

ISSN 1819-3161

ПРОБЛЕМЫ УПРАВЛЕНИЯ

4/2025

CONTROL  SCIENCES

РЕДАКЦИОННЫЙ СОВЕТ

С. Н. Васильев, академик РАН,
И. А. Каляев, академик РАН,
Н. В. Кузнецов, чл.-корр. РАН,
В. А. Левин, академик РАН,
Н. А. Махутов, чл.-корр. РАН,
А. Ф. Резчиков, чл.-корр. РАН,
Е. А. Федосов, академик РАН

РЕДКОЛЛЕГИЯ

Ф. Т. Алескеров, д-р техн. наук,
В. Н. Афанасьев, д-р техн. наук,
Н. Н. Бахтадзе, д-р техн. наук,
В. М. Вишневский, д-р техн. наук,
О. И. Дранко, д-р техн. наук,
А. О. Калашников, д-р техн. наук,
В. В. Клочков, д-р экон. наук,
С. А. Краснова, д-р техн. наук,
О. П. Кузнецов, д-р техн. наук,
А. А. Лазарев, д-р физ.-мат. наук,
В. Г. Лебедев, д-р техн. наук,
В. Е. Лепский, д-р психол. наук,
Н. Е. Максимова, канд. техн. наук
(ответственный секретарь),
А. С. Мандель, д-р техн. наук,
Р. В. Мещеряков, д-р техн. наук,
А. И. Михальский, д-р биол. наук,
Д. А. Новиков, академик РАН
(гл. редактор),
Б. В. Павлов, д-р техн. наук,
Ф. Ф. Пашченко, д-р техн. наук
(зам. гл. редактора),
Л. Б. Рапопорт, д-р физ.-мат. наук,
С. В. Ратнер, д-р экон. наук,
Е. Я. Рубинович, д-р техн. наук,
Л. Ю. Филимонюк, д-р техн. наук,
М. В. Хлебников, д-р физ.-мат. наук,
А. Г. Чхартишвили, д-р физ.-мат. наук,
И. Б. Ядыкин, д-р техн. наук

РУКОВОДИТЕЛИ РЕГИОНАЛЬНЫХ РЕДСОВЕТОВ

Владивосток – О. В. Абрамов, д-р техн. наук,
Волгоград – А. А. Воронин, д-р физ.-мат. наук,
Воронеж – С. А. Баркалов, д-р техн. наук,
Курск – С. Г. Емельянов, д-р техн. наук,
Липецк – А. К. Погодаев, д-р техн. наук,
Пермь – В. Ю. Столбов, д-р техн. наук,
Ростов-на-Дону – Г. А. Угольников,
д-р техн. наук,
Самара – М. И. Гераськин, д-р экон. наук,
Саратов – В. А. Кушников, д-р техн. наук,
Тамбов – М. Н. Краснянский, д-р техн. наук,
Уфа – Б. Г. Ильясов, д-р техн. наук,
Челябинск – О. В. Логиновский, д-р техн. наук

ADVISORY BOARD

Е. А. Fedosov, RAS¹ Academician,
I. A. Kalyaev, RAS Academician,
N. V. Kuznetsov, RAS Corr. Member,
V. A. Levin, RAS Academician,
N. A. Makhutov, RAS Corr. Member,
A. F. Rezchikov, RAS Corr. Member,
S. N. Vassilyev, RAS Academician

EDITORIAL BOARD

V. N. Afanas'ev, Dr. Sci. (Tech.),
F. T. Aleskerov, Dr. Sci. (Tech.),
N. N. Bakhtadze, Dr. Sci. (Tech.),
A. G. Chkhartishvili, Dr. Sci. (Phys.-Math.),
O. I. Dranko, Dr. Sci. (Tech.),
L. Yu. Filimonuk, Dr. Sci. (Tech.),
A. O. Kalashnikov, Dr. Sci. (Tech.),
V. V. Klochkov, Dr. Sci. (Econ.),
M. V. Khlebnikov, Dr. Sci. (Phys.-Math.),
S. A. Krasnova, Dr. Sci. (Tech.),
O. P. Kuznetsov, Dr. Sci. (Tech.),
A. A. Lazarev, Dr. Sci. (Phys.-Math.),
V. G. Lebedev, Dr. Sci. (Tech.),
V. E. Lepskiy, D. Sc. (Psych.),
A. S. Mandel, Dr. Sci. (Tech.),
N. E. Maximova, Cand. Sci. (Tech),
Executive Editor-in-Chief,
R. V. Meshcheryakov, Dr. Sci. (Tech.),
A. I. Michalski, Dr. Sci. (Biol.),
D. A. Novikov, RAS Academician,
Editor-in-Chief,
F. F. Pashchenko, Dr. Sci. (Tech.),
Deputy Editor-in-Chief,
B. V. Pavlov, Dr. Sci. (Tech.),
L. B. Rapoport, Dr. Sci. (Phys.-Math.),
S. V. Ratner, Dr. Sci. (Econ.),
E. Ya. Rubinovich, Dr. Sci. (Tech.),
A. D. Tsvirkun, Dr. Sci. (Tech.),
V. M. Vishnevsky, Dr. Sci. (Tech.),
I. B. Yadykin, Dr. Sci. (Tech)

