЕТРОВ CRM-МОДЕЛИ ДЛЯ ВСЕЙ СИСТЕМЫ ЗАВОДНЕНИЯ

Далее определим параметры CRM-модели всей системы заводнения. Рассмотрим два алгоритма: 1) непосредственное решение оптимизационной задачи (алгоритм А); 2) итерационный алгоритм В.

Алгоритм А решает задачу:

$$(6) \sum_{k=2l=1}^N \sum_{l=1}^L (q_l(k) - q_l^*(k))^2 \rightarrow \min_{a_l, f_{j(l),l}},$$

$$(7) q_l^*(k) = a_l q_l(k-1) + (1-a_l) \sum_{j(l)} f_{j(l),l} I_{j(l)}(k), \quad l=1, \dots, L,$$

$$(8) q_l^*(1) = q_l(1), \quad 0 < a_l < 1, \quad l=1, \dots, L,$$

$$(9) f_{j(l),l} \geq 0, \quad \sum_{l=1}^L f_{j(l),l} \leq 1, \quad l=1, \dots, L, \quad j(l) \in \{1, \dots, M\}.$$

Алгоритм В включает шаг 0 и циклически повторяет шаги 1 и 2 ($k=1, 2, \dots$):

Шаг 0. Вычисление коэффициентов взаимовлияния $f_{j(l),l}^{(0)}$ для всей CRM-модели по найденным при решении L задач (2)–(4) параметрам $a_1^{(0)}, \dots, a_L^{(0)}$:

$$(10) \sum_{k=2l=1}^N \sum_{l=1}^L (q_l(k) - q_l^*(k))^2 \rightarrow \min_{f_{j(l),l}},$$

$$(11) q_l^*(k) = a_l^{(0)} q_l(k-1) + (1-a_l^{(0)}) \sum_{j(l)} f_{j(l),l} I_{j(l)}(k), \quad q_l^*(1) = q_l(1),$$

$$(12) f_{j(l),l} \geq 0, \quad \sum_{l=1}^L f_{j(l),l} \leq 1, \quad l=1, \dots, L, \quad j(l) \in \{1, \dots, M\}.$$

Шаг 1^(k). Вычисление коэффициентов $a_1^{(k)}, \dots, a_L^{(k)}$ для всей CRM-модели по найденным на предыдущем шаге коэффициентам взаимовлияния $f_{j(l),l}^{(k-1)}$:

$$(13) \sum_{k=2l=1}^N \sum_{l=1}^L (q_l(k) - q_l^*(k))^2 \rightarrow \min_{a_l},$$

$$(14) q_l^*(k) = a_l (q_l(k-1) - \sum_{j(l)} f_{j(l),j}^{(k-1)} I_{j(l)}(k)) + \sum_{j(l)} f_{j(l),j}^{(k-1)} I_{j(l)}(k),$$

$$(15) \quad q_l^*(1) = q_l(1), \quad 0 < a_l < 1, \quad l = 1, \dots, L, \quad j(l) \in \{1, \dots, M\}.$$

Шаг $2^{(k)}$. Вычисление коэффициентов взаимовлияния $f_{j(l),l}^{(k)}$ для всей CRM-модели по найденным на предыдущем шаге параметрам $a_1^{(k)}, \dots, a_L^{(k)}$:

$$(16) \quad \sum_{k=2l=1}^N \sum_{l=1}^L (q_l(k) - q_l^*(k))^2 \rightarrow \min, \quad f_{j(l),l}$$

$$(17) \quad q_l^*(k) = a_l^{(k)} q_l(k-1) + (1 - a_l^{(k)}) \sum_{j(l)} f_{j(l),l} I_{j(l)}(k), \quad q_l^*(1) = q_l(1),$$

$$(18) \quad f_{j(l),l} \geq 0, \quad \sum_{l=1}^L f_{j(l),l} \leq 1, \quad l = 1, \dots, L, \quad j(l) \in \{1, \dots, M\}.$$

Цикл останавливается, если уменьшение целевой функции станет незначительным.

Алгоритмы для решения оптимизационных задач (2)–(4), (6)–(18) реализованы на языке Python с использованием глобально сходящегося метода скользящих асимптот (Method of moving asymptotes, MMA) для локальной оптимизации на основе градиента из библиотеки нелинейной оптимизации NLOpt [14]. MMA впервые был описан в [19]. Он относится к семейству консервативных выпуклых сепарабельных аппроксимаций CCSA и зарекомендовал себя в области структурной оптимизации, где вычисления функций и градиентов занимают много времени и обладает глобальной сходимостью и высоким быстродействием [14]. Суть метода состоит в том, что в каждой текущей точке генерируется локальная аппроксимация с помощью градиента функции, условий ограничения и квадратичного штрафного слагаемого, чтобы сделать аппроксимацию консервативной.

Для малых выборок данных при большой размерности задачи использованный в [10] алгоритм ISRES (Improved Stochastic Ranking Evolution Strategy) из библиотеки NLOpt [3, 14, 20] не показал одинаково устойчивой работы, поэтому вместо него был использован метод скользящих асимптот. ISRES является эволюционным алгоритмом, основанным на комбинации правила мутации (с логарифмически нормальным обновлением размера шага и экспоненциальным сглаживанием) и дифференциальной вариации (правило обновления, подобное пра-

вилу Нелдера – Мида). В оптимизационных задачах (2)–(4), (6)–(18) отсутствуют нелинейные ограничения, поэтому ранжирование пригодности осуществляется только с помощью целевой функции. ISRES использует большое количество случайно сгенерированных наборов параметров, что уменьшает скорость сходимости алгоритма.

Для оценки точности CRM-моделирования вычислялись среднеквадратические отклонения (с.к.о.) ошибок коэффициентов взаимовлияния, ошибок запаздывания и средняя абсолютная ошибка прогноза.

3. Численные эксперименты

Рассмотрим пример модели реальной системы заводнения нефтяного месторождения. Система состоит из $M = 60$ нагнетательных скважин и $L = 160$ добывающих скважин. Приемистости нагнетательных скважин – это фактические данные с реальной системы заводнения, для каждой скважины I_j , $j = 1, \dots, M$, имеется по $N = 17$ наблюдений, зафиксированных через равные интервалы времени. Матрица коэффициентов взаимовлияния f_{ji} формировалась случайным образом с учетом нормировки ($f_{ji} \geq 0$, $\sum_{j=1}^M f_{ji} \leq 1$), для каждой добывающей скважины число ненулевых

коэффициентов взаимовлияния варьировалось от 1 до 7, всего их оказалось 554. Постоянные времени τ формировались случайным образом в диапазоне от 1 до 6. Дебиты $q_i(k)$ добывающих скважин P_i формировались по формуле (1).

Для проверки устойчивости алгоритмов А и В к случайным погрешностям сформировано несколько наборов данных дебитов с различными с.к.о. случайной погрешности $\varepsilon(k)$: ($\sigma_\varepsilon = 0; 0,1; 0,2; 0,3$). Алгоритмы оценивались по таким параметрам как скорость сходимости алгоритма (время вычислений), значение целевой функции, с.к.о. ошибок оценок коэффициентов взаимовлияния и параметров запаздывания. С.к.о. ошибок коэффициентов взаимовлияния вычислялось как

$$(19) RMSE_f = \sqrt{\frac{\sum_{j=1}^M \sum_{l=1}^L (f_{jl} - \hat{f}_{jl})^2}{L \cdot M}}, \quad l = 1, \dots, L, \quad j = 1, \dots, M,$$

где f_{jl} и $\hat{f}_{jl}$ – фактические и прогнозные значения коэффициентов взаимовлияния.

После формирования предварительной маски по формулам (2)–(4) для каждой добывающей скважины уточнялось множество нагнетательных скважин с использованием энтропийного показателя (5). Пороговый уровень для δ_l зададим равным 2,5%. Показатель (5) при введении случайных погрешностей $\varepsilon(k)$ допустил порядка 30÷35 ошибок 1-го рода (значимый коэффициент взаимовлияния скважин идентифицирован как не значимый). Ошибок 2-го рода нет. Ошибки 1-го рода возникали в ситуациях, когда фактические коэффициенты взаимовлияния были очень малы (не более 0,02), т.е. они объясняются наличием сопоставимых случайных погрешностей.

Дополнительно реализован алгоритм А без использования маски взаимовлияния, т.е. минимизация выполнялась по всей матрице коэффициентов взаимовлияния. Соответственно, в задаче минимизации использовалось 9760 параметров: 160 коэффициентов запаздывания и 9600 коэффициентов взаимовлияния. В таблице 1 приведено сравнение реализованных алгоритмов.

Алгоритм А без использования маски взаимовлияния с ростом с.к.о. случайной погрешности показал значительное ухудшение качества моделирования и уступает остальным по всем показателям.

Алгоритмы А и В показали похожие результаты и продемонстрировали устойчивость к случайным погрешностям. Алгоритм А работает быстрее в 1,3÷1,7 раза, но проигрывает алгоритму В по точности оценивания (величине минимума целевой функции) в 1,1÷1,5 раза. Однако расхождения могут быть вызваны малой выборкой: всего рассматривалось по три реализации случайных погрешностей.

Таблица 1. Результаты работы алгоритмов

Показатели эффектив- ности	σ	Алгоритм А, ММА	Алгоритм А без маски взаимо- влияния, ММА	Алгоритм В, ММА	Алгоритм А, ISRES
Скорость сходимости, сек	0	33,4	97,9	56,5	369,3
	0,1	48,1	140,7	82,6	877,0
	0,2	57,4	281,0	76,1	875,3
	0,3	59,5	198,9	86,8	844,9
Значение целевой функции	0	0,0	12462,5	0,0	0,0
	0,1	48,0	11674,0	31,4	241,1
	0,2	116,6	10835,9	109,3	1476,5
	0,3	293,5	13361,7	257,3	3296,6
С.к.о. f_l	0	0,000	0,025	0,000	0,000
	0,1	0,002	0,025	0,002	0,002
	0,2	0,005	0,025	0,005	0,004
	0,3	0,006	0,025	0,006	0,005
САО прогноза, м ³	0	1,34	6,12	1,34	1,34
	0,1	1,36	6,17	1,34	1,51
	0,2	1,47	6,25	1,46	2,04
	0,3	1,65	5,99	1,63	2,59
САО прогноза, %	0	3,45	15,76	3,45	3,45
	0,1	3,50	15,89	3,45	3,89
	0,2	3,78	16,09	3,76	5,25
	0,3	4,25	15,42	4,20	6,67

Для оценки качества прогнозирования сформировано дополнительно $NP = 17$ наблюдений. Средняя абсолютная ошибка прогноза равна

$$(20) D = \frac{1}{L \cdot (NP - N)} \sum_{k=N+1}^{N+NP} \sum_{l=1}^L |q_l(k) - \hat{q}_l(k)|,$$

где $\hat{q}_l(k)$ – прогнозируемое значение дебита для скважины P_l в k -й период времени, $q_l(k)$ – фактическое значение дебита для скважины P_l в k -й период времени.

На рис. 2 в качестве иллюстрации показан пример фактических и прогнозных значений дебита по двум скважинам P_5 и P_{133} .

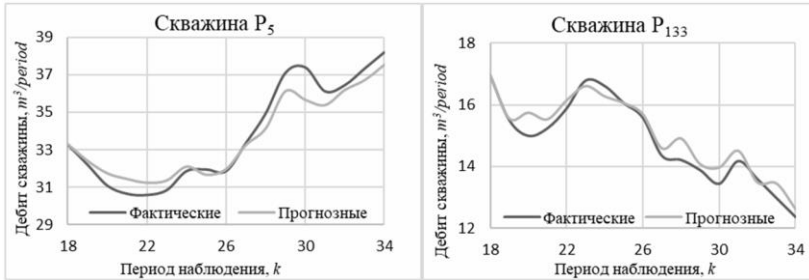


Рис. 2. Фактические и прогнозные значения дебита по скважинам P_5 и P_{133}

Для реализованных алгоритмов D составила $1,34 \text{ м}^3$, или 3,45% от среднего фактического дебита по всем скважинам в периоде прогноза. Без использования маски ошибка прогноза составила $6,61 \text{ м}^3$, или 17,02% от среднего фактического дебита по всем скважинам в периоде прогноза. Таким образом, алгоритмы с использованием маски взаимовлияния предпочтительнее для задач мониторинга и прогнозирования уровня нефтедобычи.

4. Заключение

Показано преимущество использования маски взаимовлияния для оценки параметров CRM-модели. Описан способ предварительной оценки маски с использованием энтропийного показателя взаимосвязи. Данный способ позволяет уточнить для каждой добывающей множество влияющих нагнетательных скважин.

Определение маски взаимовлияния значительно снижает размерности задач оптимизации и повышает точность CRM-моделирования. Так, в рассмотренном примере количество параметров минимизации было уменьшено в 14 раз (с 9760 до 712), что уменьшило время сходимости алгоритма и увеличило точность оценивания параметров CRM-модели в среднем в 14 раз, в отдельных случаях – от 3,7 до 42 раз.

Двукратное уменьшение обучающей выборки по сравнению с известным решением достигнуто за счет формирования предварительной маски взаимовлияния и начального приближения для CRM-модели путем отдельного рассмотрения всех добывающих скважин. Малый размер обучающей выборки соответствует реальным условиям эксплуатации систем заводнения.

Алгоритмы показали устойчивость к случайным погрешностям и достаточно точные результаты прогноза. Рассмотренные размерности соответствуют реальным системам заводнения. В рассмотренном примере апробация проводилась на модельных данных, полученных по модели (1) без учета забойного давления. В дальнейшем планируется учесть забойное давление и апробировать предложенный подход для мониторинга реальных систем заводнения и прогнозирования нефтедобычи.

Литература

1. АЗИЗ Х., СЕТТАРИ Э. Математическое моделирование пластовых систем: пер. с англ. – М.: Недра, 1982. – 408 с.
2. АФАНАСКИН И.В., КРЫГАНОВ П.В., ГЛУШАКОВ А.А., и др. *Использование CRM-моделей интерференции скважин для оценки фильтрационно-емкостных свойств пласта по данным разработки* // Успехи кибернетики. – 2020. – №1(1). – С. 17–27.
3. *Библиотека алгоритмов NLOpt для нелинейной оптимизации.* – URL: https://nlopt.readthedocs.io/en/latest/NLOpt_Algorithms.
4. ДАНЬКО М.Ю., БРИЛЛИАНТ Л.С., ЗАВЬЯЛОВ А.С. *Применение метода динамического материального баланса и CRM-метода к подсчету запасов ачимовских и баженовских коллекторов* // Недропользование XXI век. – 2019. – №4(80). – С. 76–85.
5. НОВИКОВ Д.А. *Сетевые структуры и организационные системы.* – М.: ИПУ РАН, 2003. – 102 с.
6. ПОСПЕЛОВА Т.А., СТЕПАНОВ С.В., СТРЕКАЛОВ А.В. и др. *Математическое моделирование для принятия реше-*

ний по разработке месторождений. – М.: Недра, 2021. – 427 с.

7. СТЕПАНОВ С.В., БЕКМАН А.Д., РУЧКИН А.А. и др. *Сопровождение разработки нефтяных месторождений с использованием моделей CRM.* – Тюмень: ИПЦ «Экспресс», 2021. – 300 с.
8. СТЕПАНОВ С.В., ТЫРСИН А.Н., РУЧКИН А.А. и др. *Использование энтропийного моделирования для анализа эффективности системы заводнения* // Нефтяное хозяйство. – 2020. – №6. – С. 62–67.
9. ТЫРСИН А.Н. *Энтропийное моделирование сетевых структур* // Автоматика и телемеханика. – 2022. – № 10. – С. 144–155. DOI: 10.31857/S0005231022100130
10. ТЫРСИН А.Н., СТЕПАНОВ С.В., РУЧКИН А.А. и др. *Повышение достоверности моделирования взаимовлияния скважин для анализа эффективности системы заводнения* // Математическое моделирование. – 2023. – Т. 35, №6. – С. 63–80.
11. ХАТМУЛЛИН И.Ф., ЦАНДА А.П., АНДРИАНОВА А.М. и др. *Полуаналитические модели расчета интерференции скважин на базе класса моделей CRM* // Нефтяное хозяйство. – 2018. – №12. – С. 38–41.
12. ЮШКОВ И.Р., ХИЖНЯК Г.П., ИЛЮШИН П.Ю. *Разработка и эксплуатация нефтяных и газовых месторождений.* – Пермь: ПНИПУ, 2013. – 177 с.
13. HOLANDA R.W., GILDIN E., JENSEN J.L. et al. *A State-of-the-Art Literature Review on Capacitance Resistance Models for Reservoir Characterization and Performance Forecasting* // Energies. – 2018. – Vol. 11. – 3368. – 45 p.
14. JOHNSON S.G. *The NLopt nonlinear-optimization package.* – URL: <http://github.com/stevengj/nlopt>.
15. KIM J.S., LAKE L., EDGAR T.F. *Integrated Capacitance-Resistance Model for Characterizing Waterflooded Reservoirs* // Proc. of the 2012 IFAC Workshop on Automatic Control in Off-shore Oil and Gas Production, Norwegian University of Science and Technology, Trondheim, Norway, May 31 – June 1. – 2012. – P. 19–24.

16. LINDENLAUB I., PRUMMER A. *Network Structure and Performance Get access Arrow* // The Economic Journal. – 2021. – Vol. 131, No. 634. – P. 851–898.
17. PULLAN W. *Structure*. – Cambridge: Cambridge University Press, 2000.
18. RUNARSSON T.P., XIN YAO. *Search Biases in Constrained Evolutionary Optimization* // IEEE Trans. on Systems Man and Cybernetics – Part C: Applications and Reviews. 2005. – Vol. 35, No. 2. – P. 233–243.
19. SVANBERG K. *The Method of Moving Asymptotes – A New Method for Structural Optimization* // Int. Journal for Numerical Methods in Engineering. – 1987. Vol. 24. – No. 2. – P. 359–373.
20. TUO CHEN, JIANCHUAN XIANYU. *Application of stochastic ranking based evolutionary strategy in environmental management of microgrids* // Journal of Physics: Conference Series. 2023. – Vol. 2477. – 012076.
21. YOUSEF A.A., GENTIL P.H., JENSEN J.L. et al. *A Capacitance Model to Infer Interwell Connectivity from Production and Injection Rate Fluctuations* // SPE Annual Technical Conference and Exhibition, 9–12 October 2005, Dallas, Texas.

WELL INTERACTION MODELING TO ANALYZE FLOODING SYSTEMS EFFICIENCY ON SMALL DATA SAMPLES

Alexander Tyrsin, Science and Engineering Center “Reliability and Resource of Large Systems and Machines”, Ural Branch of RAS; Ural Federal University, Yekaterinburg, Doctor of Science, professor (at2001@yandex.ru).

Stanislav Kashcheev, South-Ural State University, Chelyabinsk (kashcheevs@susu.ru).

Abstract: At the final stage of oil field development, an urgent problem is to maintain acceptable levels of oil production through operational flooding management. The complexity is compounded by the increasing number of wells operating in the field and the variability of the process of their interaction. This requires new approaches that consider these trends in oil production. A popular approach for analyzing the effectiveness of oilfield flooding systems in recent years has been the use of proxy models of the CRM family (capacitance-resistive models), which are math-

ematical models of material balance. At the same time, the inverse problem is solved to determine the model parameters. However, the small size of the data samples and the large number of functioning wells in the flooding system limits the effective practical application of this approach. The purpose of the article is to increase the efficiency of monitoring water flooding systems by reducing the size of the training data sample and expanding the scale of the analyzed systems from several tens to hundreds of wells. Two algorithms focused on large dimensions and small data samples are proposed. They were tested on model data in which there were 60 injection and 160 production wells, and 17 observations and random errors were present. The injectivity of injection wells is actual data from a real water flooding system. The flow rates of production wells are model values, taking into account random errors present in practice. These algorithms have demonstrated acceptable characteristics both in terms of accuracy and speed, and if possible, their application for forecasting.

Keywords: material balance model, flooding system, well interference mask, entropy, forecast, capacitance resistive model.