LEADERS OF REGIONAL BOARDS

Chelyabinsk – O. V. Loginovskiy, Dr. Sci. (Tech.),
Kursk – S. G. Emelyanov, Dr. Sci. (Tech.),
Lipetsk – A. K. Pogodaev, Dr. Sci. (Tech.),
Perm – V. Yu. Stolbov, Dr. Sci. (Tech.),
Rostov-on-Don – G. A. Ougolnitsky,
Dr. Sci. (Tech.),
Samara – M. I. Geraskin, Dr. Sci. (Econ.),
Saratov – V. A. Kushnikov, Dr. Sci. (Tech.),
Tambov – M. N. Krasnyanskiy, Dr. Sci. (Tech.),
Ufa – Б. Г. Ильясов, Dr. Sci. (Tech.),
Vladivostok – O. V. Abramov, Dr. Sci. (Tech.),
Volgograd – A. A. Voronin, Dr. Sci. (Phys.-Math.),
Voronezh – S. A. Barkalov, Dr. Sci. (Tech.)

¹Russian Academy of Sciences.



CONTROL SCIENCES
Научно-технический
журнал

6 номеров в год

ISSN 1819-3161 (Print)

ISSN 2712-8687 (Online)

Издается с 2003 года

УЧРЕДИТЕЛЬ и ИЗДАТЕЛЬ

Федеральное государственное
бюджетное учреждение науки
Институт проблем управления
им. В.А. Трапезникова РАН

Главный редактор
академик РАН
Д.А. Новиков

Заместитель главного редактора
Ф.Ф. Пащенко

Ответственный секретарь
Н.Е. Максимова

Выпускающий редактор
Л.В. Петракова

Адрес редакции
117997, ГСП-7, Москва,
ул. Профсоюзная, д. 65, к. 410

Тел./факс (495) 198-17-20, доб. 1410

E-mail: pu@ipu.ru

Интернет: <http://pu.mtas.ru>
<http://controlsciences.org>

Опубликовано: 16 сентября 2025 г.

Свидетельство о регистрации
ПИ № ФС 77-49203 от 30 марта 2012 г.
выдано Министерством Российской
Федерации по делам печати,
телерадиовещания и средств массовых
коммуникаций

Свидетельство о регистрации
Эл № ФС 77-80482 от 17 февраля 2021 г.
выдано Федеральной службой
по надзору в сфере связи,
информационных технологий и
массовых коммуникаций

Журнал входит в RSCI на платформе
Web of Science и Перечень
рецензируемых научных изданий ВАК

Журнал включен в Российский индекс
научного цитирования (РИНЦ).
На сайте Научной электронной
библиотеки (www.elibrary.ru) доступны
полные тексты статей.

© Федеральное государственное
бюджетное учреждение науки
Институт проблем управления
им. В.А. Трапезникова РАН

ПРОБЛЕМЫ УПРАВЛЕНИЯ

4.2025

СОДЕРЖАНИЕ

Обзоры

Буков В. Н., Бронников А. М., Попов А. С., Шурман В. А.