УДК 622.276.43: 519.237: 519.8

ББК 33.36 + 22.172 + 22.19

DOI: 10.25728/ubs.2024.111.10

*Статья представлена к публикации
членом редакционной коллегии А.Г. Лосевым.*

Поступила в редакцию 03.04.2024.

Опубликована 30.09.2024.

МЕТОД ВЫБОРА ЭНЕРГОЭКОНОМИЧНОГО ПУТИ ОБХОДА СЛОЖНОГО ПРЕПЯТСТВИЯ АВТОНОМНЫМ НЕОБИТАЕМЫМ ПОДВОДНЫМ АППАРАТОМ¹

Мартынова Л. А.², Павлов А. А.³

(АО «Концерн «ЦНИИ «Электроприбор», Санкт-Петербург)

Для выбора пути обхода сложного препятствия разработан метод, основанный на данных батиметрических и физических карт Мирового океана. Определены условия применимости батиметрических и физических карт с мезорельефом для формирования пути обхода. Для перехода от макро- и мезорельефа к микрорельефу предложено использовать алгоритм Кригинга; на примере показана его работа. Для определения наиболее энергоэкономичного пути обхода сложных препятствий предложено рассматривать сетку глубин в виде ориентированного взвешенного графа. Определены условия соответствия кратчайшего пути на графе наиболее энергоэкономичному пути. Разработан алгоритм перебора вариантов пути обхода с отсечением, базирующийся на алгоритме «поиска на графе в глубину». На примере обхода сложного препятствия показано преимущество использования предложенного метода выбора варианта обхода препятствия по сравнению с традиционно предлагаемыми обходами сверху или сбоку препятствия. Приведен пример использования разработанного метода для определения наиболее энергоэкономичного пространственного пути обхода сложного препятствия. Результаты проведенных численных экспериментов подтвердили правильность предложенного решения: сокращение пути для рассмотренного варианта препятствия составило 15–20%, что позволяет реализовать предложенный метод выбора энергоэкономичного пространственного пути обхода сложного препятствия в системе управления автономного необитаемого подводного аппарата.

Ключевые слова: автономный необитаемый подводный аппарат, рельеф дна, батиметрическая карта, алгоритм Кригинга, поиск кратчайшего пути на графе, энергоэкономичный путь.

¹ Исследование выполнено за счет гранта Российского научного фонда №23-29-00803, <https://rscf.ru/project/23-29-00803/>.

² Любовь Александровна Мартынова, д.т.н., в.н.с. (martynowa999@bk).

³ Алексей Александрович Павлов, к.т.н., ведущий инженер-программист (ialex.science@gmail.com).

1. Введение

В настоящее время совершенствование технологий привело к возможности создания АНПА для решения широкого круга задач подводного наблюдения [5] при выполнении протяженного маршрутного задания (МЗ) в течение длительного времени. АНПА дальнего плавания в ходе выполнения МЗ способен преодолевать значительные расстояния – свыше 10 тыс. км в течение нескольких месяцев.

Несмотря на прокладку маршрута в обход областей навигационных опасностей, по ряду причин на пути следования АНПА возможно возникновение препятствий:

- не всегда при прокладке маршрута существует возможность обхода зон навигационной опасности;
- из-за нарастания с течением времени погрешности навигационных определений АНПА возможно его отклонение от МЗ;
- прогноз условий внешней среды на длительный период невозможен, а их влияние может привести к отклонению АНПА от МЗ;
- из-за наличия неучтенных препятствий, связанных с неточностями существующих навигационных карт.

Таким образом, необходимо допускать возможность попадания АНПА в зону с препятствиями, что определяет актуальность проводимых исследований.

Получение первичной информации о возникновении препятствия по курсу движения АНПА происходит по данным гидролокатора секторного обзора (ГСО), зондирующего пространство перед АНПА. Регистрация приемником ГСО отраженного сигнала означает, что впереди по курсу движения АНПА находится препятствие. Из-за ограничения по углу сектора обзора ГСО, а также из-за ограниченной дальности действия ГСО, система управления АНПА не способна сразу оценить размер и форму препятствия и определить энергоэкономичный путь его обхода. Под энергоэкономичным путем обхода препятствия будем понимать такой путь, при движении по которому происходит наименьший расход энергоресурса АНПА.

В имеющихся публикациях, посвященных обходу препятствия с использованием АНПА, рассматривается реакционный подход к выбору пути обхода, т.е. без учета формы всего препятствия: АНПА вынужден выбирать безопасное направление по текущим показаниям ГСО [4]. Так, в ряде источников [1, 3] предлагается обходить препятствие сверху, однако препятствие может возвышаться над морской поверхностью и попытка обойти его сверху становится бессмысленной, что приводит к неоправданному перерасходу энергоресурса. В случае невозможности обхода препятствия сверху в [2, 4] предлагается попытаться обойти его сбоку, наугад выбирая направление обхода слева или справа, определяясь в процессе обхода с глубиной движения [2].

Публикации по обходу препятствия АНПА связаны с определением пути по текущим измерениям отраженного от препятствия сигнала и детально описаны в [9, 16, 17, 21, 22, 24]. Так, в [17] представлена разработка алгоритма обхода препятствий для АНПА на основе нечеткой реактивной архитектуры для различных скоростей движения аппарата. В [21] основное внимание уделено онлайн-планированию обхода препятствий. Для улучшения автономной способности и интеллекта при планировании обхода препятствий предлагается рекуррентная нейронная сеть со сверткой. В [24] предложен алгоритм обхода препятствий по трехмерной траектории. В [16] предложен подход к построению архитектуры системы управления АНПА для предотвращения столкновений и движения АНПА в неструктурированных средах. Подход сочетает в себе как реактивный, так и совещательный компоненты, а традиционная трехуровневая архитектура расширена за счет включения четвертого уровня сценария. В [22] предложен новый алгоритм планирования пути АНПА в трехмерном заранее неизвестном пространстве, основанный на нечеткой логике, в котором задача трехмерного планирования пути разбивается на две независимые задачи планирования в горизонтальной и вертикальной плоскостях соответственно. Полученные решения объединяются с помощью весового нечеткого логического контроллера для генерации окончательной команды движения. В [13] для колесного робота пред-

ложен выбор кратчайшего пути достижения целевой точки в условиях препятствий на основе нейронных сетей.

Таким образом, применительно к подводной среде имеющиеся публикации посвящены обходу препятствий мезорельефа и микрорельефа. Однако при выполнении протяженного МЗ исчисление ведется сотнями километров, и в этом случае используется информация макрорельефа.

Для того чтобы АНПА смог сформировать обход препятствия по наиболее энергоэкономичному пути, ему необходимо:

- идентифицировать препятствие по той информации, которая имеется у него в базе данных; в базу данных могут быть загружены батиметрические данные тех участков дна [25], через которые проложено МЗ;
- оценить свое местоположение относительно препятствия;
- получить максимально возможную информацию о размере и форме препятствия.

Для безопасности АНПА целесообразно перед попаданием в зону с препятствиями запланировать обсервацию по сигналам спутниковых навигационных систем.

При наличии у АНПА информации о своем местоположении и батиметрических данных карты макрорельефа дна и физической карты морской поверхности АНПА способен идентифицировать препятствие и оценить свое местоположение относительно него. Этой информации достаточно для выбора макро-направления обхода препятствия. Так при выходе препятствия на морскую поверхность бессмысленно пытаться обойти его сверху, а при существенном превышении пути обхода препятствия слева выбрать путь обхода справа и наоборот.

В настоящее время благодаря развитию средств измерений глубин с привязкой результатов измерений к навигационным определениям и комплексной обработки результатов измерений можно использовать батиметрические карты, тем более что маршруты АНПА дальнего плавания прокладываются, как правило, по заранее обследованным акваториям.

Для того чтобы АНПА смог по данным батиметрической съемки сформировать энергоэкономичный путь, необходимо предварительно проанализировать возможность получения по

данным мезорельефа уточненные данные микрорельефа, загрузить массив узловых точек макрорельефа и мезорельефа в бортовой компьютер АНПА по всем препятствиям, которые потенциально могут встретиться на долгом пути АНПА с учетом отклонения АНПА от МЗ из-за накопления погрешности навигационных определений. Для применения в ходе определения энергоэкономичного пути алгоритма Кригинга [10, 11, 15] для измельчения сетки глубин необходимо соблюдать [20, 23] принцип несмещенности среднего: взятые все вместе значения на карте должны иметь правильное среднее значение. Глобальная несмещенность формально обеспечивается за счет повышения низких значений и уменьшения высоких. Тогда при правильных выбранных априорных предположениях алгоритм Кригинга дает наилучшее линейное несмещенное предсказание промежуточных значений. Для сокращения памяти бортового компьютера при загрузке батиметрической карты с измельченной сеткой глубин можно использовать технологию запекания текстур [26], транспонирующую текстуру и детали с высокополигональной модели на низкополигональную. Это позволит перейти от 3D-модели рельефа к двумерной.

При наличии перечисленной выше информации и соблюдения ограничений система управления АНПА имеет все данные о форме препятствия, что позволяет обоснованно выбрать энергоэкономичный путь обхода обнаруженного препятствия.

В связи с отсутствием в публикациях описания выбора пространственного пути обхода препятствия морским робототехническим комплексом – АНПА – на основе информации о его форме целью работы явилось определение наиболее энергоэкономичного пути обхода препятствия.

2. Постановка задачи

Имеется АНПА, выполняющий протяженное МЗ. МЗ состоит из участков l_i , $i=1, \dots, M$; каждый из участков характеризуется координатами начала – конца и параметрами движения АНПА.

Пусть предварительно перед выполнением МЗ проведена съемка рельефа дна и по результатам съемки сформирована карта микрорельефа в виде сетки со значениями глубин h_i в узловых точках; значения узловых точек внесены в базу данных системы управления АНПА. Для сокращения объема вносимой в базу данных АНПА информации о препятствии проведено сжатие данных, в результате которого сформирована информация о мезорельефе дна в месте положения препятствия.

По каждому из режимов движения АНПА известны затраты энергоресурса; наименьшие затраты возникают при движении АНПА в режиме экономичного.

Предполагается, что перед входением АНПА в зоны с препятствиями проводится обсервация его местоположения по сигналам спутниковых навигационных систем.

Необходимо: сформировать энергоэкономичный путь обхода препятствия, т.е. определить координаты начала и конца участков пути обхода так, чтобы суммарный расход энергии по всем участкам пути обхода был минимальным на множестве альтернативных решений:

$$(1) \quad E = \sum_{i=1}^{M_r} e_i \rightarrow \min_{L_j},$$

где $i = 1, \dots, M_r$; M_r – количество участков пути обхода; L_j – множество альтернативных решений пути обхода; j – количество альтернативных вариантов путей обхода.

3. Предлагаемое решение

Для определения координат участков пути обхода необходимо сформировать путь обхода препятствия.

Для формирования энергоэкономичного пути необходимо использовать данные микрорельефа. Для этого предлагается алгоритмами системы управления АНПА измельчить грубую сетку батиметрических данных о препятствии из базы данных АНПА. Для измельчения сетки предлагается использовать алгоритм Кригинга. Это позволит по данным полученного микрорельефа сформировать множество путей обхода препятствия, из

которых затем выбрать наиболее энергоэкономичный путь, обеспечивающий минимальный расход энергии при обходе препятствия.

Алгоритм Кригинга основан на том, что расстояние или направление между узловыми точками отражает пространственная корреляция, которую используют для объяснения изменений рельефа дна. Для этого полагаем, что обеспечены наименьшая разница между фактическими и оцененными содержаниями (минимизация дисперсии погрешности) и отсутствие погрешности (несмещенность оценки).

Алгоритм Кригинга включает поисковый статистический анализ данных, моделирование вариограммы и создание поверхности морского дна. Для построения модели структуры измеряемых точек сформируем эмпирическую вариограмму, вычисленную для всех пар узловых точек, разделенных расстоянием h , с использованием уравнения

$$(2) \quad f(d_{ij}) = \frac{1}{2} \frac{(h_i - h_j)^2}{N},$$

где N – количество пар узловых точек; i и j – номера пар узловых точек; h_i, h_j – глубины в узловых точках; $f(d_{ij})$ – среднее значение глубины по результатам обработки пары точек i и j .

Пусть, например, рельеф дна имеет вид, представленный на рис. 1.

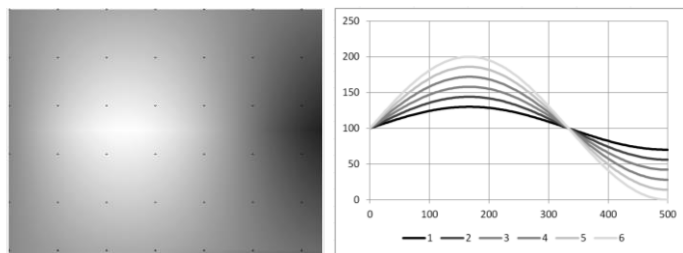


Рис. 1. Фрагмент рельефа морского дна и профиля глубин

На рис.1 слева оттенками серого цвета отражены глубины: с увеличением глубины цвет приобретает более темный оттенок. Черными точками показаны положения узлов батиметриче-

ской карты. Справа на рис.1 изображены профили глубины для вертикальных разрезов препятствия плоскостью. Из рис. 1 видно, что количество узловых точек, покрывающих препятствие, исчисляется единицами и не позволяет детально построить путь обхода данного препятствия. Для этого сформируем сетку с более крупным масштабом.

Вид вариограммы поверхности морского дна, изображенного на рис. 1, показан точками на рис. 2.

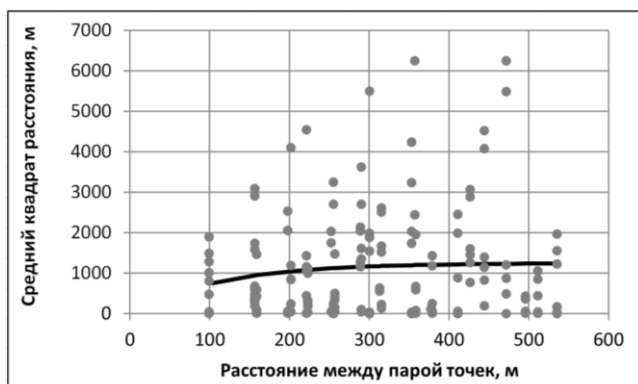


Рис. 2. Вид вариограммы в виде точек

Для прогнозирования глубин в узлах измельченной сетки морского дна сформируем математическую модель глубин в виде экспоненциальной функции вида

$$(3) \quad y(h) = \begin{cases} c_0 + c_1 \left(1 - \exp\left(\frac{-h}{c_2}\right) \right), & h > 0, \\ 0, & h = 0; \end{cases}$$

где c_0 , c_1 , c_2 – коэффициенты функции; h – глубина.

С использованием метода наименьших квадратов запишем уравнение

$$(4) \quad \Delta y = (y(h) - (c_0 + c_1(1 - \exp(-h/c_2))))^2 \rightarrow \min.$$

Найденные значения коэффициентов: $c_0 = 0$, $c_1 = 1250$, $c_2 = 110$; уравнение регрессии приняло вид:

$$y = 1250 \cdot (1 - \exp(-h/110)).$$

Зависимость, соответствующая этому уравнению, показана на рис. 2 сплошной линией.

Предсказание для неизмеренного местоположения глубины $Z(s_0)$ сформируем как взвешенную сумму данных:

$$(5) \quad \hat{Z}(s_0) = \sum_{i=1}^n \lambda_i Z(s_i),$$

где $Z(s_i)$ – измеряемое значение в местоположении i ; λ_i – неизвестный вес для измеряемого значения в местоположении i ; s_0 – местоположение прогноза, n – количество точек, используемых при определении глубины в произвольной точке.

Для расчета весов n узловых точек сформируем систему линейных уравнений:

$$(6) \quad \begin{bmatrix} y(x_{11}) & y(x_{12}) & \dots & y(x_{1n}) & 1 \\ y(x_{21}) & y(x_{22}) & \dots & y(x_{2n}) & 1 \\ \dots & \dots & \dots & \dots & \dots \\ y(x_{n1}) & y(x_{n2}) & \dots & y(x_{nn}) & 1 \\ 1 & 1 & 1 & 1 & 0 \end{bmatrix} \times \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \dots \\ \lambda_n \\ \mu \end{bmatrix} = \begin{bmatrix} y(x_{10}) \\ y(x_{20}) \\ \dots \\ y(x_{n0}) \\ 1 \end{bmatrix},$$

где $y(x_{11}), y(x_{12}), \dots, y(x_{nn})$ – значения вариограмм для рассчитанных расстояний между узловыми точками; $\lambda_1, \lambda_2, \dots, \lambda_n$ – весовые коэффициенты, которые необходимо найти; $y(x_{10}), y(x_{20}), \dots, y(x_{n0})$ – значения вариограмм для рассчитанных расстояний между узловыми точками и искомой точкой; μ – множитель Лагранжа.

В результате решения системы уравнений (5) определим весовые коэффициенты: $\lambda_1, \lambda_2, \dots, \lambda_n$. Подставляя весовые коэффициенты и значения узловых точек в выражение (4), определим глубину в искомой точке. По результатам обработки всех измеренных узловых точек получим значения глубин в промежуточных точках и тем самым уменьшим шаг сетки батиметрической карты.

Таким образом, применение алгоритма Кригинга позволило сформировать сетку глубин микрорельефа дна в районе выявления препятствия вперед смотрящим гидролокатором секторного обзора. Данные сетки глубин микрорельефа позволяют системе управления АНПА сформировать полную информацию о форме

препятствия и по алгоритмам системы управления АНПА [7, 8] выбрать наиболее энергоэкономичный путь обхода препятствия.

В данном методе обнаружение фрагмента препятствия ГСО можно считать получением первичной информации о том, что впереди по курсу движения есть препятствие. По этим данным с учетом информации о текущем положении АНПА появляется возможность идентификации своего взаимного положения относительно препятствия. Одновременно в системе управления АНПА с использованием алгоритма Кригинга происходит формирование модели рельефа всего препятствия. Сформированная таким образом вторичная информация берется за основу для определения энергоэкономичного пути обхода препятствия. И далее осуществляется движение АНПА по сформированному пути обхода препятствия. ГСО в этом случае играет вспомогательную роль, работая исключительно в целях безопасности, чтобы АНПА в процессе обхода не врезалось в выступ и т.п.

Для выбора энергоэкономичного пути проанализируем особенности траектории движения АНПА при переходе с одного участка пути обхода на другой и оценим возможность движения АНПА в энергоэкономичном режиме.

4. Определение энергоэкономичного пути

После получения более полной информации о геометрической форме и размерах препятствия с данными узловых точек осуществим выбор наиболее энергоэкономичного пути обхода препятствия.

Для определения пути обхода сложного препятствия измельченную сетку глубин представим в виде взвешенного ориентированного графа G , узлами Y_i которого являются узлы измельченной сетки батиметрической карты, а ребрами $S_{i,j}$ – пространственные расстояния между узлами. Каждый i -й узел характеризуется координатами (x_i, y_i) своего положения в горизонтальной плоскости и глубиной h_i в вертикальной плоскости.

Назначим ребрам веса, равные расстояниям между узлами трехмерной сетки (широта, долгота, глубина). Такой вариант позволяет формализовать задачу определения наиболее энерго-

экономичного пути обхода препятствия как двумерную задачу поиска кратчайшего пути в графе.

Путь L в графе G – это последовательность узлов $L = (Y_1, \dots, Y_m, \dots, Y_M) \in Y$: Y_m смежен с Y_{m+1} для $1 \leq m \leq M$, M – количество узлов. Такой путь L является путем из узла Y_1 в Y_M , где Y_1 – положение АНПА в момент начала обхода препятствия, Y_M – целевая точка на маршрутной траектории после обхода препятствия.

Обозначим весовые функции $f_j: T \rightarrow R$, отображающие на ребра их веса, и весовые функции $f_m: Y \rightarrow R$, отображающие на узлы их веса. Весами являются затраты энергии. Задача сводится к определению в графе G такого пути L^* движения АНПА в обход препятствия, который обеспечивал бы минимизацию затрат энергоресурса АНПА.

Остановимся подробнее на расчете затрат электроэнергии на каждом из участков пути.

Расход энергоресурса определяется скоростью движения и используемыми исполнительными механизмами: рулями, подруливающими устройствами, уравнильно-дифферентной системой. Наиболее энергоэкономичным является движение АНПА в экономичном режиме с использованием рулей и без использования энергоемких подруливающих устройств и уравнильно-дифферентной системы. Опишем путь АНПА моделью Номото [18]: каждый участок МЗ охарактеризуем расходом электроэнергии на установившееся движение и на движение в ходе переходного процесса при маневрировании по курсу, скорости и глубине:

$$(7) \quad E_r = E_{const} + E_{var},$$

где E_{const} и E_{var} – затраты энергии на устойчивое движение АНПА и на переходные процессы соответственно. Затраты энергии в ходе движения определим по результатам действия на АНПА гидродинамических сил, сил плавучести и сил, возникающих в результате работы исполнительных механизмов. Для этого представим динамическое состояние АНПА в процессе движения в морской среде в виде системы дифференциальных уравнений вида

$$(8) \quad \frac{dv_i}{dt}(M + M_{\text{пр}}) = \sum F_i,$$

$$(9) \quad \frac{d\omega_i}{dt}(J + J_{\text{пр}}) = \sum M_i,$$

где i – индикатор проекции на оси связанной с АНПА системы координат $Oxyz$, $i = x, y, z$; M – масса АНПА, кг; $M_{\text{пр}}$ – присоединенные массы воды, кг; $\sum F_i$ – суммы проекций всех сил, действующих на объект, на оси связанной системы координат, Н; $\sum M_i$ – суммы проекций всех моментов инерции, действующих на объект, на оси связанной системы координат, Н·м; v_i – проекции (составляющие) вектора скорости на оси связанной системы координат, м/с; ω_i – проекции вектора угловой скорости аппарата, с⁻¹.

По результатам решения системы уравнений (8)–(9) определяем длину пройденного пути.

Формирование взвешенного графа позволяет определить кратчайший путь в смысле минимизации энергозатрат в соответствии с (1).

Существует большое количество всевозможных алгоритмов поиска кратчайшего пути в графе: алгоритм Дейкстры, алгоритм Беллмана – Форда, алгоритм поиска A^* , алгоритм Флойда – Уоршелла, алгоритм Джонсона, алгоритм Ли (волновой алгоритм), алгоритм поиска кратчайшего пути на основе алгоритма Килдала. Однако практически все перечисленные алгоритмы являются эвристическими, в связи с чем для поиска кратчайшего пути предлагается использовать точные алгоритмы, а именно – путем перебора всех возможных вариантов пути от начальной точки Y_1 до конечной Y_M .

Для этого можно использовать, например, широко распространенные методы обхода графа: поиск в глубину (англ. Depth-first search, DFS) [12, 14, 19] или поиск в ширину (Breadth-first search).

Стратегия поиска в глубину (Iterative deepening) состоит в том, чтобы идти «вглубь» графа, насколько это возможно. Алгоритм поиска описывается рекурсивно: перебираем все исходящие из рассматриваемой вершины ребра. Если ребро ведет

в вершину, которая не была рассмотрена ранее, то запускаем алгоритм от этой нерассмотренной вершины, а после возвращаемся и продолжаем перебирать ребра. Возврат происходит в том случае, если в рассматриваемой вершине не осталось ребер, которые ведут в нерассмотренную вершину. Если после завершения алгоритма не все вершины были рассмотрены, то необходимо запустить алгоритм от одной из нерассмотренных вершин.

Стратегия поиска в ширину заключается в систематическом обходе всех ребер $G = (V, E)$ для «открытия» всех вершин, достижимых из S , вычисляя при этом расстояние (минимальное количество ребер) от S до каждой достижимой из S вершины.

В соответствии с поиском в глубину формируем пути с ребрами, альтернативными последнему l_n ребру. В соответствии с алгоритмом требуется просчитать все такие пути. Однако для сокращения вычислительных ресурсов и времени решения задачи выбора пути обхода препятствия усовершенствуем алгоритм поиска в глубину.

Для этого зададимся начальным значением длины пути, пройдя одну из веток графа от Y_1 до Y_M . Запомним это значение и значения ребер $l_1, l_2, \dots, l_n$, через которые проходит этот путь (через n вершин). Если из всех ребер последней глубины найдутся такие, вес которых превышает уже рассмотренное, то такие ребра отсекаем из дальнейшего рассмотрения.

Точно так же поступим на каждой из глубин по мере вычисления протяженности пути. В том случае, если текущее значение пути оказалось меньше посчитанного на предыдущих шагах алгоритма, этот путь берется за основу и поиск кратчайшего пути продолжается.

Кроме того, при оценке протяженности пути по мере движения АНПА от узла к узлу предлагается проводить постоянное сравнение вычисляемой в каждом узле протяженности текущего пути с протяженностью рассчитанных ранее путей. И если оказалось, что протяженность рассматриваемого пути превысила протяженность сформированного ранее пути, то рассмотрение текущего пути прекращается и происходит отсечение этого пути. Отсечение неперспективных вариантов позволяет эконо-

мить время и вычислительные ресурсы по сравнению с полным перебором всех вариантов путей обхода препятствий.

Для тестирования предложенного алгоритма управления АНПА при обходе сложного препятствия по пространственной траектории разработана математическая модель и ее программная реализация, с использованием которой проведены численные эксперименты.

5. Результаты численных экспериментов

Оценим выигрыш предложенного решения по сравнению с приведенными в литературе рекомендациями обходить сбоку или сверху при условии, что сверху его можно обойти.

Оценим выигрыш на примере сложного препятствия, представляющего собой комбинацию двух фигур: в основании препятствия – пологая протяженная фигура, в верхней части препятствия – пикообразная фигура. Профиль препятствия представлен на рис. 3 слева. Он в основном совпадет с профилем пологого протяженного препятствия, представленного на рис. 1, с той лишь разницей, что в верхней части препятствия появился резкий вертикальный выступ: на рис. 3 справа пунктирной линией показан профиль пологого протяженного препятствия, сплошной линией – профиль комбинированного препятствия. Рассматриваемое препятствие можно трактовать как сложное, и очевидного ответа на вопрос о направлении его обхода – сверху или сбоку – в данном случае нет.

Рассмотрим три варианта обхода препятствия: первый – рекомендованный в литературе [2, 4] сверху, второй – сбоку, выбирая наугад направление, третий – комбинированный по предложенному в работе методу с использованием батиметрической и физической карт. В этом случае для построения сетки микро-рельефа используется грубая сетка мезорельефа и применяется алгоритм Кригинга. По результатам определения формы всего препятствия АНПА методом перебора с отсечением определяется энергоэкономичный пространственный путь обхода препятствия.

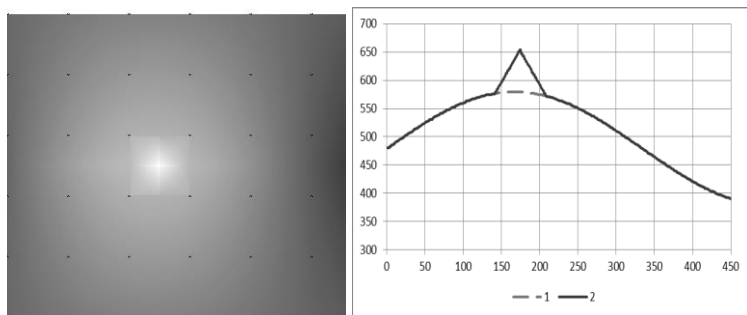


Рис.3. План и профиль рельефа комбинированного сложного препятствия

На рис. 4 приведен вид сверху на рассмотренные варианты обхода препятствия: сбоку без изменения глубины, сверху без отклонения от генерального направления МЗ и комбинированный вариант.

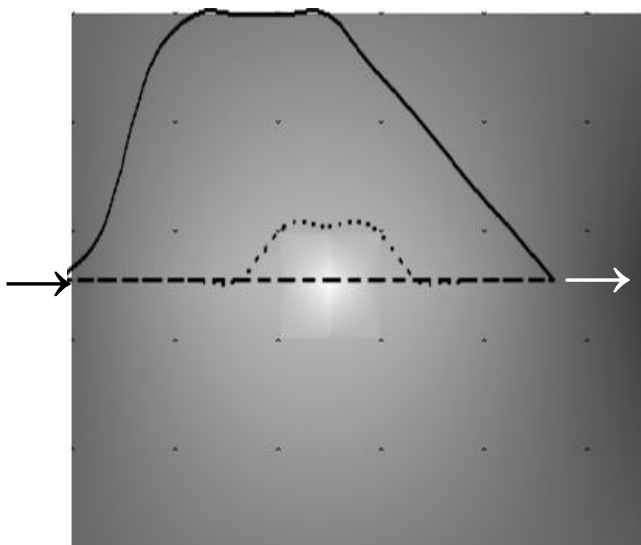


Рис.4. Вид сверху на траектории обхода препятствия: обход сбоку (сплошная линия), обход сверху (пунктирная линия) и обход комбинированным вариантом (точечная линия)

Таким оказался путь, сочетающий последовательность трех участков: первый связан с незначительным всплытием вверх, второй – с движением по окружности вокруг пикообразного выступа в верхней части препятствия и третий – заглубление до первоначальной глубины движения АНПА.

По результатам моделирования [6] оказалось, что путь обхода сверху и путь обхода сбоку существенно энергозатратнее, чем путь, определенный с применением разработанного метода обхода: длина пути обхода препятствия сверху составила 523 м, обхода сбоку – 563 м, а при использовании предложенного метода – 447 м, что соответственно на 15% и 20 % короче традиционных вариантов обхода. Поскольку траектория движения не требовала поворотов на прямые и острые углы, то выбранный по методу путь может быть выполнен с экономичной скоростью, что означает, что этот путь оказался более энергоэкономичным.

Таким образом, пространственный обход комбинированного сложного препятствия привел к более энергоэкономичному пути по сравнению с обходом этого же препятствия сбоку или сверху, что подтверждает правильность предложенного решения.

6. Заключение

Для выбора пути обхода сложного препятствия разработан метод, основанный на данных батиметрических и физических карт Мирового океана. Определены условия применимости батиметрических и физических карт с мезорельефом для формирования пути обхода. Для перехода от мезорельефа к микрорельефу предложено использовать алгоритм Кригинга; на примере показана его работа. Для определения наиболее энергоэкономичного пути обхода сложных препятствий предложено рассматривать сетку глубин в виде ориентированного взвешенного графа. Определены условия соответствия кратчайшего пути на графе наиболее энергоемкому пути. Разработан алгоритм перебора вариантов пути обхода с отсечением, базирующийся на алгоритме «поиска на графе в глубину». На примере обхода

сложного препятствия показано преимущество использования предложенного метода выбора варианта обхода препятствия по сравнению с традиционно предлагаемыми обходами сверху или сбоку препятствия. Приведен пример использования разработанного метода для определения наиболее энергоэкономичного пространственного пути обхода сложного препятствия. Результаты проведенных численных экспериментов подтвердили правильность предложенного решения: сокращение пути для рассмотренного варианта препятствия составило 15–20%, что позволяет реализовать предложенный метод выбора энергоэкономичного пространственного пути обхода сложного препятствия в системе управления АНПА.

Литература

1. АГАРКОВ С.А., ПАШЕНЦЕВ С.В. *Параметрическая идентификация обобщенной модели Номото с помощью аппарата вариационного исчисления* // Вестник МГТУ. – Т. 18. – №1. – 2015. – С. 7–11.
2. БЫКОВА В.С., МАШОШИН А.И., ПАШКЕВИЧ И.В. *Алгоритм обеспечения безопасности плавания автономного необитаемого подводного аппарата* // Гирскопия и навигация. – 2021. – Т. 29, №1(112). – С. 97–110.
3. ЗАВЬЯЛОВ В.В., КЛЮЕВА С.Ф., ЛАБЮК Ф.И. *Анализ точности построения и использования цифровой модели дна в задачах навигации по полю глубин* // Transport Business in Russia. – №6. – 2015. – С. 211–213.
4. ИНЗАРЦЕВ А.В., БАГНИЦКИЙ А.В. *Алгоритмы обхода локальных донных объектов для автономного подводного робота* // Шестая Всерос. науч.- техн. конф. «Технические проблемы освоения мирового океана» (ТПМО-6). Владивосток. – 2015. – С. 450–454.
5. МАРТЫНОВА Л.А. *Решение задачи подводного наблюдения в условиях применения интеллектуальных помех* // Информационно-управляющие системы. – 2018. – №1(92). – С. 31–41.
6. МАРТЫНОВА Л.А., ГРИНЕНКОВ А.В., ПРОНИН А.О. и др. *Имитационное моделирование функционирования мультиагентной системы управления автономного необитаемого подводного аппарата* // В сб.: «Имитационное моделирование. Теория и практика. Восьмая Всероссийская научно-практическая конференция

- по имитационному моделированию и его применению в науке и промышленности». – 2017. – С. 474–479.
7. МАРТЫНОВА Л.А., КИСЕЛЕВ Н.К., МЫСЛИВЫЙ А.А. *Метод выбора архитектуры мультиагентной системы управления автономного необитаемого подводного аппарата* // Информационно-управляющие системы. – 2020. – №4(107). – С. 31–41.
 8. МАРТЫНОВА Л.А., МАШОШИН А.И., ПАШКЕВИЧ И.В. и др. *Алгоритмы, реализуемые интегрированной системой управления АНПА* // Известия ЮФУ. Технические науки. – 2015. – №1(162). – С. 50–58.
 9. ПШИХОПОВ В.Х. *Интеллектуальное планирование траекторий подвижных объектов в средах с препятствиями*. – М.: Физматлит, 2015. – 300 с.
 10. ALCARAS E., PARENTE C., VALLARIO A. *From electronic navigational chart data to sea-bottom models: Kriging approaches for the Bay of Pozzuoli* // Acta IMEKO. – Vol. 10, No. 4, Article 9, December 2021, ID: IMEKO-ACTA-10 (2021)-04-09/.
 11. AMOROSO P.P., FALCHI U., FIGLIOMENI F.G. et al. *The Influence of Interpolation Methods and point density on the Accuracy of a Bathymetric Model* // IEEE Int. Workshop on Metrology for the Sea; Learning to Measure Sea Health Parameters (MetroSea-2023), 04–06 October 2023. – DOI: 10.1109/MetroSea58055.2023.10317127.
 12. BIRX A., DISSER Y., HOPP A.V. et al. *An improved lower bound for competitive graph exploration* // Theoretical Computer Science. – May 2021. – No. 868. – P. 65–86. – arXiv:2002.10958. – DOI: 10.1016/j.tcs.2021.04.003.
 13. BRAHMI H., AMMAR B., ALIMI A.M. *Intelligent path planning algorithm for autonomous robot based on recurrent neural networks* // Int. Conf. on Advanced Logistics and Transport, 2013. – P. 199–204. – DOI: 10.1109/ICAdLT.2013.6568459.
 14. BRANDT S., FOERSTER K.-T., MAURER J. et al. *Online graph exploration on a restricted graph class: Optimal solutions for tadpole graphs* // Theoretical Computer Science. – November 2020. – No. 839. – P. 176–185. – arXiv:1903.00581. – DOI: 10.1016/j.tcs.2020.06.007.
 15. CHANG J.-H., HART D.R., MUNROE D.M. et al. *Bias Correction of Ocean Bottom Temperature and Salinity Simulations From a Regional Circulation Model Using Regression Kriging* // Journal of Geophysical Research: Oceans. – 2021. – DOI: 10.1029/2020JC017140.

16. EVANS J., PATRÓN P., SMITH B. et al. *Design and evaluation of a reactive and deliberative collision avoidance and escape architecture for autonomous robots* // Autonomous Robots. – 2008. – Vol. 24, No 3. – P. 247–266.
17. GALARZA C., MASMITJA I., PRAT J. et al. *Design of obstacle detection and avoidance system for Guanay II AUV* // Appl. Sci. – 2020. – Vol. 10. – P. 32–37.
18. GOKARN R.P. *A Study of Ship Manoeuvrability*. – Springer Nature Singapore Pte Ltd., 2024. – P. 105–115.
19. IYANDA J.A. *Comparative Analysis of Breadth First Search (BFS) and Depth First Search (DFS) Algorithms*. – May 2023. Preprint. – URL: https://www.researchgate.net/publication/370751322_Title_A_Comparative_Analysis_of_Breadth_First_Search_BFS_and_Depth_First_Search_DFS_Algorithms (дата обращения: 17.06.2024).
20. LEE S.Y., MALLICK B. *Bayesian Hierarchical Modeling: Application Towards Production Results in the Eagle Ford Shale of South Texas* // Sankhya B 84. – 2022. – Vol. 84. – P. 1–43. – DOI: 10.1007/s13571-020-00245-8.
21. LIN C., WANG H., YUAN J. et al. *An improved recurrent neural network for unmanned underwater vehicle online obstacle avoidance* // IEEEJ. Ocean. Eng. – 2019. – Vol. 44. – P. 120–133.
22. LIU S., WEI Y., GAO Y. *3D path planning for AUV using fuzzy logic* // Computer Science and Information Processing (CSIP), 2012. – URL: <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6308925> (дата обращения: 17.06.2024).
23. RANFTL S., VON DER LINDEN W. *Bayesian Surrogate Analysis and Uncertainty Propagation*. // Physical Sciences Forum. – 2021. – No. 3(1):6. – arXiv:2101.04038. – DOI: 10.3390/psf2021003006. ISSN 2673-9984.
24. SAMI A., AYMAN M.M., BRISHA M. *Analysis and Simulation of 3D Trajectory with Obstacle Avoidance of an Autonomous Underwater Vehicle for Optimum Performance* // Int. Journal of Computer Science and Network Security. – 2012. – Vol. 12, No. 3. – P. 43–50.
25. *The General Bathymetric Chart of the Oceans (GEBCO)*. – URL: <https://www.gebco.net/> (дата обращения: 17.06.2024).
26. VALENZA E. *Blender 2.6 Cycles: Materials and Textures Cookbook: Over 40 Recipes to Help You Create Stunning Materials and Textures Using the Cycles Rendering Engine With Blender*. – Birmingham, UK: Packt Pub Ltd, 2013. – 265 p.

METHOD CHOOSING AN ENERGY ECONOMICAL WAY TO BYPASS A COMPLEX OBSTACLE BY AUTONOMOUS

Lyubov Martynova, JSC Concern Central Research Institute Elektropribor, Saint Petersburg, Doctor of Technical Sciences, Senior Researcher (martynowa999@bk).

Alexey Pavlov, JSC Concern Central Research Institute Elektropribor, Saint Petersburg, Ph.D., Lead Software Engineer (ialex.science@gmail.com).

Abstract: A method based on bathymetric and physical maps of the World Ocean has been developed to select a bypass path around a complex obstacle. It has been proposed to use the Kriging algorithm to go from mesorelief to minirelief. It has been proposed to consider a depth grid in the form of an oriented weighted graph to determine the most energy-efficient bypass path around complex obstacles. The conditions for the shortest path on the graph to be the most energy efficient way have been determined. An algorithm for enumerating bypass path options with pruning has been developed, based on the "depth-first graph search" algorithm. The advantage of using the proposed method for selecting a bypass option over the traditionally proposed bypasses from above or from the side of the obstacle has been shown in an example of bypassing a complex obstacle. The results of the conducted numerical experiments confirmed the correctness of the proposed solution: the reduction of the path for the considered variant of the obstacle was 15-20%, which makes it possible to implement the proposed method of selecting an energy-efficient spatial path to bypass a complex obstacle in the control system of an autonomous underwater vehicle.

Keywords: autonomous underwater vehicle, bottom relief, bathymetric map, Kriging algorithm, search for the shortest path on a graph.

УДК 004.896

ББК 32.965.07

DOI: 10.25728/ubs.2024.111.11

*Статья представлена к публикации
членом редакционной коллегии А.И. Алчиновым.*

Поступила в редакцию 19.05.2024.

Опубликована 30.09.2024.