Методы мониторинга технического состояния систем

в целях управления их избыточностью. Ч. 3. Неклассические

модели в технической диагностике 3

Системный анализ

Калянов Г. Н., Новиков Д. А. Классификация процессов

деятельности 21

Математические проблемы управления

Кононов Д. А., Фуругян М. Г. Управление комплексом

работ с неопределенными моментами поступления запросов

на выполнение 31

Управление в социально-экономических системах

Цыганов В. В. Механизмы устойчивого развития

транспортной корпорации с контролируемым обучением 40

Сыч В. В. Исследование взаимосвязи активности

в социальной сети и академической успеваемости студентов

технического университета 52

Управление подвижными объектами и навигация

Московский А. Д. Применение недоопределенных моделей

в задаче глобальной локализации мобильного робота 64



CONTROL SCIENCES
Scientific Technical
Journal

6 issues per year

ISSN 1819-3161 (Print)

ISSN 2712-8687 (Online)

Published since 2003

FOUNDER and PUBLISHER

V.A. Trapeznikov

Institute of Control Sciences
of Russian Academy of Sciences

Editor-in-Chief

D.A. Novikov, RAS Academician

Deputy Editor-in-Chief

F.F. Pashchenko

Executive Editor-in-Chief

N.E. Maximova

Editor

L.V. Petrakova

Editorial address

65 Profsoyuznaya st., office 410,
Moscow 117997, Russia

☎/📠 +7(495) 198-17-20, ext. 1410

✉ pu@ipu.ru

URL: <http://pu.mtas.ru>

<http://controlsciences.org>

Published: September 16, 2025

Registration certificate of

ПИ № ФС 77-49203 of 30 March 2012

issued by the Ministry of Press,

Broadcasting, and Mass Media

of the Russian Federation

Registration certificate of

Эл № ФС 77-80482 of 17 February 2021

issued by the Federal Service

for Supervision of Communications,

Information Technology, and Mass Media

The Journal is indexed in RSCI (Russian

Science Citation Index) on the platform

Web of Science and in the list of peer-

reviewed scientific publications of HAC

On the website of the Scientific electronic

library (www.elibrary.ru) full texts of

articles are available

© V.A. Trapeznikov

Institute of Control Sciences

of Russian Academy of Sciences

CONTROL SCIENCES

4.2025

CONTENTS

Surveys

Bukov, V. N., Bronnikov, A. M., Popov, A. S.,

and Shurman, V. A. Technical Condition Monitoring

Methods to Manage the Redundancy of Systems. Part III:

Nonclassical Models in Fault Diagnosis 3

Systems Analysis

Kalyanov, G. N. and Novikov, D. A. Classification

of Activity Processes 21

Mathematical Problems of Control

Kononov, D. A. and Furugyan, M. G. Managing a Complex

of Jobs with Uncertain Execution Request Arrivals 31

Control in Social and Economic Systems

Tsyganov, V. V. Sustainable Development Mechanisms

for a Transport Corporation with Supervised Learning 40

Sych, V. V. The Correlation Between the Social Network

Activity and Academic Performance of Technical University

Students: A Case Study of *Vkontakte* 52

Control of Moving Objects and Navigation

Moscowsky, A. D. Application of Subdefinite Models

in the Global Localization of a Mobile Robot 64

МЕТОДЫ МОНИТОРИНГА ТЕХНИЧЕСКОГО СОСТОЯНИЯ СИСТЕМ В ЦЕЛЯХ УПРАВЛЕНИЯ ИХ ИЗБЫТОЧНОСТЬЮ.

Ч. 3. Неклассические модели в технической диагностике

В. Н. Буков*, А. М. Бронников**, А. С. Попов***, В. А. Шурман****

*АО «Бортовые аэронавигационные системы», г. Москва,

**Московский государственный технический университет им. Н. Э. Баумана,

***ВУНЦ ВВС «Военно-воздушная академия им. проф. Н. Е. Жуковского и Ю. А. Гагарина», г. Воронеж,

****АО «НИИ авиационного оборудования», г. Жуковский, Московская обл.

*✉ v_bukov@mail.ru, **✉ bronnikov_a_m@mail.ru, ***✉ saga30@yandex.ru, ****✉ shurman@niiao.ru

Аннотация. Управление избыточностью технической системы предполагает использование процедуры мониторинга для реконфигурирования системы по мере необходимости. В обзоре, состоящем из четырех частей, приводятся современные методы диагностирования текущего состояния динамических систем как неотъемлемой функции мониторинга. В третьей части обзора изложены методы диагностирования, использующие нейронные сети, нечеткие модели, структурные модели и модели в виде множеств, а также статистический подход. Приводятся основы создания и обучения нейросети с ориентацией на выполнение функции диагностирования. Изложен подход с нечеткими моделями, общие правила формирования моделей и особенности их использования в задаче диагностирования. Промонстрирован подход со структурными моделями и описаны особенности обнаружения отказов на его основе. Представлены основы методов с использованием аппарата теории множеств, в частности формализма зонотопов. Кратко изложен подход на основе статистического распознавания образов.