РАЗРАБОТКА МЕТОДИКИ И СРЕДСТВ ПРОВЕДЕНИЯ РЕСУРСНЫХ ИСПЫТАНИЙ ДЛЯ ПОСТРОЕНИЯ СИСТЕМЫ РАННЕЙ ДИАГНОСТИКИ ТЕХНИЧЕСКОГО СОСТОЯНИЯ ЭЛЕКТРОМЕХАНИЧЕСКОГО РУЛЕВОГО ПРИВОДА БЕСПИЛОТНОГО ВОЗДУШНОГО СУДНА¹

**Баженов С. Г.^{2аb}, Вересников Г. С.^{3а}, Голев А. В.^{4а},
Гончаренко В. И.^{5а}, Ерофеев Е. В.^{6аb}, Лазурин Г. А.^{7б},
Скрябин А. В.^{8аb}, Тимофеева А. Д.^{9аb}, Феденюк В. А.^{10б}**

*(^а ФГБУН Институт проблем управления
им. В.А. Трапезникова РАН, Москва, ^б ФАУ «ЦАГИ», Жуковский)*

Работа посвящена созданию технологии получения и накопления экспериментальных данных при ресурсных испытаниях электромеханического рулевого привода (ЭМРП), необходимых для разработки алгоритмов интеллектуального анализа, обеспечивающих раннюю диагностику ЭМРП для проведения технического обслуживания летательного аппарата по состоянию. Разработаны программа и методика ресурсных испытаний, позволяющие ускорить процессы износа в компонентах ЭМРП и получить данные о функционировании при воспроизведении циклограмм нагружения, соответствующих наземной отработке в ходе предполетной подготовки («активный эксперимент») и в ходе полета летательного аппарата («пассивный эксперимент»). Создан прототип системы ранней диагностики, включающий в себя стендовую установку, обеспечивающую упругое или весовое механическое нагружение ЭМРП беспилотного воздушного судна (БВС), и программно-аппаратный комплекс сбора данных, обеспечивающий измерение и регистрацию физических величин, связанных с рабочими процессами ЭМРП в различных условиях работы и различных технических состояниях, обусловленных износом. В качестве диагностических сигналов предполагается использовать физические величины электрической, механической, тепловой и виброакустической природы. Приведены результаты

¹ Исследование частично выполнено за счет гранта Российского научного фонда (проект № 23-19-00464).

² Сергей Георгиевич Баженов, д.т.н., доцент, с.н.с., начальник отделения (sergey.bazhenov@tsagi.ru).

³ Георгий Сергеевич Вересников, д.т.н., в.н.с., зав. лабораторией (veresnikov@mail.ru).

⁴ Артем Владимирович Голев, н.с. (oiw23@mail.ru).

⁵ Владимир Иванович Гончаренко, д.т.н., доцент, вед. инженер (vladimirgonch@mail.ru).

⁶ Евгений Валерьевич Ерофеев, н.с., начальник установки (evgeniy.erofeev@tsagi.ru).

⁷ Григорий Андреевич Лазурин, н.с. (flight15@tsagi.ru).

⁸ Алексей Валерьевич Скрябин, к.т.н., н.с., вед. инженер (aleksey.skryabin@tsagi.ru).

⁹ Анастасия Дмитриевна Тимофеева, м.н.с., инженер (anastasiya.timofeeva@tsagi.ru).

¹⁰ Василий Андреевич Феденюк, инженер (flight15@tsagi.ru).

тестовых ресурсных испытаний образцов ЭМРП БВС, которые позволили уточнить программу испытаний в части действующих нагрузок.

Ключевые слова: электромеханический привод, беспилотное воздушное судно, диагностика, ресурсные испытания.

1. Введение

В связи с повышением степени электрификации бортовых систем при создании «более электрического самолета» (БЭС) необходимо обеспечить их безопасную эксплуатацию и эффективное техническое обслуживание. При решении задач стабилизации и управления пространственным положением летательного аппарата [6, 11] одной из наиболее важных и сложных систем с электрическим энергопитанием является система рулевых приводов, включающая в себя электромеханические рулевые приводы (ЭМРП) для отклонения аэродинамических поверхностей. Для эффективной эксплуатации системы рулевых приводов, сохранение функциональности которой критически важно для обеспечения безопасного полета, целесообразно создание технологии ранней диагностики технического состояния для выявления скрытых ухудшений (деградаций) характеристик до момента наступления отказа ЭМРП [10, 13].

Отечественные ЭМРП для БЭС в настоящее время существуют в виде прототипов и имеют уровень готовности технологии 5–6 [2, 3, 20]. По состоянию на сегодняшний день системы рулевых приводов с электрическим энергопитанием широко используются на беспилотных воздушных судах (БВС) в классах «нано», «микро», «мини», «ближнего и среднего радиусов действия» по классификации UVS International, поэтому создание и отработка технологии ранней диагностики возможны для существующих серийных ЭМРП в этих классах объектов.

Разработка современных систем ранней диагностики сложных технических систем предполагает использование алгоритмов интеллектуального анализа данных [14, 15, 17], которые позволяют решить задачу классификации технического состояния [8] и прогнозировать отказные ситуации. Системы ранней диагностики, построенные с использованием алгоритмов интеллектуального анализа данных [9, 18, 21], требуют наличия выборки

данных, характеризующих техническое состояние ЭМРП [19]. Такая выборка может быть получена при проведении ресурсных испытаний и записи сигналов датчиков физических величин, связанных с работой ЭМРП [13, 16].

Для подготовки к ресурсным испытаниям партии однотипных ЭМРП и накопления статистических данных, характеризующих изменение технического состояния, была разработана программа и методика отработки заданных циклограмм и создана экспериментальная установка, обеспечивающая механическое нагружение и запись диагностических сигналов, которые используются в алгоритмах интеллектуального анализа для формирования признаков развития неисправностей.

2. Программа ресурсных испытаний

Для оценки технического состояния ЭМРП с использованием алгоритмов интеллектуального анализа в ходе ресурсных испытаний предлагается аккумулировать данные [7] двух типов:

- статические и динамические характеристики, формируемые при отработке контрольных осциллограмм («активный эксперимент»);
- статические и динамические характеристики, формируемые при воспроизведении циклограмм, приближенных к нормальному режиму функционирования при типовом полете («пассивный эксперимент»).

Расходование ресурса ЭМРП предлагается проводить при отработке режима малой постоянной скорости выходного звена и ограничения действующей нагрузки.

Ресурсные испытания проводятся сериями, состоящими из последовательности сеансов, длительность каждого из которых не превышает 8 часов. Типовая программа сеанса ресурсных испытаний показана в таблице 1.

Номинальное и пониженное напряжения питания ЭМРП соответствуют уровням в бортовой сети электроснабжения от аккумуляторной батареи в заряженном и разряженном состояниях.

Таблица 1. Программа сеанса ресурсных испытаний

Тип осциллограммы	Нагрузка	Напряжение питания	Форма сигнала	Амплитуда, % от максимальной	Частота сигнала, Гц
1. Контрольные осциллограммы (автономная отработка)					
Статическое позиционирование	Без нагрузки	Номинальное	Треугольный	1	0,01
ЛАФЧХ			Синусоидальный	1; 3	
Переходные процессы			Меандр Треугольный	Максимальная	0,5 / 1
ЛАФЧХ		Пониженное	Синусоидальный	1; 3	0,1...1,0 / 1...15
Переходные процессы			Меандр Треугольный	Максимальная	0,5 / 1
2. Нормальный режим функционирования (полунатурное моделирование)					
Вираз	Пружина	Номинальное	Из модели полета		
Горизонтальный полет – набор высоты					
Воздействие ветровых порывов					
Вираз		Пониженное			
Горизонтальный полет – набор высоты					
Воздействие ветровых порывов					
3. Форсированный режим функционирования (ускорение износа)					
Износ	Грузы	Номинальное	Треугольный	Максимальная	0,2
Контрольная запись					

3. Методика ресурсных испытаний

Разработанная методика ресурсных испытаний используется для исследования процессов износа и формирования массива данных, содержащего информацию о функционировании ЭМРП и развитии деградаций в его компонентах.

Оценка технических состояний ЭМРП может проводиться при сравнении статических и динамических характеристик, полученных в результате обработки контрольных осциллограмм (переходных процессов при ступенчатом, треугольном или синусоидальном управляющем сигнале различной амплитуды и частоты). Выполнение подобной оценки возможно как в ручном режиме с привлечением эксперта, так и автоматизированно с использованием специализированных программно-аппаратных комплексов [12].

Основными параметрами для оценки технического состояния ЭМРП по статическим и динамическим характеристикам являются:

- характер и повторяемость (в рамках одного сеанса испытаний) статической характеристики позиционирования выходного звена, позволяющие оценить люфт, трение и чувствительность ЭМРП;
- максимальная скорость, определяемая по переходному процессу при ступенчатом или треугольном управляющем сигнале на участке с постоянной скоростью;
- амплитудные и фазовые искажения позиционного сигнала в рабочем диапазоне частот;
- характер переходного процесса выходного сигнала при обработке управляющего сигнала ступенчатой формы.

Статическая характеристика позиционирования – зависимость положения выходного звена от величины управляющего сигнала – может быть получена при воспроизведении управляющего сигнала треугольной формы и малой частоты (не более 0,01 Гц) при обработке «в малом» (амплитуда не более 1% максимального хода ЭМРП) и «в большом» (в окрестности 100% максимальной амплитуды). Для идентификации величины люфта особое значение представляют участки, на которых изменяется направление движения выходного звена [5]. Процесс износа контактирующих пар редуктора сопровождается отделением металлической стружки, которая способствует возникновению трещин, их разрастанию и в конечном счете разрушению или заклиниванию редуктора. Деградации такого рода могут также наблюдаться на статической характеристике позиционирования.

Для получения амплитудных и фазовых искажений применяется метод определения амплитудных и фазовых искажений, при возбуждении исследуемой системы входным гармоническим воздействием $u(t) = u_0 \cos(\omega t)$, $t \geq 0$, и вычислении интегралов Фурье от ее выходного сигнала $y(t)$:

$$(1) \quad \alpha(\omega) = \frac{2}{NT_0} \int_{t_{st}}^{t_{st} + NT_0} y(t) \cos \omega t dt,$$

$$(2) \quad \beta(\omega) = \frac{2}{NT_0} \int_{t_{st}}^{t_{st} + NT_0} y(t) \sin \omega t dt,$$

где $T_0 = 2\pi/\omega$ – период сигнала возбуждения; N – число периодов усреднения; t_{st} – время между началом подачи сигнала возбуждения и началом измерения. Это время выбирается таким образом, чтобы к началу измерения практически затухло собственное движение системы. Значения амплитудной и фазовой характеристик системы на частоте ω – $A(\omega)$ и $\phi(\omega)$ вычисляются по формулам:

$$(3) \quad A(\omega) = \sqrt{\alpha^2(\omega) + \beta^2(\omega)} / u_0,$$

$$(4) \quad \phi(\omega) = -\arctg \left[\frac{\beta(\omega)}{\alpha(\omega)} \right].$$

Формулы (1)–(4) обеспечивают точное определение определения частотных характеристик линейной непрерывной системы, а для нелинейной непрерывной системы позволяют точно определить отношение первой гармоники выходного сигнала к амплитуде входного сигнала (коэффициент гармонической линеаризации).

При деградации редуктора изменяются значения развиваемой максимальной скорости и потребляемого тока, что может быть определено на участках постоянной скорости при воспроизведении переходного процесса при ступенчатом управляющем сигнале максимальной амплитуды. Переходный процесс при ступенчатом управляющем сигнале также позволяет оценить изменения динамических характеристик на участках разгона и торможения выходного звена.

Отработка контрольных циклограмм (автономная отработка) для исследования алгоритмов диагностики технического

состояния выполняется при отсутствии механической нагрузки, что позволяет оценить качество функционирования системы ЭМРП при отсутствии внешних факторов воздействия (напр. определить люфт, который при наличии нагрузки «выбирается»). При успешной реализации подобной технологии в дальнейшем предлагается внедрить ее в процедуру технического обслуживания при наземной отработке системы рулевых приводов в ходе «активного эксперимента».

Для оценки технического состояния ЭМРП в ходе полета предлагается при проведении полунатурного моделирования [4] обеспечить режим, приближенный к нормальному функционированию. Результаты отработки алгоритмов диагностики на данных ресурсных испытаний в подобном режиме могут в дальнейшем использоваться в бортовой системе БВС при проведении «пассивного эксперимента». Режим нормального функционирования характеризуется воспроизведением управляющего сигнала (как правило, рабочий ход выходного звена не превышает 10 %), формируемого системой управления при выполнении типовых режимов полета БВС. В связи с тем, что бортовые системы обладают ограниченными возможностями для регистрации сигналов датчиков (количество сигналов и частота сбора данных), то для оценки технического состояния ЭМРП целесообразно использовать энергетические параметры, объединяющие в себе информацию о преобразовании энергии в электромеханическом тракте: ток и напряжение питания, частота вращения и момент на валу. Данные параметры наиболее информативны при наличии механической нагрузки – шарнирного момента, возникающего на рулевой поверхности, находящейся в набегающем потоке, который определяется выражениями

$$(5) \quad M_{\text{ш}} = m_{\text{ш}} q S b_a ,$$

$$(6) \quad m_{\text{ш}} = m_{\text{ш}0} + m_{\text{ш}}^{\alpha} \alpha + m_{\text{ш}}^{\delta} \delta_{\text{руля}} ,$$

здесь $m_{\text{ш}}$ – коэффициент шарнирного момента органа управления; q – скоростной напор; S – габаритная площадь органа управления; b_a – средняя аэродинамическая хорда органа управления; $m_{\text{ш}}^{\alpha}$, $m_{\text{ш}}^{\delta}$ – коэффициенты шарнирных моментов; α , δ – углы атаки и отклонения рулевой поверхности.

Для имитации шарнирного момента в ходе ресурсных испытаний предлагается использовать пружину растяжения для воспроизведения линейной нагрузки, которая имитирует рост шарнирного момента при отклонении рулевой поверхности.

Для ускорения расходования ресурса в ходе испытаний предлагается воспроизводить процессы износа при недопущении выхода на предельные режимы по нагрузке и скорости с использованием известного метода ускоренного воспроизведения рабочих циклов, при котором работа, выполняемая ЭМРП в процессе испытаний, соответствует типовой полетной циклограмме, воспроизводимой на протяжении нормальной эксплуатации [1]. С этой целью в качестве управляющего сигнала использован периодический треугольный сигнал низкой частоты f_{Cmd} (0,2 Гц) и максимальной амплитуды φ_{Cmd} , обрабатываемой под действием нагрузки M_{Cmd} , составляющей 20% от максимального момента ЭМРП. В этом режиме механическая нагрузка воспроизводится с использованием грузов (весовая нагрузка), а инерционная сила, возникающая при изменении направления скорости движения вала привода, является незначительной при низкой частоте позиционного управляющего сигнала.

Для треугольного сигнала общая затраченная в ходе испытаний энергия и длительность испытаний на износ определяются выражениями

$$(7) \quad E_{\Sigma. \text{Исп. Треуг.}} = 2 \cdot \varphi_{Cmd} \cdot f_{Cmd} \cdot M_{Cmd} \cdot t_{\text{исп}},$$

$$(8) \quad t_{\text{исп.треуг}} = \frac{E_{\Sigma. \text{ЖЦ}}}{2 \cdot \varphi_{Cmd} \cdot f_{Cmd} \cdot M_{Cmd}};$$

Здесь $E_{\Sigma. \text{ЖЦ}}$ – общая энергия, затраченная ЭМРП в ходе жизненного цикла в составе системы управления летательного аппарата.

4. Прототип установки проведения ресурсных испытаний

Для проведения испытаний была разработана рабочая конструкторская документация на стендовую установку, схема которой показана на рис. 1. Функционирование установки позволяет проводить отработку ЭМРП в трех конфигурациях: без нагрузки,

под действием упругой нагрузки и под действием весовой нагрузки.

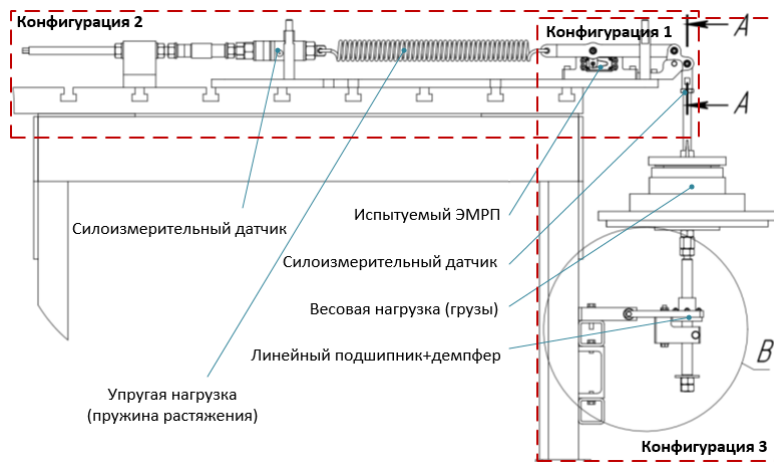


Рис. 1. Схема узлов крепления ЭМРП при проведении ресурсных испытаний

Таблица 2. Средства измерения диагностических сигналов

№	Сигнал	Источник сигнала	Ед. изм.	Погрешность, %
1	Заданное положение ЭМРП	ЭМРП	°	0,2
2	Текущее положение ЭМРП	ЭМРП	°	0,2
3	Линейная нагрузка (канал 1)	Тензометрический датчик силы	Н	0,5
4	Линейная нагрузка (канал 2)	Тензометрический датчик силы	Н	0,5
5	Напряжение питания	Источник питания	В	0,5
6	Потребляемый ток	Датчик Холла	А	1,5
7	Акустический шум	Измерительный микрофон	дБ	Сигнал/шум 70 дБ
8	Виброускорение	Трехосевой MEMS-акселерометр	м/с ²	0,5
9	Температура ЭМРП	Термистор	°С	0,5
10	Температура помещения	Термистор	°С	0,5
11	Термограмма	Тепловизор	°С	3

Для отработки прототипа системы диагностики в качестве диагностируемого объекта был выбран покупной ЭМРП, для которого были определены средства измерения, показанные в таблице 2. Схема подключения компонентов программно-аппаратного комплекса сбора данных показана на рис. 2.

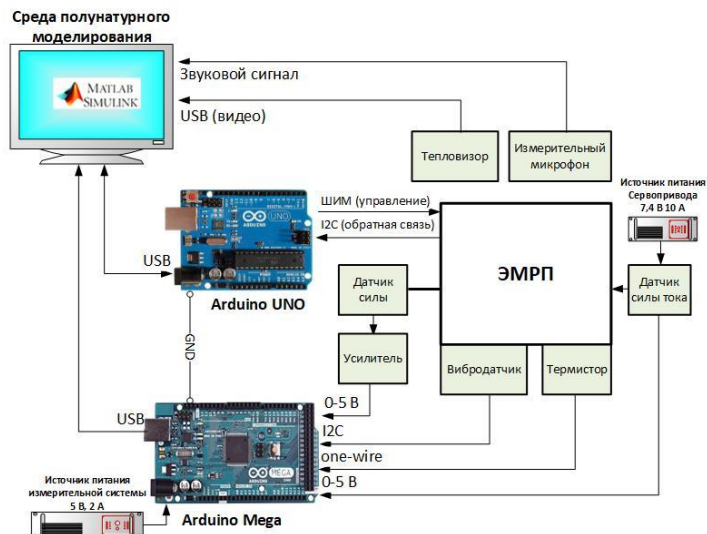


Рис. 2. Схема подключения программно-аппаратного комплекса сбора данных

Для автоматизации проведения ресурсных испытаний на базе среды моделирования MATLAB/Simulink разработан программно-аппаратный комплекс сбора данных, обеспечивающий измерение и регистрацию сигналов: позиционных, механических, электрических, виброакустических и тепловых. Разработанное специализированное программное обеспечение включает в себя:

- программно-математическую модель, реализованную в среде Simulink и включающую в себя s-функции обмена данными с двумя контроллерами Arduino (обмен данными с ЭМРП и сбор данных с внешних датчиков), блок опроса звуковой карты,

блоки генерации управляющего сигнала и обработки выходных сигналов и драйвер реального времени;

- программное обеспечение контроллера (прошивка) Arduino Mega 2560, обеспечивающее опрос силоизмерительных датчиков (АЦП), датчика тока (АЦП), акселерометра (I2C), термистора (one-wire) и передачу показаний в модель Simulink посредством s-функции, разработанной на базе библиотеки firmata;

- программное обеспечение контроллера (прошивка) Arduino UNO, обеспечивающее управление ЭМРП ШИМ-сигналом, опрос встроенного в ЭМРП позиционного датчика обратной связи (I2C) и обеспечивающее обмен данными с моделью Simulink посредством s-функции, разработанной на базе библиотеки firmata.

Для записи термограмм использовалось программное обеспечение IRCAM Thermal Viewer 4.4.0, разработанный на базе MATLAB Runtime 2022a.

5. Результаты тестовых испытаний и отказные ситуации

Созданная экспериментальная установка и испытуемый ЭМРП показаны на рис. 3. В соответствии с разработанной методикой на макете стендовой установки были проведены тестовые ресурсные испытания для различных случаев нагружения выходного вала привода: свободный вал (контрольные осциллограммы), упругая нагрузка (нормальный режим функционирования), весовая нагрузка (форсированный режим функционирования).

Примеры записанных осциллограмм при отработке трехугольных сигналов на форсированном режиме показаны на рис. 4. В этом режиме установка обеспечивает точность поддержания весовой нагрузки 5% номинального момента испытуемого ЭМРП, что обуславливается направлением движения (помогающая или препятствующая нагрузка) и вызывает потребление тока от 0,5 до 1,1 А. В таком режиме рост температуры достигает установившегося значения 60 °С (рис. 5). Таким образом, в нормальных условиях (при температуре окружающей среды 23 °С) форсированный режим обеспечивает возможность длительной

отработки ЭМРП под нагрузкой, нагрев достигает уровня, не превышающего допустимую температуру внешнего воздействия.

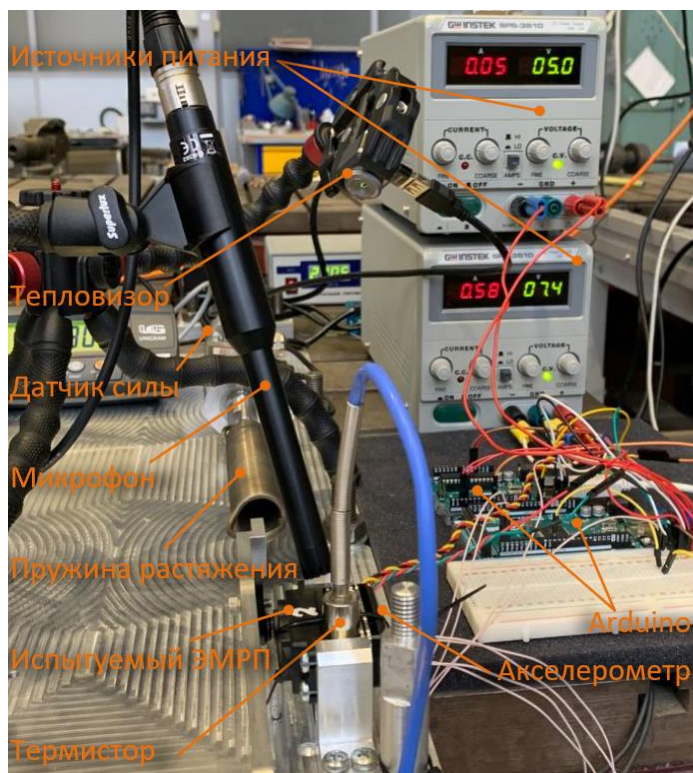


Рис. 3. Проведение испытаний ЭМРП БВС на макете установки

При тестовой отработке двух образцов ЭМРП БВС были исследованы предельные режимы их работы, в ходе которых наблюдались отказные ситуации электрической и механической природы, установлены последствия отказов на функционирование ЭМРП.

При длительном удержании постоянно действующей нагрузки (несколько минут), соответствующей моменту торможения, произошел электрический отказ силового транзистора электронного блока (рис. 6). В результате ЭМРП потерял управление, а выходное звено при отключенном питании и действии

внешнего момента стало поворачиваться с большим усилием, что свидетельствует о коротком замыкании обмоток электродвигателя через силовой транзистор и критическом повышении демпфирования, практически соответствующего состоянию заклинившего редуктора.

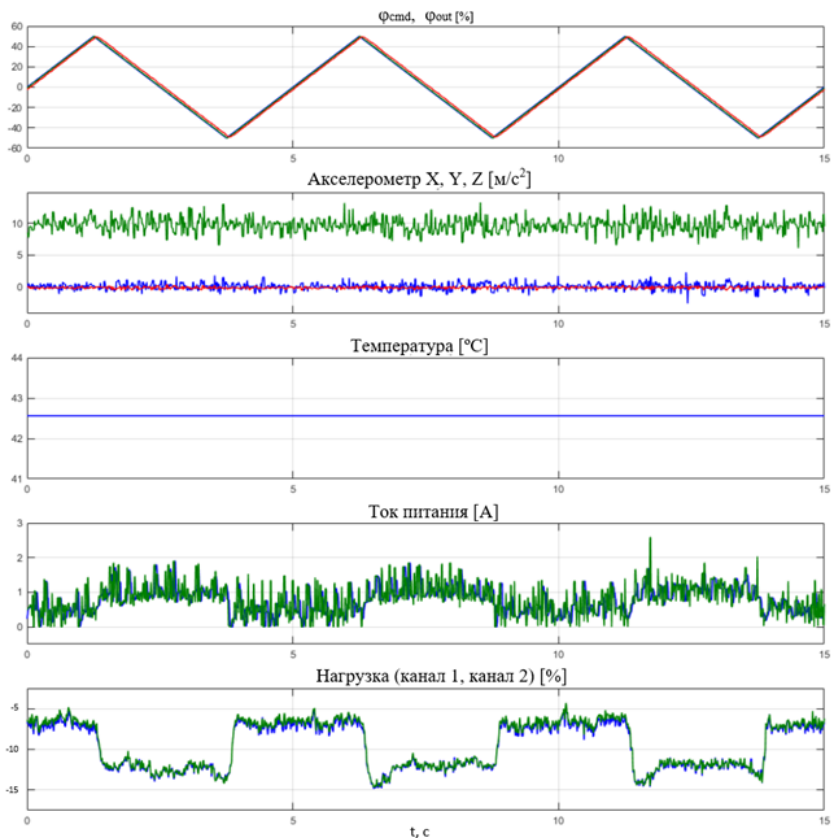


Рис. 4. Осциллограммы при отработке форсированного режима: позиционные сигналы, виброускорения по трём осям, температура, сила тока питания, нагрузка на валу

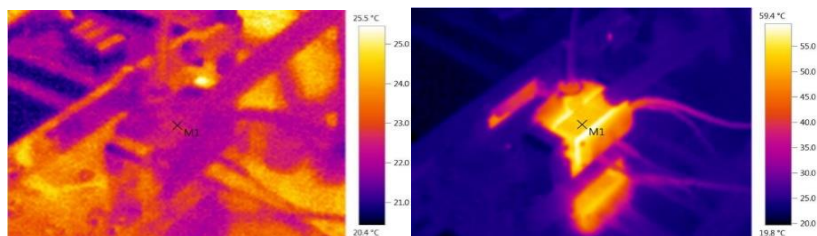


Рис. 5. Термограммы ЭМРП при форсированном режиме: до испытаний (слева) и через 2 часа после их начала (справа)



Рис. 6. Вид печатной платы ЭМРП с отказавшим блоком электроники при длительном статичном удержании момента торможения

При отработке управляющего сигнала типа «меандр» и под действием момента торможения ЭМРП не смог преодолеть препятствующую нагрузку, а при отклонении с помогающей нагрузкой вышел на заданное положение и потерял управление с последующей потерей способности воспроизводить управляющие сигналы и ростом потребляемого тока до 2 А. (рис. 7). При внешнем воздействии выходное звено стало свободно поворачиваться с подклиниваниями, таким образом произошло разрушение элементов прямоугольного редуктора с рассоединением кинематической связи выходного звена и электродвигателя.

Отказы ЭМРП позволили уточнить методику проведения испытаний в части ограничения скорости нарастания управляющего сигнала и действующей нагрузки.

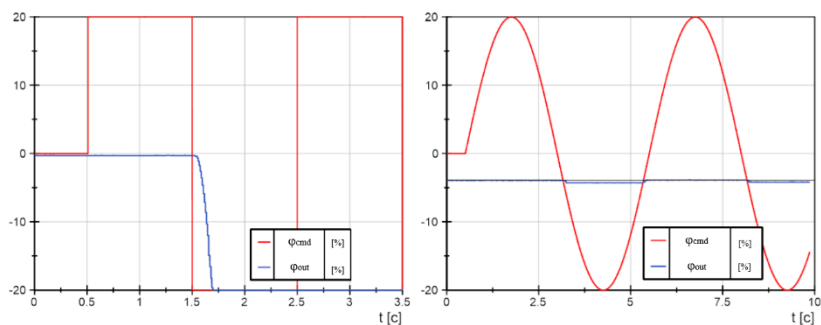


Рис. 7. Поломка редуктора при воспроизведении меандра под максимальным моментом (слева) и реакция ЭМРП при синусоидальном сигнале без нагрузки после поломки (справа)

6. Заключение

Разработанный прототип системы диагностики позволяет отработать в стендовых условиях часть технологии ранней диагностики системы рулевых приводов – процедуру сбора и накопления данных о работе ЭМРП в различных технических состояниях при развитии процессов износа в компонентах. Данные, полученные в ходе ресурсных испытаний, будут использованы для разработки и верификации алгоритмов интеллектуального анализа, которые позволяют как классифицировать техническое состояние отдельного агрегата, так и определить тренд изменения технического состояния и время до наступления отказа. Технология ранней диагностики, отработанная с использованием ЭМРП, применяемых на БВС, может применяться на пилотируемых летательных аппаратах для продления ресурса и организации технического обслуживания по состоянию электромеханических систем, широкое внедрение которых предполагается с развитием технологии «более электрического самолета» (БЭС).

При проведении тестовых ресурсных испытаний в отношении образцов ЭМРП БВС обеспечен автоматизированный сбор данных с датчиков, встроенных в привод, и дополнительных датчиков параметров среды (тепловых, механических и виброакустических). Наблюдавшиеся в ходе испытаний отказы являются следствием электрических и механических рабочих процессов,

протекающих на предельных режимах работы ЭМРП, и могут приводить к потере управления рулевой поверхностью, в том числе и при заклинивании или рассоединении выходного звена ЭМРП от электродвигателя.

Разработанные программа и методика ресурсных испытаний приближены к условиям натурной эксплуатации ЭМРП в составе летательного аппарата и учитывают сбор данных о работе ЭМРП как в ходе нормального функционирования во время полета (пассивный эксперимент), так и обработку типовых осциллограмм для оценки статических и динамических характеристик (активный эксперимент). Эти процессы сбора данных могут выполняться автоматически, не требуют участия специализированного технического персонала для накопления данных и решения задач диагностики системы ЭМРП в составе борта.

Литература

1. БЕССОЛОВА О.А., РАЙХЕР В.Л., УСТИНОВ А.С. *Расчет усталостной повреждаемости при циклическом и случайном нагружении с ненулевым средним значением* // Ученые записки ЦАГИ. – 1989. – Т. XX, №3. – С. 72–80.
2. ЕРМАКОВ С.А., КОНСТАНТИНОВ С.В., КУЗНЕЦОВ И.П. и др. *Структура и компоненты рулевых приводов систем управления перспективных самолетов с повышенным уровнем электрификации* // Общероссийский научно-технический журнал «Полет». – 2016. – №4. – С. 35–47.
3. ЕРОФЕЕВ Е.В., КУВШИНОВ В.М., СКРЯБИН А.В. и др. *Построение силовой системы управления самолетов с использованием рулевых приводов с электрическим силовым питанием для реализации концепции «более электрического самолета»* // Труды ЦАГИ. – 2019. – № 2785 «Динамика полета и системы управления летательных аппаратов. Сборник статей» / Под ред. В.Л. Суханова. – С. 49–70.
4. САМСОНОВИЧ С.Л., РОЖНИН Н.Б., ЛАРИН А.П. и др. *Результаты моделирования работы резервированных электро-механических приводов летательных аппаратов* // Известия ТулГУ. Технические науки. – 2021. – №11. – С. 376–385.

5. СЕРЕБРЕННЫЙ В.В., БОШЛЯКОВ А.А., ОГОРОДНИК А.И. *Математическая модель исполнительных модулей захватных устройств роботов* // Вестник БГТУ им. В.Г. Шухова. – 2019. – №6. – С. 123–135.
6. ANNAZ F., KALUARACHCHI M.M. *Progress in Redundant Electromechanical Actuators for Aerospace Applications* // Aerospace. – 2023. – Vol. 10(9): 787. – DOI: 10.3390/aerospace 10090787.
7. ARELLANO-ESPITIA F., GONZALEZ-ABREU A.D., PRIETO M.D. et al. *Analysis of Machine Learning based Condition Monitoring Schemes Applied to Complex Electromechanical Systems* // 25th IEEE Int. Conf. on Emerging Technologies and Factory Automation (ETFA). Vienna, Austria, 2020. – DOI: 10.1109/ETFA46521.2020.9212026. – P. 1419–1422.
8. BAZHENOV S., SKRYABIN A., VERESNIKOV G. *The Development of Algorithms for EMA Fault Early Detection System* // Proc. of 32nd Congress of the Int. Council of the Aeronautical Sciences (ICAS 2020). – 2021. – Shanghai, China, CSAA.
9. BALABAN E. et. al.. *Prognostic Health-Management System Development for Electromechanical Actuators* // Journal of Aerospace Information Systems. – 2015. – Vol. 12, No. 3.
10. BOSCH C., HAJEK M., ISMAIL M.A.A. *Preliminary system safety assessment of electromechanical actuation architectures for unmanned aerial vehicles* // 5th Int. Conf. on Control and Fault-Tolerant Systems (SysTol'21). Saint-Raphaël, France, 2021. – P. 21297562.
11. DREYER N., IMMLER T., WEBER G. *Preparing electric actuation technology for upcoming applications* // Int. Conf. on More Electric Aircraft MEA'24, Toulouse, France, 2024. – 5 p.
12. EROFEEV E., SKRYABIN A., STEBLINKIN A. et al. *Methodologies and test-rig configurations for the experimental improvement of flight control actuation systems* // Int. Conf. on Recent Advances in Aerospace Actuation Systems and Component, Toulouse, France, 2018. – P. 109–116.
13. FREEMAN P., BALAS G.J. *Actuation Failure Modes and Effects Analysis for a small UAV* // American Control Conference (ACC-2014), Portland, Oregon, USA, 2014. – P. 1292–1297.

14. KHELIFI A., MANSOUR N., LAKHAL B. et al. *Artificial Neural Network-based Fault Detection* // 5th Int. Conf. on Control, Decision and Information Technologies (CoDIT), Thessaloniki, Greece, 2018. – P. 1017–1022. – DOI: 10.1109/CoDIT.2018.8394963.
15. LI C., WANG Z., BU S. et al. *Semi-Supervised Adaptive Parzen Gentleboost Algorithm for Fault Diagnosis* // Proc. of the 21st Int. Conf. on Pattern Recognition (ICPR 2012), Tsukuba, Japan, 2012. – P. 2290–2293.
16. MOHAMED A.A. ISMAIL, BALABAN E., WINDELBERG J. *Spall Fault Quantification Method for Flight Control Electromechanical Actuator* // Actuators. – 2022. – Vol. 11(2), No. 29. – DOI: 103390/act11020029.
17. PATIL S.S., PATHAN S.K. *A Novel approach of fault detection using artificial neural network (ann)* // Int. Journal of Advanced Research in Computer Engineering & Technology (IJARCET). – 2015. – Vol. 4, Iss. 6. – P. 2715–2720.
18. PHAM T.-H., BIERIG A. *First step towards a robust vibration-based condition monitoring algorithm for electro-mechanical flight control actuators* // 7th Int. Conf. on Recent Advances in Aerospace Actuation Systems and Components (R3ASC-2016), Toulouse, France, 2016. – P. 33–40.
19. SAUCEDO-DORANTES J.J., OSORNIO-RIOS R.A., ROMERO-TRONCOSO R.J. et al. *Novel condition monitoring approach based on hybrid feature extraction and neural network for assessing multiple faults in electromechanical systems* // IEEE 12th Int. Symposium on Diagnostics for Electrical Machines, Power Electronics and Drives (SDEMPED), Toulouse, France, 2019. – P. 466–473.
20. URSU V., EROFEEV E., STEBLINKIN A. *The development and testing of the wave gear rotary EMA for aerospace applications* // 6th Int. Workshop on Aircraft System Technologies, Hamburg, Germany, 2017. – P. 95–103.
21. VAN DER LINDEN F, DREYER N, DORKEL A. *EMA Health monitoring: an overview* // 7th Int. Conf. on Recent Advances in Aerospace Actuation Systems and Components (R3ASC-2016), Toulouse, France, 2016. – P. 21–26.

DEVELOPMENT OF A METHODS AND INSTRUMENTS FOR CONDUCTING LIFE TESTS TO BUILD AN EARLY DIAGNOSTICS SYSTEM FOR UAV ELECTROMECHANICAL ACTUATOR

Sergey Bazhenov, V.A. Trapeznikov Institute of Control Sciences of RAS, Moscow, Central Aerohydrodynamic Institute (TsAGI), Zhukovsky, Doctor of Technical Sciences, Assistant Professor (sergey.bazhenov@tsagi.ru).

Georgiy Veresnikov, V.A. Trapeznikov Institute of Control Sciences of RAS, Moscow, Doctor of Technical Sciences (veresnikov@mail.ru).

Artem Golev, V.A. Trapeznikov Institute of Control Sciences of RAS, Moscow, Researcher (oiw23@mail.ru).

Vladimir Goncharenko, V.A. Trapeznikov Institute of Control Sciences of RAS, Moscow, Doctor of Technical Sciences, Assistant Professor (vladimirgonch@mail.ru).

Evgeniy Erofeev, V.A. Trapeznikov Institute of Control Sciences of RAS, Moscow, Central Aerohydrodynamic Institute (TsAGI), Zhukovsky, Reserarcher (evgeniy.erofeev@tsagi.ru).

Grigoriy Lazurin, Central Aerohydrodynamic Institute (TsAGI), Zhukovsky, Researcher (flight15@tsagi.ru).

Aleksey Skryabin, V.A. Trapeznikov Institute of Control Sciences of RAS, Moscow Central Aerohydrodynamic Institute (TsAGI), Zhukovsky, Candidat of Technical Sciences, Researcher (aleksey.skryabin@tsagi.ru).

Anastasiya Timofeeva, V.A. Trapeznikov Institute of Control Sciences of RAS, Moscow, Central Aerohydrodynamic Institute (TsAGI), Zhukovsky, Junior Researcher (anastasiya.timofeeva@tsagi.ru).

Vasiliy Fedenyuk, Central Aerohydrodynamic Institute (TsAGI), Zhukovsky, Engineer (flight15@tsagi.ru).

Abstract: The work is devoted to the creation of a technology for obtaining and accumulating experimental data during electromechanical actuator (EMA) life tests, necessary for the intelligent analysis algorithms development that provide EMA early diagnostics for aircraft condition-based maintenance. A life tests program and methodology have been developed that make it possible to accelerate EMA components wear processes and obtain data on functioning when reproducing loading cyclograms corresponding to ground testing during pre-flight preparation («active experiment»)

and during the flight of the aircraft («passive experiment»). An early diagnostics system prototype has been created, which includes a bench installation that provides elastic or weight mechanical loading of the unmanned aerial vehicle (UAV) EMA, and a software and hardware data collection complex that provides measurement and registration of physical quantities associated with the EMA work processes in various operating conditions and various technical conditions caused by wear. It is proposed to use physical quantities of electrical, mechanical, thermal and vibroacoustic nature as diagnostic signals. The results of UAV EMA life tests samples are presented, which made it possible to clarify the test program in terms of existing loads.

Keywords: electromechanical actuator, UAV, diagnostics, life-cycle tests.