Ключевые слова: искусственный нейрон, нейронная сеть, нечеткие модели, функция принадлежности, нечеткая кластеризация, структурные модели, декомпозиция Дулмаджа – Мендельсона, диагностируемость, зонотопы, градации в дискретном пространстве признаков, эмпирическая оценка отношения правдоподобия.

ВВЕДЕНИЕ

Основным ограничением подходов с аналитическими моделями является необходимость точного знания этих моделей (как их структуры, так и параметров), что в практических применениях зачастую неосуществимо. В случае неполной определенности математических моделей есть отработанные процедуры их параметрической или структурной идентификации [1–3]. Однако здесь внимание сосредоточено на «изначальном» (т. е. присутствующем самому подходу) преодолении указанного недостатка. К таким подходам относится исполь-

зование нейронных сетей, нечетких моделей, структурных моделей, аппарата теории множеств, а также статистических оценок.

Сторонники таких методов намеренно или в силу недостаточности информации отказываются от детального описания протекающих в диагностируемой системе процессов. Это так или иначе делает используемые модели грубыми (неточными, приближенными), однако позволяет придать диагностированию большую гибкость, снижающую или полностью обходящую негативные эффекты, чем и мотивировано применение этих подходов для эффективного контроля работоспособности динамических систем [4].

1. НЕЙРОСЕТЕВЫЕ МЕТОДЫ ДИАГНОСТИРОВАНИЯ

Нейронные сети (НС) получили широкое распространение в задачах технической диагностики [5–9]. Вариант таких сетей [10, 11] использует модели нелинейной авторегрессии с экзогенными входами (англ. *Nonlinear AutoRegressive with exogenous inputs*, NARX) в сочетании с алгоритмами обучения.

1.1. Формирование нейросети типа ARX

Один из подходов заключается в разработке и обучении набора нейросетевых оценщиков для воспроизведения поведения исследуемой системы. В структуре i -го отдельного нейрона [12] используется MISO-система (англ. *Multiple Input Single Output* – много входов, один выход), в которой выходной сигнал y_i вычисляется как функция взвешенной суммы всех входов u_i нейрона $u_{i,1}, \dots, u_{i,n_i}$ с соответствующими весами $w_{i,1}, \dots, w_{i,n_i}$, как показано на рис. 1. Функция f называется функцией активации.

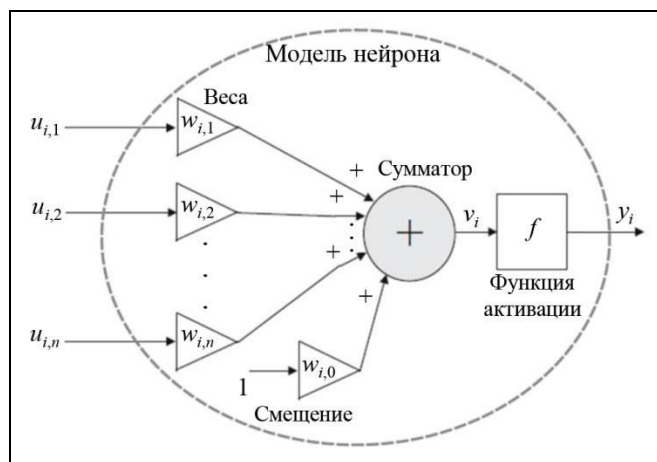


Рис. 1. Пример искусственного нейрона

Нейросети классифицируются по способам соединения элементов друг с другом [13, 14]. В сети с прямой связью элементы сгруппированы в однонаправленные слои. Первый из них, а именно входной слой, получает информацию непосредственно с сетевых входов, затем каждый последующий скрытый слой принимает входные данные от нейронов предыдущего слоя и передает выходные данные нейронам следующего слоя, вплоть до последнего выходного слоя, на котором генерируются окончательные сетевые выходные данные. Следовательно, нейроны подключаются от одного

слоя к другому, но не внутри одного и того же слоя. Единственным ограничением является количество нейронов в выходном слое, которое должно быть равно количеству фактических каналов выходных данных сети.

С другой стороны, рекуррентные сети [13] представляют собой многослойные сети, в которых выходные данные некоторых нейронов передаются обратно нейронам, принадлежащим предыдущим слоям, таким образом, информация передается как вперед, так и назад и обеспечивает динамическую память внутри сети.