УДК 629.7.062

ББК 30.2-5-05

DOI: 10.25728/ubs.2024.111.12

*Статья представлена к публикации
членом редакционной коллегии Н.А. Коргиным.*

*Поступила в редакцию 07.05.2024.
Опубликована 30.09.2024.*

О МОДЕЛИРОВАНИИ ПОВЕДЕНИЯ ОДНОГО МАРКОВСКОГО ПРОЦЕССА НА ОСНОВЕ МЕТОДА МОДЕЛИРОВАНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН С ПОМОЩЬЮ ИНТЕНСИВНОСТЕЙ

Зверкина Г. А.¹

*(ФГБУН Институт проблем управления
им. В.А. Трапезникова РАН, Москва)*

Кошелев А. А.²

*(ФГБУН Институт проблем управления
им. В.А. Трапезникова РАН, Москва, Московский
государственный университет им. М.В. Ломоносова, Москва)*

Ранее авторами был предложен метод имитационного моделирования случайной величины по интенсивности – одной из характеристик функции распределения. В данной работе представлены результаты тестирования этого метода для имитационного моделирования поведения стохастической модели, в данном случае – модели пары восстанавливаемых зависимых элементов. Характеристики периодов безотказной работы и периодов восстановления подобраны таким образом, что возможно аналитическое исследование поведения исследуемой модели. Результаты имитационного моделирования «классическим» методом и методом моделирования случайных величин по интенсивности были сравнены с аналитическим решением задачи о поведении коэффициента готовности исследуемой модели. В результате численных экспериментов показано, что моделирование поведения случайного процесса с методом моделирования случайных величин не уступает по точности «классическому» методу моделирования.

Ключевые слова: интенсивность случайной величины, численное моделирование Марковского процесса, вопросы оптимизации моделирования, моделирование случайной величины по интенсивности.

1. Введение

Для прогнозирования и анализа поведения сложных стохастических моделей теории массового обслуживания (ТМО), теории надёжности (ТН) и смежных областей в случае, когда теоретическими методами этого сделать невозможно или очень слож-

¹ Галина Александровна Зверкина, к.ф.-м.н., доцент (zverkina@gmail.com).

² Александр Анатольевич Кошелев, (koshelev030698@yandex.ru).

но, используется имитационное моделирование поведения таких моделей. Анализ полученных статистических данных по большому количеству таких моделирований позволяет прогнозировать поведение реальных сложных стохастических моделей.

Как правило, в таких моделях предполагается интенсивность (скорость) поступления заявок, обслуживания, времени ремонта/простоя и пр. постоянной или, в случае переменной интенсивности, – с заданным заранее «графиком» её изменения, определяемой функцией распределения (ф.р.), например, времени обслуживания (ремонта и пр.) – см. ниже определение ???. В этом случае имитационное моделирование проводится хорошо известными методами с использованием моделирования сл.в. по функции распределения этих сл.в. – см., например, [9, 12, 11].

Но хорошо известно, что в реальных сетях связи интенсивность (частота) появления новых соединений меняется как в зависимости от времени, так и в связи с внешними событиями; например, количество телефонных вызовов резко возрастает во время знаковых событий (важный спортивный матч, стихийное бедствие, авария на дороге и пр.).

Поэтому важно уметь моделировать поведение сложных стохастических систем в случае, когда интенсивность появления интересующих исследователя событий зависит от состояния всей системы (а также от неких внешних событий).

По-видимому, впервые асимптотический анализ поведения стохастических систем с зависящими от полного состояния системы интенсивностями был предложен в [19, 20]; вопросы вычисления оценок поведения такого рода стохастических систем исследовались, например, в [16, 17, 18, 21, 22, 23] и др. Но полученные в этих работах результаты дают только грубые оценки, а более точные оценки могут быть получены с помощью имитационного моделирования.

Имитационное моделирование поведения стохастических систем обслуживания и др. обычно сводится к последовательному моделированию случайных величин (длин периодов рабо-

ты/ремонта/обслуживания и пр.). При этом моделирование сл.в. чаще всего строится с использованием обратной функции к функции распределения (см., например, [9] и ниже раздел 2.2).

Однако для стохастических систем с зависящими от полного состояния системы интенсивностями стандартные методы моделирования малоприменимы: распределение периодов работы/ремонта/обслуживания и пр. не известны заранее, они могут меняться по ходу изменения полного состояния системы.

Поэтому в публикациях [5, 6, 7] был предложен метод моделирования сл.в. по её интенсивности. (Этот метод кратко упоминается в [14], но алгоритма использования интенсивности для моделирования сл.в. не приводится.)

В настоящей работе предполагается сравнить моделирование небольшой стохастической системы с зависящими от полного состояния системы интенсивностями «классическим» методом (методом обратной функции) и методом моделирования сл.в. по её интенсивности с явно вычисленными характеристиками поведения этой системы.

Естественно, будет выбрана небольшая система, поведение которой может быть описано с помощью цепи Маркова в непрерывном времени; для таких систем существует хорошо известный аппарат исследования их поведения. В следующих исследованиях будут изучаться системы с более сложной зависимостью поведения элементов от полного состояния системы.

2. Интенсивность и метод моделирования сл.в. по интенсивности

Сначала напомним определение интенсивности отказов/восстановлений/входящих потоков/обслуживания и пр.

2.1. Интенсивность

Рассмотрим протекающий во времени процесс, поведение которого определяется некоторыми периодами, длины которых – сл.в. с непрерывной ф.р. (простейший случай – процесс восстановления, см., например, [3, 10, 15]).

Одной из величин, характеризующих распределение сл.в., является интенсивность.

Определение 1. Интенсивностью случайной величины ξ с функцией распределения $F(t)$ называется функция

$$(1) \quad \lambda(t) = \lim_{\Delta \rightarrow 0} \frac{\mathbf{P}\{\xi \in (t, t + \Delta) \mid \xi > t\}}{\Delta} = \frac{F'(t)}{F(t)}.$$

Замечание 1. Здесь мы предполагаем, что $F(t)$ непрерывна и $F'(t)$ существует п.в., т.е. распределение сл.в. ξ абсолютно непрерывно.

Понятие интенсивности может быть распространено и на неабсолютно непрерывные (но не сингулярные) распределения (см., например, [13]), однако здесь мы ограничиваемся рассмотрением только абсолютно непрерывных распределений.

Замечание 2. Поскольку

$$F(t) = 1 - \exp \left(- \int_0^t \lambda(v) dv \right),$$

интенсивность определяет распределение неотрицательной сл.в.

Далее мы будем рассматривать только неотрицательные случайные величины.

Замечание 3. Поскольку $F(+\infty) = 1$, то

$$(2) \quad \int_0^{\infty} \lambda(t) dt = +\infty,$$

т.е. некоторая функция $\lambda(t)$ будет задавать интенсивность некоторой сл.в., если выполнено условие (2).

2.2. Схема алгоритма моделирования сл.в. по интенсивности

Предположим, что $\lambda(t)$ – интенсивность сл.в. ξ .

Возьмём $\{\mathcal{U}_i\}_{i=0,1,2,\dots}$ – последовательность независимых сл.в., равномерно распределённых на отрезке $[0, 1)$.

Рассмотрим точки $\{t_i\}_{i=0,1,2,\dots}$, $t_0 = 0$, $t_i < t_{i+1}$; обозначим $\Delta_i \stackrel{\text{def}}{=} t_{i+1} - t_i$.

Если $\lambda(t_0)\Delta_0 > \mathcal{U}_0$, то полагаем $\xi = 0$.

В противном случае, если $\lambda(t_1)\Delta_1 > \mathcal{U}_1$, то полагаем $\xi = t_1$.

В противном случае, если $\lambda(t_2)\Delta_2 > \mathcal{U}_2$, то полагаем $\xi = t_2$.

...

...если $\lambda(t_i)\Delta_i > \mathcal{U}_i$, то полагаем $\xi = t_i$...

Эта процедура с вероятностью 1 конечна, и её качество зависит от выбора длины шагов Δ_i – см. [5, 6, 7]. В работе [6] с помощью методов математической статистики было получено, что шаг моделирования зависит от значения интенсивности в рассматриваемый в момент t и вычисляется по формуле

$$\Delta(t) = \frac{K}{\lambda(t)}, \quad K = 0,001,$$

но данная процедура применима только к положительным сл.в.

Замечание 4. Если известна ф.р. $F(t)$ сл.в. ξ , то для компьютерного моделирования сл.в. ξ удобно применять следующее соотношение:

$$(3) \quad F^{-1}(\mathcal{U}) \stackrel{\mathcal{D}}{=} \xi,$$

где $\mathcal{U}$ равномерно распределена на $[0; 1]$,

$$F^{-1}(y) \stackrel{\text{def}}{=} \inf\{t \in \mathbb{R} : F(t) \geq y\},$$

об этом см., например, [9].

Хотелось бы отметить различие возможности применения в сложных системах массового обслуживания моделирования реализаций сл.в. (времени работы, восстановления и пр.) методом моделирования по интенсивности и методом обратной функции.

Моделируя значения некоторого случайного промежутка времени в системе массового обслуживания методом обратной функции, мы можем не учесть изменения состояния нашей системы массового обслуживания в течение этого промежутка (из-за сложности системы массового обслуживания). Метод моделирования по интенсивности автоматически учитывает любые сложности системы массового обслуживания, поскольку реализация всех действующих в каждый момент времени t сл.в. производится одновременно на малом промежутке $(t, t + \Delta t)$.

Очевидно, что моделирование сл.в. с помощью интенсивности требует гораздо большего числа операций, чем моделирование с помощью обратной функции. Именно поэтому требуется

тестирование метода моделирования сл.в. с помощью интенсивности на примере моделирования поведения стохастической системы и сравнение результатов как с результатом моделирования с помощью обратных функций, так и с теоретически вычисленными характеристиками поведения этой стохастической системы.

3. Выбор стохастической системы для тестирования метода имитационного моделирования её поведения с помощью интенсивностей

Рассматриваем систему надёжности, состоящую из двух зависимых восстанавливаемых элементов. В качестве примера можно привести две кассы в магазине, две взлётно-посадочные полосы аэропорта и пр. Зависимость состоит в том, что во время нерабочего состояния одного элемента второй должен работать с бóльшей нагрузкой, а значит, и изнашивается быстрее, т.е. повышается вероятность его отказа в ближайшее время. С другой стороны, если оба элемента находятся в отказавшем (нерабочем) состоянии, то скорость ремонтных работ изменяется, так как мощности ремонтников имеют некоторые ограничения.

Иначе говоря, интенсивности отказов и восстановлений обоих элементов зависят от полного состояния системы, которое включает в себя показатели состояния обоих элементов (работа или ремонт), а также прошедшее время нахождения обоих элементов в этом состоянии.

Пространство состояний такой системы – это произведение $\mathcal{S} := \{0; 1\} \times \mathbb{R}^+ \times \{0; 1\} \times \mathbb{R}^+$ со стандартной σ -алгеброй.

Элементами $\mathcal{S}$ являются векторы $Z = (i, x; j, y)$ с $i, j = 0, 1$ и $x, y \geq 0$.

Значение $i = 0$ (или 1) означает, что первый элемент системы находится в рабочем состоянии (или в состоянии ремонта); значение x обозначает время, прошедшее с момента последнего изменения первой переменной i . Аналогично значения j и y интерпретируются для второго элемента системы.

Процесс Y_t на пространстве $\mathcal{S}$ задан случайными величинами

нами: $X_t^1, \xi_t^1, X_t^2, \xi_t^2$. X_t^1 принимает значения 0 и 1 в зависимости от состояния первого прибора (работы/ремонта/пр.), а ξ_t^1 – случайное время с последнего изменения состояния прибора до момента t . Величины X_t^2, ξ_t^2 вводятся аналогичным образом для второго прибора. Тогда обозначим по определению $Y_t = (X_t^1, \xi_t^1, X_t^2, \xi_t^2)$ – состояние всей системы в момент времени t , и вероятности переходов (изменения компонент i, j) задаются формулами

$$\begin{aligned} \mathbf{P}\{X_{t+\Delta}^1 = |i_t - 1|, \xi_{t+\Delta}^1 \in (0; \Delta), X_{t+\Delta}^2 = j_t, \xi_{t+\Delta}^2 = y_t + \Delta \mid \\ |X_t^1 = i_t, \xi_t^1 = x_t, X_t^2 = j_t, \xi_t^2 = y_t\} = \\ = \rho_1(Y_t)\Delta + o(\Delta); \end{aligned}$$

$$\begin{aligned} \mathbf{P}\{X_{t+\Delta}^1 = i_t, \xi_{t+\Delta}^1 = x_t + \Delta, X_{t+\Delta}^2 = |j_t - 1|, \xi_{t+\Delta}^2 \in (0; \Delta) \mid \\ |X_t^1 = i_t, \xi_t^1 = x_t, X_t^2 = j_t, \xi_t^2 = y_t\} = \\ = \rho_2(Y_t)\Delta + o(\Delta); \end{aligned}$$

$$\begin{aligned} \mathbf{P}\{X_{t+\Delta}^1 = i_t, \xi_{t+\Delta}^1 = x_t + \Delta, X_{t+\Delta}^2 = j_t, \xi_{t+\Delta}^2 = y_t + \Delta \mid \\ |X_t^1 = i_t, \xi_t^1 = x_t, X_t^2 = j_t, \xi_t^2 = y_t\} = \\ = 1 - \rho_1(Y_t)\Delta - \rho_2(Y_t)\Delta + o(\Delta), \end{aligned}$$

где $\rho_1(Y_t), \rho_2(Y_t)$ – интенсивности переходов для первого и второго элемента соответственно. Обозначения $\rho_1(Y_t), \rho_2(Y_t)$ как раз подразумевают $\rho_1(Y_t) = \rho_1(i_t, x_t, j_t, y_t), \rho_2(Y_t) = \rho_2(i_t, x_t, j_t, y_t)$ для $i_t, j_t = 0, 1$.

Понятно, что в общем случае решить соответствующие поведению этой системы уравнения Колмогорова в явном виде не представляется возможным.

Поэтому в этой работе мы ограничимся случаем, когда $\rho_1(Y_t) = \rho_1(i_t, j_t)$ и $\rho_2(Y_t) = \rho_2(i_t, j_t)$; обозначим X_t – компоненты (i_t, j_t) процесса Y_t : $X_t = (i_t, j_t)$. Для упрощения формул обозначаем:

$\rho_1(0, 0) = \lambda_{11}$ (интенсивность отказа первого элемента, второй элемент исправен);

$\rho_1(1, 0) = \mu_{11}$ (интенсивность восстановления первого элемента,

второй элемент исправен);

$\rho_1(0, 1) = \lambda_{12}$ (интенсивность отказа первого элемента, второй элемент неисправен);

$\rho_1(1, 1) = \mu_{12}$ (интенсивность восстановления первого элемента, второй элемент неисправен).

$\rho_2(0, 0) = \lambda_{21}$ (интенсивность отказа второго элемента, первый элемент исправен);

$\rho_2(0, 1) = \mu_{21}$ (интенсивность восстановления второго элемента, первый элемент исправен);

$\rho_2(1, 0) = \lambda_{22}$ (интенсивность отказа второго элемента, первый элемент неисправен);

$\rho_2(1, 1) = \mu_{22}$ (интенсивность восстановления второго элемента, первый элемент неисправен).

4. Поведения изучаемой системы надёжности

Итак, процесс $X_t = (i_t, j_t)$, $i_t, j_t \in \{0; 1\}$, описывает поведение изучаемой системы надёжности; здесь i_t описывает состояние первого элемента, а j_t – второго; нулю соответствует рабочее состояние элемента, единице – состояние ремонта. Процесс X_t является цепью Маркова в непрерывном времени с пространством состояний $\mathcal{X} = \{(0,0), (1,0), (0,1), (1,1)\}$ (см. диаграмму на рис. 1). Предполагается, что в начальный момент времени $t = 0$ оба элемента исправны.

Изучение поведения процесса X_t – это изучение поведения вероятностей состояний $\{(0,0), (1,0), (0,1), (1,1)\}$; обозначим эти вероятности:

$$\begin{aligned} p_1(t) &= \mathbf{P}\{X_t = (0,0)\}, & p_2(t) &= \mathbf{P}\{X_t = (1,0)\}, \\ p_3(t) &= \mathbf{P}\{X_t = (0,1)\}, & p_4(t) &= \mathbf{P}\{X_t = (1,1)\}. \end{aligned}$$

Важной характеристикой системы, состоящей из восстанавливаемых элементов, является коэффициент готовности – вероятность того, что в данный момент времени система находится в работоспособном состоянии: $K_{\Gamma}(t) = 1 - p_4(t) = p_1(t) + p_2(t) + p_3(t)$.

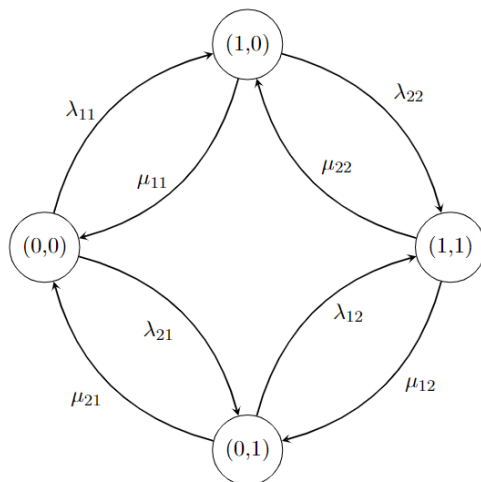


Рис. 1. Диаграмма процесса X_t

Для изучаемой системы надёжности система прямых уравнений Колмогорова с учётом начальных условий и условия нормировки имеет вид

$$(4) \quad \begin{cases} p_1'(t) = -(\lambda_{11} + \lambda_{21})p_1(t) + \mu_{11}p_2(t) + \mu_{21}p_3(t), \\ p_2'(t) = \lambda_{11}p_1(t) - (\mu_{11} + \lambda_{22})p_2(t) + \mu_{22}p_4(t), \\ p_3'(t) = \lambda_{21}p_1(t) - (\mu_{21} + \lambda_{12})p_3(t) + \mu_{12}p_4(t), \\ p_4'(t) = \lambda_{22}p_2(t) + \lambda_{12}p_3(t) - (\mu_{22} + \mu_{12})p_4(t), \\ p_1(t) + p_2(t) + p_3(t) + p_4(t) = 1, \quad p_1(0) = 1, \\ p_2(0) = p_3(0) = p_4(0) = 0. \end{cases}$$

Используя нормировочное условие, получаем дифференциальные уравнения для $p_1(t)$, $p_2(t)$, $p_3(t)$:

$$(5) \quad \begin{cases} p_1'(t) = -(\lambda_{11} + \lambda_{21})p_1(t) + \mu_{11}p_2(t) + \mu_{21}p_3(t), \\ p_2'(t) = (\lambda_{11} - \mu_{22})p_1(t) - \\ \quad - (\mu_{11} + \lambda_{22} + \mu_{22})p_2(t) - \mu_{22}p_3(t) + \mu_{22}, \\ p_3'(t) = (\lambda_{21} - \mu_{12})p_1(t) - \mu_{12}p_2(t) - \\ \quad - (\mu_{21} + \lambda_{12} + \mu_{12})p_3(t) + \mu_{12}, \\ p_1(0) = 1, \quad p_2(0) = p_3(0) = 0. \end{cases}$$

Из свойств X_t , являющейся невырожденной цепью Маркова в непрерывном времени с конечным числом состояний следует, что решения уравнений (5) $p_i(t)$ имеют предел при $t \rightarrow \infty$ (см. [8]).

5. Решение уравнения (5) и стационарные значения вероятностей

Поскольку цепь Маркова X_t имеет конечное число состояний и возвратна, то существуют пределы

$$(6) \quad \lim_{t \rightarrow +\infty} p_i(t) = \pi_i, \quad i = \overline{1, 4},$$

которые можно получить, решив систему

$$(7) \quad \begin{cases} -(\lambda_{11} + \lambda_{21})\pi_1 + \mu_{11}\pi_2 + \mu_{21}\pi_3 = 0, \\ (\lambda_{11} - \mu_{22})\pi_1 - (\mu_{11} + \lambda_{22} + \mu_{22})\pi_2 - \mu_{22}\pi_3 = -\mu_{22}, \\ (\lambda_{21} - \mu_{12})\pi_1 - \mu_{12}\pi_2 - (\mu_{21} + \lambda_{12} + \mu_{12})\pi_3 = -\mu_{12}. \end{cases}$$

Обозначив

$$(8) \quad \begin{aligned} \ell_{11} &= -\lambda_{11} - \lambda_{21}, & \ell_{12} &= \mu_{11}, & \ell_{13} &= \mu_{21}, \\ \ell_{21} &= \lambda_{11} - \mu_{22}, & \ell_{22} &= -\mu_{11} - \lambda_{22} - \mu_{22}, & \ell_{23} &= -\mu_{22}, \\ \ell_{31} &= \lambda_{21} - \mu_{12}, & \ell_{32} &= -\mu_{12}, & \ell_{33} &= -\mu_{21} - \lambda_{12} - \mu_{12}, \end{aligned}$$

имеем

$$(9) \quad \begin{cases} \ell_{11}\pi_1 + \ell_{12}\pi_2 + \ell_{13}\pi_3 = 0, \\ \ell_{21}\pi_1 + \ell_{22}\pi_2 + \ell_{23}\pi_3 = \ell_{23}, \\ \ell_{31}\pi_1 + \ell_{32}\pi_2 + \ell_{33}\pi_3 = \ell_{32}. \end{cases}$$

Или в векторном виде

$$(10) \quad \mathbb{L}\vec{\pi} = (0, \ell_{23}, \ell_{32})^T, \quad \mathbb{L} = (\ell_{ij}).$$

Решение системы (9) будет иметь вид

$$\begin{aligned} \pi_1 &= \frac{1}{\Delta_\ell} (\ell_{23}\ell_{32}(\ell_{13} + \ell_{12}) - \ell_{32}\ell_{22}\ell_{13} - \ell_{23}\ell_{12}\ell_{33}), \\ \pi_2 &= \frac{1}{\Delta_\ell} (\ell_{11}\ell_{23}(\ell_{33} - \ell_{32}) + \ell_{21}\ell_{32}\ell_{13} - \ell_{31}\ell_{22}\ell_{13}), \\ \pi_3 &= \frac{1}{\Delta_\ell} (\ell_{11}\ell_{22}\ell_{32} + \ell_{31}\ell_{12}\ell_{23} - \ell_{11}\ell_{32}\ell_{23} - \ell_{21}\ell_{12}\ell_{33}), \\ \Delta_\ell &= \det \mathbb{L}. \end{aligned}$$

Для решения системы (5) воспользуемся преобразованием Лапласа для решения системы обыкновенных дифференциальных уравнений с начальными условиями: $p_i(t) \doteq \Pi_i(r)$, $i = 1, 2, 3$. Используя замену (8), получаем

$$(11) \quad \begin{cases} (\ell_{11} - r)\Pi_1(r) + \ell_{12}\Pi_2(r) + \ell_{13}\Pi_3(r) = -1, \\ \ell_{21}\Pi_1(r) + (\ell_{22} - r)\Pi_2(r) + \ell_{23}\Pi_3(r) = \frac{\ell_{23}}{r}, \\ \ell_{31}\Pi_1(r) + \ell_{32}\Pi_2(r) + (\ell_{33} - r)\Pi_3(r) = \frac{\ell_{32}}{r}. \end{cases}$$

В векторном виде система (11) имеет вид

$$(12) \quad (\mathbb{L} - r\mathbb{E})\vec{\Pi} = \left(-1, \frac{\ell_{23}}{r}, \frac{\ell_{32}}{r}\right)^T,$$

где $\mathbb{E}$ – единичная матрица. Решением этой системы будут дробно-рациональные функции, зависящие от r :

$$(13) \quad \begin{aligned} \Pi_1(r) &= \frac{r^3 + a_1r^2 + b_1r + c_1}{r(r^3 + ar^2 + br + c)}, \\ \Pi_i(r) &= \frac{a_ir^2 + b_ir + c_i}{r(r^3 + ar^2 + br + c)}, \quad i = 2, 3. \end{aligned}$$

где

$$\begin{aligned} a &= -\ell_{11} - \ell_{22} - \ell_{33}, c = -\det \mathbb{L}, \\ b &= -\ell_{12}\ell_{21} + \ell_{11}\ell_{22} - \ell_{13}\ell_{31} - \ell_{23}\ell_{32} + \ell_{11}\ell_{33} + \ell_{22}\ell_{33}; \\ a_1 &= -\ell_{22} - \ell_{33}, b_1 = -\ell_{12}\ell_{23} - \ell_{13}\ell_{32} - \ell_{23}\ell_{32} + \ell_{22}\ell_{33}, \\ c_1 &= \ell_{13}\ell_{22}\ell_{32} - \ell_{12}\ell_{23}\ell_{32} - \ell_{13}\ell_{23}\ell_{32} + \ell_{12}\ell_{23}\ell_{33}, \\ a_2 &= \ell_{21} - \ell_{23}, b_2 = \ell_{11}\ell_{23} + \ell_{23}\ell_{31} - \ell_{23}\ell_{32} - \ell_{21}\ell_{33} + \ell_{23}\ell_{33}, \\ c_2 &= \ell_{13}\ell_{23}\ell_{31} - \ell_{13}\ell_{21}\ell_{32} + \ell_{11}\ell_{23}\ell_{32} - \ell_{11}\ell_{23}\ell_{33}, \\ a_3 &= \ell_{31} - \ell_{32}, b_3 = -\ell_{22}\ell_{31} + \ell_{11}\ell_{32} + \ell_{21}\ell_{32} + \ell_{22}\ell_{32} - \ell_{23}\ell_{32}, \\ c_3 &= -\ell_{12}\ell_{23}\ell_{31} + \ell_{12}\ell_{21}\ell_{32} - \ell_{11}\ell_{22}\ell_{32} + \ell_{11}\ell_{23}\ell_{32}. \end{aligned}$$

Для обращения преобразований Лапласа $\Pi_i(r)$ необходимо разложить на множители знаменатель $r^3 + ar^2 + br + c$, т.е. найти корни уравнения

$$(14) \quad r^3 + ar^2 + br + c = 0.$$

Поскольку система (5) является невырожденной, то величина c не обращается в ноль. Воспользуемся тригонометрическими

формулами Виета для вычисления корней кубического уравнения (14) (см., например, [1]).

Вычислим

$$Q = \frac{a^2 - 3b}{9}, \quad R = \frac{2a^3 - 3ab + 27c}{54}, \quad S = Q^3 - R^2.$$

В зависимости от значений Q, R, S имеют место следующие случаи.

1. $S > 0$. В этом случае есть три различных вещественных корня:

$$r_1 = -2\sqrt{Q} \cos \varphi - \frac{a}{3}, \quad r_{2,3} = -2\sqrt{Q} \cos \left(\varphi \pm \frac{2\pi}{3} \right) - \frac{a}{3},$$

$$\varphi = \frac{1}{3} \arccos \frac{R}{\sqrt{Q^3}},$$

и решения уравнений Колмогорова (5) примут вид

$$(15) \quad p_1(t) = -\frac{c_1}{r_1} \left(\frac{1}{r_2 r_3} - \frac{e^{r_2 t}}{r_2(r_3 - r_2)} - \frac{e^{r_3 t}}{r_3(r_2 - r_3)} \right) +$$

$$+ \frac{r_1^3 + a_1 r_1^2 + b_1 r_1 + c_1}{r_1} \left(\frac{e^{r_1 t}}{(r_2 - r_1)(r_3 - r_1)} + \right.$$

$$+ \frac{e^{r_2 t}}{(r_1 - r_2)(r_3 - r_2)} + \frac{e^{r_3 t}}{(r_1 - r_3)(r_2 - r_3)} \Bigg) +$$

$$+ \frac{(a_1 + r_1 + r_2)e^{r_2 t} - (a_1 + r_1 + r_3)e^{r_3 t}}{r_2 - r_3},$$

$$(16) \quad p_i(t) = \frac{(a_i r_1^2 + b_i r_1 + c_i)e^{r_1 t}}{r_1(r_2 - r_1)(r_3 - r_1)} + \frac{(a_i r_1 r_2 + b_i r_1 + c_i)e^{r_2 t}}{r_1(r_1 - r_2)(r_3 - r_2)} +$$

$$+ \frac{(a_i r_1 r_3 + b_i r_1 + c_i)e^{r_3 t}}{r_1(r_1 - r_3)(r_2 - r_3)} - \frac{c_i}{r_1} \left(\frac{1}{r_2 r_3} - \frac{e^{r_2 t}}{r_2(r_3 - r_2)} - \right.$$

$$\left. - \frac{e^{r_3 t}}{r_3(r_2 - r_3)} \right), \quad i = 2, 3.$$

2. $S = 0$. В этом случае есть три вещественных корня, причем один из корней имеет кратность 2:

$$r_1 = -2\sqrt[3]{R} - \frac{a}{3}, \quad r_2 = \sqrt[3]{R} - \frac{a}{3},$$

и решения уравнений Колмогорова (5) примут вид

$$(17) \quad p_1(t) = (1 + r_2 t) e^{r_2 t} + \\ + \frac{(r_1^3 + a_1 r_1^2 + b_1 r_1 + c_1)(e^{r_1 t} - e^{r_2 t})}{r_1(r_1 - r_2)^2} - \\ - \frac{(r_1^2 r_2 + a_1 r_1 r_2 + b_1 r_1 + c_1) e^{r_2 t}}{(r_1 - r_2)} + \frac{c_1}{r_1 r_2^2} (e^{r_2 t} (1 - r_2 t) - 1),$$

$$(18) \quad p_i(t) = \frac{(a_i r_1^2 + b_i r_1 + c_i)(e^{r_1 t} - e^{r_2 t})}{r_1(r_1 - r_2)^2} - \\ - \frac{(a_i r_1 r_2 + b_i r_1 + c_i) t e^{r_2 t}}{r_1(r_1 - r_2)} + \\ + \frac{c_i}{r_1 r_2^2} (e^{r_2 t} (1 - r_2 t) - 1), \quad i = 2, 3.$$

3. $S = 0$, $R = 0$. В этом случае есть один вещественный корень кратности 3:

$$r_0 = -\frac{a}{3},$$

и решения уравнений Колмогорова (5) примут вид

$$(19) \quad p_1(t) = (r_0^2 + b_1 r_0 + c_1) \frac{t^2 e^{r_0 t}}{2 r_0} + (1 + (a_1 + 2 r_0) t) e^{r_0 t} + \\ + \frac{c_1}{r_0^3} (e^{r_0 t} (1 - r_0 t) - 1),$$

$$(20) \quad p_i(t) = a_i t e^{r_0 t} + (a_i r_0^2 + b_i r_0 + c_i) \frac{t^2 e^{r_0 t}}{2 r_0} + \\ + \frac{c_i}{r_0^3} (e^{r_0 t} (1 - r_0 t) - 1), \quad i = 2, 3.$$

4. $S < 0$. В этом случае есть один вещественный корень r_1 и два комплексно сопряжённых корня $r_{2,3} = \alpha \pm i\beta$. В зависимости от значения величины Q имеется три случая.

а) $Q > 0$. В этом случае

$$r_1 = -2 \operatorname{sgn} R \sqrt{Q} \operatorname{ch} \varphi - \frac{a}{3}, \quad \alpha = \operatorname{sgn} R \sqrt{Q} \operatorname{ch} \varphi - \frac{a}{3}, \\ \beta = \sqrt{3Q} \operatorname{sh} \varphi, \quad \varphi = \frac{1}{3} \operatorname{Arch} \frac{|R|}{\sqrt{Q^3}}.$$

б) $Q < 0$. В этом случае

$$r_1 = -2 \operatorname{sgn} R \sqrt{|Q|} \operatorname{sh} \varphi - \frac{a}{3}, \quad \alpha = \operatorname{sgn} R \sqrt{|Q|} \operatorname{sh} \varphi - \frac{a}{3},$$

$$\beta = \sqrt{3|Q|} \operatorname{ch} \varphi, \quad \varphi = \frac{1}{3} \operatorname{Arsh} \frac{|R|}{\sqrt{|Q|^3}}.$$

в) $Q = 0$. В этом случае

$$r_1 = -\sqrt[3]{c - \frac{a^3}{27}} - \frac{a}{3}, \quad \alpha = -\frac{a + r_1}{2},$$

$$\beta = \frac{1}{2} \sqrt{|(a - 3r_1)(a + r_1) - 4b|},$$

и решения уравнений Колмогорова (5) примут вид

$$(21) \quad p_1(t) = \frac{1}{(r_1 - \alpha)^2 + \beta^2} \times$$

$$\times \left\{ \left[r_1^2 e^{r_1 t} + ((\alpha - r_1)^2 + \beta^2 - r_1^2) e^{\alpha t} \cos \beta t + \right. \right.$$

$$\left. + \left(r_1 \beta + \alpha \left(\beta + \frac{(\alpha - r_1) \alpha}{\beta} \right) \right) e^{\alpha t} \sin \beta t \right] +$$

$$+ a_1 \left[r_1 e^{r_1 t} - r_1 e^{\alpha t} \cos \beta t + \right.$$

$$+ \left(\beta + \frac{(\alpha - r_1) \alpha}{\beta} \right) e^{\alpha t} \sin \beta t \left. \right] + \frac{b_1 r_1 + c_1}{r_1} \left[e^{r_1 t} - e^{\alpha t} \cos \beta t + \right.$$

$$\left. + \frac{\alpha - r_1}{\beta} e^{\alpha t} \sin \beta t \right] \Big\} -$$

$$- \frac{c_1}{r_1(\alpha^2 + \beta^2)} \left(1 - e^{\alpha t} \left(\cos \beta t - \frac{\alpha}{\beta} \sin \beta t \right) \right),$$

$$(22) \quad p_i(t) = \frac{1}{(r_1 - \alpha)^2 + \beta^2} \left\{ a_i \left[r_1 e^{r_1 t} - r_1 e^{\alpha t} \cos \beta t + \right. \right.$$

$$\left. + \left(\beta + \frac{(\alpha - r_1) \alpha}{\beta} \right) e^{\alpha t} \sin \beta t \right] +$$

$$+ \frac{b_i r_1 + c_i}{r_1} \left[e^{r_1 t} - e^{\alpha t} \cos \beta t + \frac{\alpha - r_1}{\beta} e^{\alpha t} \sin \beta t \right] \Big\} -$$

$$- \frac{c_i}{r_1(\alpha^2 + \beta^2)} \left(1 - e^{\alpha t} \left(\cos \beta t - \frac{\alpha}{\beta} \sin \beta t \right) \right), \quad i = 2, 3.$$

6. Результаты моделирования методом, основанном на интенсивностях, и методом обратной функции

С помощью указанного выше в разделе 2.2 метода моделирования сл.в. по интенсивности ([5, 6, 7]) и метода моделирования с помощью обратной функции (3) было проведено моделирование различного количества N реализаций процесса X_t с различными параметрами $\lambda_{k\ell}, \mu_{k\ell}$ на достаточно небольших интервалах, поскольку скорость сходимости к стационарному значению в формулах (15), (16), (17), (18), (19), (20), (21), (22), экспоненциальна со скоростью выше, чем $\min_{i=1,2,3} \{-\operatorname{Re} r_i\}$. Количество N реализаций процесса X_t , необходимое для получения достаточно близкого к теоретическому поведению $K_\Gamma(t)$ графика эмпирических значений $K_\Gamma(t)$, существенно зависит от $\min_{i=1,2,3} \{-\operatorname{Re} r_i\}$: чем меньше это значение, тем больше требуется реализаций процесса X_t , что можно увидеть на нижеприведённых данных.

Реализация методов моделирования, как и в работах [5, 6], проводилась на языке программирования С. Количество реализаций бралось разное; для больших значений $\min\{\lambda_{k\ell}, \mu_{k\ell}\}$, естественно, требуется меньшее количество реализаций для получения достаточно близкого к теоретическому графика поведения эмпирических значений $K_\Gamma(t)$; при этом $\lim_{t \rightarrow \infty} K_\Gamma(t) = \widetilde{K}_\Gamma = \pi_1 + \pi_2 + \pi_3$ (см. (6)), и к этому значению должны приближаться (а потом около него осциллировать) графики эмпирических значений $K_\Gamma(t)$.

Важно отметить, что, как правило, интенсивность отказов восстанавливаемых элементов обычно существенно меньше интенсивности восстановления: обычно восстанавливаемый элемент работает долго, а восстанавливается быстро. Поэтому при имитационном моделировании сложных стохастических систем, как правило, для получения достоверного результата потребуется очень большое количество реализаций моделируемого процесса, см. ниже графики на рис. 6, 7.

На графиках ниже приведены результаты моделирования коэффициента готовности для систем с различными параметрами.

Везде на графиках поведения коэффициента готовности: голубой цвет – моделирование по интенсивности, оранжевый цвет – моделирование методом обратной функции, фиолетовый цвет – аналитическое решение, зелёный цвет – предельное значение коэффициента готовности. Напомним, что коэффициент готовности $K_{\Gamma}(t)$ – это вероятность того, что в момент времени t система работоспособна, т.е. хотя бы один элемент находится в исправном состоянии. Обнаружилось, что, несмотря на существенно большее число вычислений при моделировании по интенсивности, результаты моделирования обоими способами близки между собой.

6.1. Случай комплексных корней уравнения (14)

Результаты моделирования поведения коэффициента готовности в случае параметров $\lambda_{11} = 2$; $\lambda_{12} = 3$; $\lambda_{21} = 4$; $\lambda_{22} = 3$; $\mu_{11} = 7$; $\mu_{12} = 8$; $\mu_{21} = 6$; $\mu_{22} = 7$ представлены на рис. 2, 3, 4. Решение уравнения (5) соответствует формулам (21), (22). При этом $Q = 12,444$; $S = -207,583$; $R = 46,203$; $\widetilde{K}_{\Gamma} = 0,911450$.

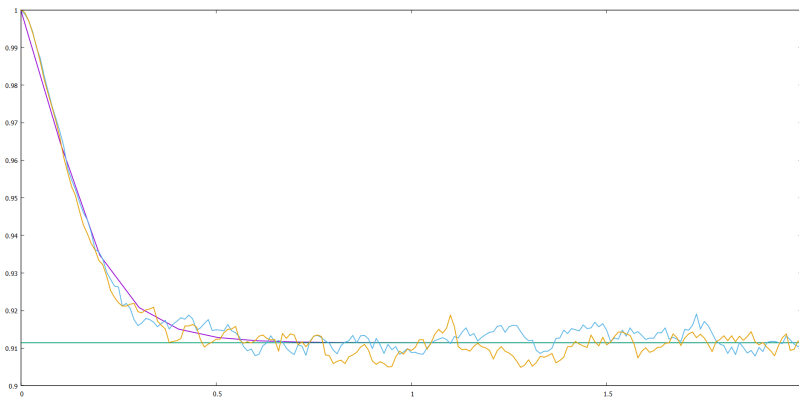


Рис. 2. $N = 10^4$; осцилляция численных решений после достижения стационарного значения $\widetilde{K}_{\Gamma}$ в пределах $\pm 0,01$

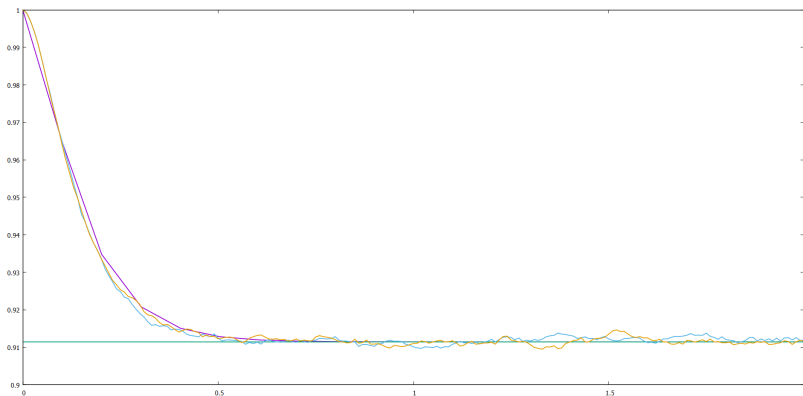


Рис. 3. $N = 10^5$; осцилляция численных решений после достижения стационарного значения $\widetilde{K}_\Gamma$ в пределах $\pm 0,003$

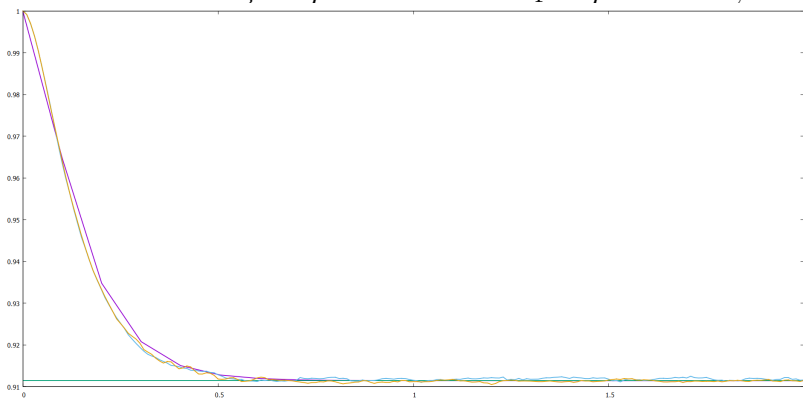


Рис. 4. $N = 10^6$; осцилляция численных решений после достижения стационарного значения $\widetilde{K}_\Gamma$ в пределах $\pm 0,0008$

6.2. Случай двух действительных корней уравнения (14) (один корень кратный)

Результаты моделирования поведения коэффициента готовности в случае параметров $\lambda_{11} = \lambda_{12} = \lambda_{21} = \lambda_{22} = \mu_{11} = \mu_{12} = \mu_{21} = \mu_{22} = 1$ представлены на рис. 5, 6, 7. Решение уравнения (5) соответствует формулам (17), (18). При этом $Q = 0,444$; $S = 0$; $R = 0,296$; $\widetilde{K}_\Gamma = 0,75$.

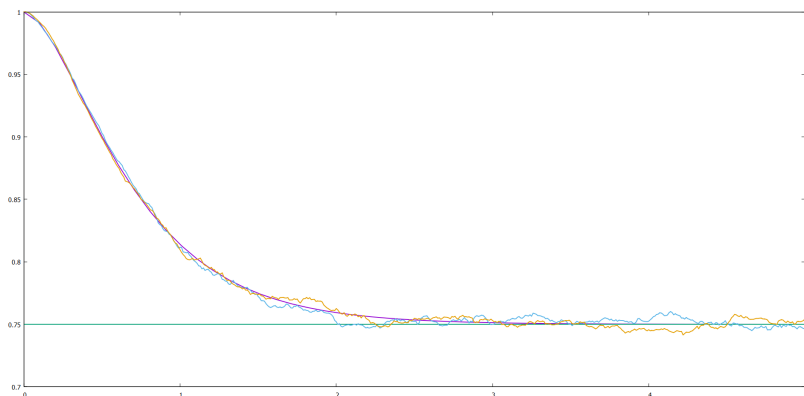


Рис. 5. $N = 10^4$; осцилляция численных решений после достижения стационарного значения $\widetilde{K}_\Gamma$ в пределах $\pm 0,01$

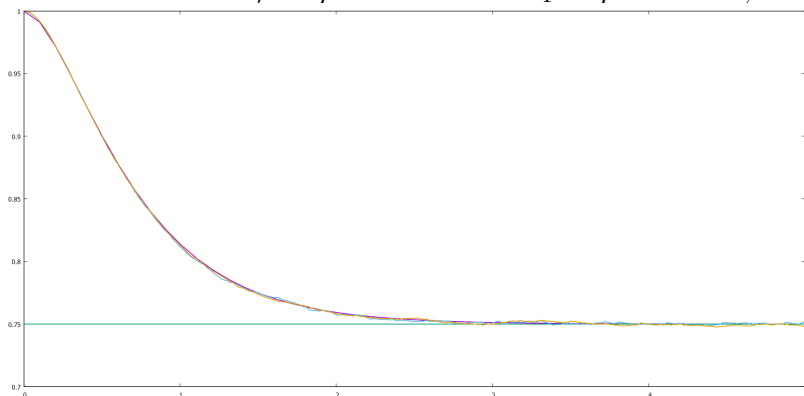


Рис. 6. $N = 10^5$; осцилляция численных решений после достижения стационарного значения $\widetilde{K}_\Gamma$ в пределах $\pm 0,002$

6.3. Случай трёх действительных корней уравнения (14)

6.3.1. Вариант 1.

Результаты моделирования поведения коэффициента готовности в случае параметров $\lambda_{11} = 0,1$; $\lambda_{12} = 0,2$; $\lambda_{21} = 0,4$; $\lambda_{22} = 0,1$; $\mu_{11} = 10$; $\mu_{12} = 20$; $\mu_{21} = 50$; $\mu_{22} = 20$ представлены на рис. 8, 9. Решение уравнения (5) соответствует формулам (15), (16). При этом $Q = 148,236$; $S = 1711710,592$; $R = -1243,245$; $\widetilde{K}_\Gamma = 0,999936$.

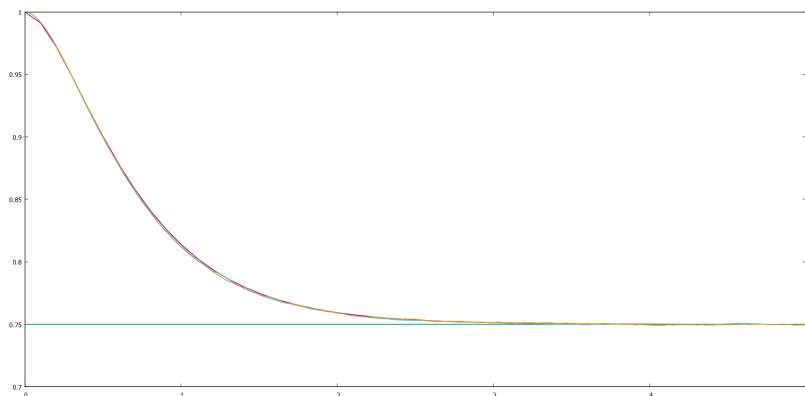


Рис. 7. $N = 10^6$; осцилляция численных решений после достижения стационарного значения $\widetilde{K}_\Gamma$ в пределах $\pm 0,0005$

Этот случай наиболее близок к реальной ситуации, когда интенсивности отказов существенно меньше интенсивностей восстановления (иначе говоря, восстанавливаемые элементы безотказно работают очень долго, а ремонтируются очень быстро).

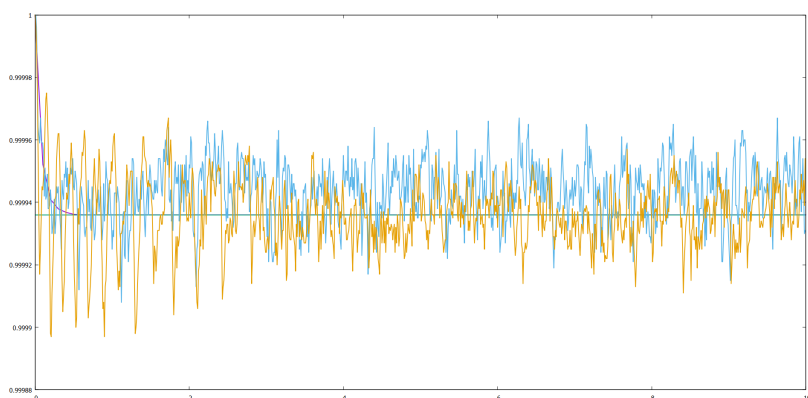


Рис. 8. $N = 10^6$; осцилляция численных решений после достижения стационарного значения $\widetilde{K}_\Gamma$ в пределах $\pm 0,0003$

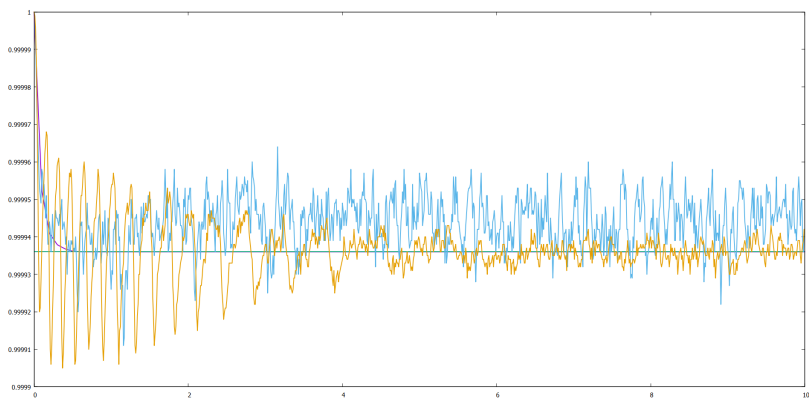


Рис. 9. $N = 10^7$; осциляция численных решений после достижения стационарного значения $\widetilde{K}_\Gamma$ в пределах $\pm 0,0003$

6.3.2. Вариант 2.

Результаты моделирования поведения коэффициента готовности в случае параметров $\lambda_{11} = 2$; $\lambda_{12} = 1$; $\lambda_{21} = 2$; $\lambda_{22} = 2$; $\mu_{11} = 3$; $\mu_{12} = 6$; $\mu_{21} = 3$; $\mu_{22} = 6$ представлены на рис. 10, 11, 12. Решение уравнения (5) соответствует формулам (15), (16). При этом $Q = 8,777$; $S = 175,888$; $R = 22,370$; $\widetilde{K}_\Gamma = 0,935644$.

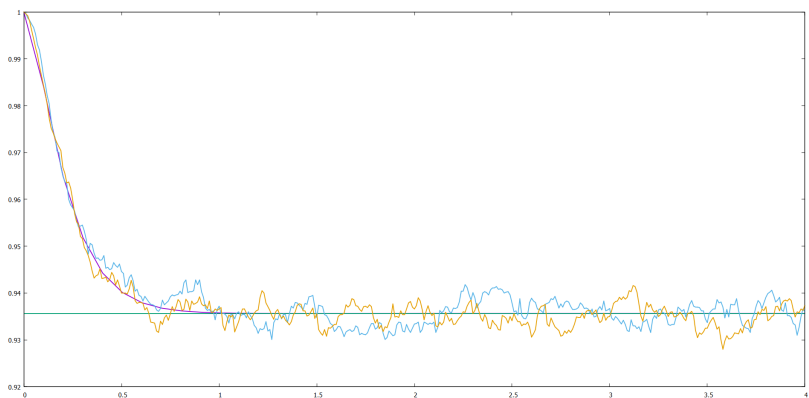


Рис. 10. $N = 10^4$; осциляция численных решений после достижения стационарного значения $\widetilde{K}_\Gamma$ в пределах $\pm 0,01$

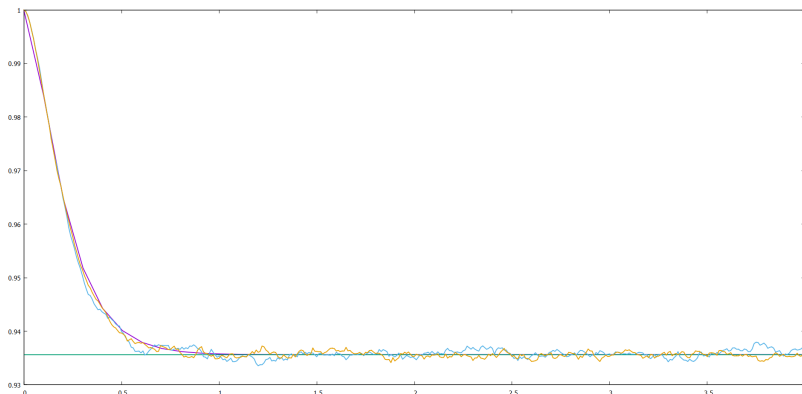


Рис. 11. $N = 10^5$; осцилляция численных решений после достижения стационарного значения $\widetilde{K}_\Gamma$ в пределах $\pm 0,001$

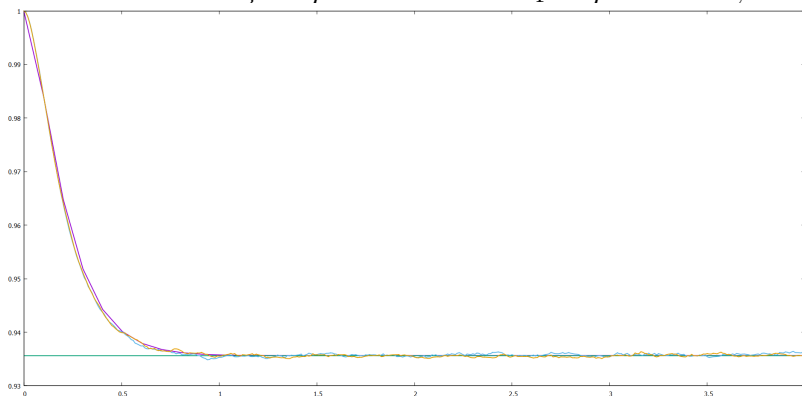


Рис. 12. $N = 10^6$; осцилляция численных решений после достижения стационарного значения $\widetilde{K}_\Gamma$ в пределах $\pm 0,0005$

7. Выводы

Анализ представленных результатов численных экспериментов приводит к следующему выводу. Несмотря на то, что моделирование случайных величин с помощью интенсивностей требует существенно большего числа операций, результаты моделирования процесса X_t с его помощью практически не отличаются от результатов моделирования «классическим» методом обратной функции. Точность моделирования зависит от соотношения чис-

ла реализаций моделированного процесса X_t и $\min\{\lambda_{kl}, \mu_{kl}\}$.

В дальнейшем планируется исследовать качество моделирования обоими способами на модели системы массового обслуживания и/или модели надёжности, поведение которой не может быть описано цепью Маркова в непрерывном времени.

Благодарность

Авторы выражают глубокую признательность В.В. Козлову за ценные консультации и полезные рекомендации в ходе работы над статьёй.

Литература

1. БРОНШТЕЙН И.Н., СЕМЕНДЯЕВ К.А. *Справочник по математике для инженеров и учащихся втузов: учебное пособие*. – СПб.: Лань, 2022. – 608 с.
2. ГНЕДЕНКО Б.В., БЕЛЯЕВ Ю.К., СОЛОВЬЁВ А.Д. *Математические методы в теории надёжности: Основные характеристики надёжности и их статистический анализ*. – М.: URSS. 2017. – 584 с.
3. ГНЕДЕНКО Б.В., КОВАЛЕНКО И.Н. *Введение в теорию массового обслуживания*. – М.: URSS, 2021. – 400 с.
4. ЗВЕРКИНА Г.А. *Об экспоненциальной скорости сходимости распределения одной нерегенерирующей системы надёжности* // *Фундаментальная и прикладная математика*. – 2020. – Т. 23, Вып. 1. – С. 145–160.
5. ЗВЕРКИНА Г.А., КОШЕЛЕВ А.А. *Об имитационном моделировании случайных величин с помощью интенсивности* // *Управление большими системами: сборник трудов*. – 2021. – Вып. 94. – С. 33–49.
6. ЗВЕРКИНА Г.А., КОШЕЛЕВ А.А. *О поиске оптимального шага при имитационном моделировании случайной величины с помощью интенсивности* // *Управление большими системами: сборник трудов*. – 2022. – Вып. 100. – С. 261–274.

7. ЗВЕРКИНА Г.А., КОШЕЛЕВ А.А. *Численное моделирование случайных величин с использованием их интенсивностей* // Труды 19-ой Всероссийской школы-конференции молодых ученых «Управление большими системами» (УБС'2023, Воронеж). – Воронеж: ФГБОУ ВО «Воронежский государственный технический университет», 2023. – С. 509–516.
8. КЛИМОВ Г.П. *Теория вероятностей и математическая статистика*. – М.: Изд-во Московского университета. – 2011. – 368 с.
9. ЛАГУТИН М.Б. *Наглядная математическая статистика: учебное пособие*. – М.: БИНОМ. Лаборатория знаний, 2019. – 472 с.
10. COX D.R. *Renewal Theory*. – Methuen & Co., 1967.
11. DEVROYE L. *Non-Uniform Random Variate Generation*. – Springer Science+Business Media, LLC, 1986.
12. HORMANN W., LEYDOLD J., DERFLINGER G. *Automatic Nonuniform Random Variate Generation*. – Springer, 2004.
13. KALIMULINA E., ZVERKINA G. *On generalized intensity function and its application to the backward renewal time estimation for renewal processes* // Proc. of the 5th Int. Conf. on Stochastic Methods (ICSM-5, 2020), 2020, Moscow, Russia, RUDN, Moscow. – P. 306–310.
14. SHELDON M.R. *Introduction to Probability Models*. – Academic Press; June 30, 2023. United Kingdom, Elsevier Inc., 2024. – 870 p.
15. SMITH W.L. *Renewal theory and its ramifications* // J. Roy. Statist. Soc. Ser. B. – 1958. – Vol. 20:2. – P. 243–302.
16. VERETENNIKOV A.YU. *On convergence rate for Erlang-Sevastyanov type models with infinitely many servers. In memory and to the 90th anniversary of A.D. Solov'yev (06.09.1927–06.04.2001)* // Theory Stoch. Process. – 2017. – Vol. 22(38):1. – P. 89–103.
17. VERETENNIKOV A. *On recurrence and availability factor for single-server system with general arrivals* // Theory and

- Applications (RT&A). – 2016. – Vol. 11, No. 3(42). – P. 49–58.
18. VERETENNIKOV A. *On polynomial recurrence for reliability system with a warm reserve* // Markov Processes and Related Fields. – 2019. – Vol. 25. – P. 745–761.
 19. VERETENNIKOV A. *On the rate of convergence for infinite server Erlang-Sevastyanov's problem* // Queueing Systems. – 2014. – Vol. 76(2). – P. 181–203.
 20. VERETENNIKOV A. *On the rate of convergence to the stationary distribution in the single-server queueing system* // Autom. Remote Control. – 2013. – Vol. 74(10). – P. 1620–1629.
 21. VERETENNIKOV A.YU., ZVERKINA G.A. *Simple Proof of Dynkin's Formula for Single-Server Systems and Polynomial Convergence Rates* // Markov Processes Relat. Fields. – 2014. – Vol. 20. – P. 479–504.
 22. VERETENNIKOV A.YU., ZVERKINA G.A. *On Polynomial Bounds of Convergence for the Availability Factor* // Distributed Computer and Communication Networks (Communications in Computer and Information Science). – 2016. – Vol. 601. – P. 358–369.
 23. ZVERKINA G.A. *A System with Warm Standby* // Computer Networks. CN 2019. Communications in Computer and Information Science. – 2019. – Vol. 1039. – P. 387–399.

ON THE MODELING THE BEHAVIOUR OF ONE MARKOV PROCESS USING THE METHOD OF MODELING RANDOM VARIABLES USING INTENSITIES

Galina Zverkina, V.A. Trapeznikov Institute of Control Sciences of RAS, Moscow, Cand.Sc., assistant professor (zverkina@gmail.com).

Alexandr Koshelev, V.A. Trapeznikov Institute of Control Sciences of RAS, Moscow, Lomonosov Moscow State University, Moscow, (koshelev030698@yandex.ru).

Abstract: In the past, the authors proposed a method for modeling of random variable using intensity which is one of the characteristics of the distribution function. The results of testing this method to simulate the behaviour of a stochastic model, specifically a model of a pair of recoverable dependent elements, are presented. The model under study's behavior can be studied analytically by selecting the characteristics of periods of failure-free operation and recovery periods. Simulation modeling using both the "classical" method and the method of modeling random variables by intensity yielded results. The availability factor behaviour of the model under study was compared to an analytical solution based on these results. The analytical solution was compared to numerical experiments to arrive at the following conclusion: the classical modeling method does not outperform the accuracy of modeling the behaviour of a random process using the method of modeling random variables using intensities.

Keywords: random variable intensity, numerical simulation of Markov process, discussion of simulation optimization, modeling of random variable using intensity.

УДК 519.2, 519.6

ББК 22.171, 22.176, 22.19

DOI: 10.25728/ubs.2024.111.13

*Статья представлена к публикации
членом редакционной коллегии А.С. Манделем.*

Поступила в редакцию 30.04.2024.

Дата опубликования 30.09.2024.