Промежуточным решением является НС с подключенной линией задержки. Сети такого типа подходят для моделирования или прогнозирования эволюции динамической системы. В частности, разомкнутая сеть NARX при надлежащем обучении может оценить текущие (или последующие) результаты на основе полученных прошлых измерений входных и выходных данных системы.

Вообще говоря, для MIMO-систем (англ. *Multiple Input Multiple Output* – много входов, много выходов) сети NARX с открытым контуром следуют закону

$$\hat{y}_t = f_{\text{net}}(u_t, \dots, u_{t-d_u}, y_{t-1}, \dots, y_{t-d_y}), \quad (1)$$

где $\hat{y}_t$ – оценка выходных данных системы на шаге t ; u и y – измеренные системные входы и выходы; d_u и d_y – количество задержек входов и выходов соответственно; f_{net} – функция, реализуемая сетью и зависящая от архитектуры слоев, количества нейронов, их весов и функций активации. Структура открытой сети NARX (с разомкнутым контуром), используемой в качестве оценщика, показана на рис. 2.

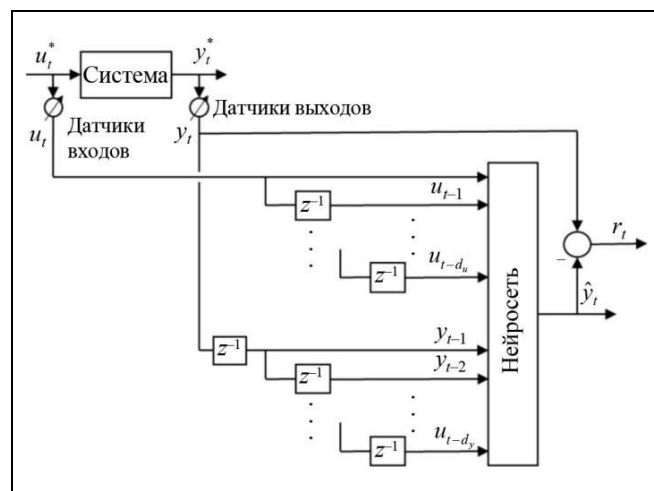


Рис. 2. Открытая динамическая нейросеть

Разница измеренного значения выхода y_t и его оценки (1) принимается в качестве невязки диагностирования r_t .

Когда доступны только входные измерения, сеть NARX может стать динамической НС при замыкании обратной связи сетевых выходов на входы, как показано на рис. 3.

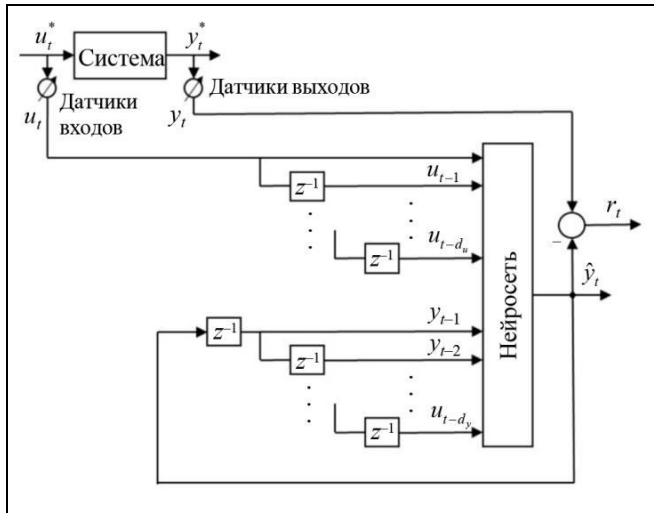


Рис. 3. Замкнутая динамическая нейросеть

Разработчик системы диагностирования может варьировать количество нейронов и связей между ними, в то время как значение весов $w_{i,j}$ внутри каждого нейрона получается путем обучения сети.

Поскольку компоненты НС зачастую выбираются субъективно, они могут быть не вполне адекватны моделируемой системе или процессу, что отражается и на реализациях сетей.

1.2. Обучение нейросети

Нейросеть становится инструментом диагностирования в результате ее обучения (подбора значений весовых коэффициентов w) на предварительно отобранных примерах функционирования диагностируемой системы как в работоспособном состоянии, так и с потерей работоспособности в результате, в частности, возникновения отдельной неисправности или их комбинации. Помимо основного варианта с фиксированными после обучения весовыми коэффициентами возможны варианты с уточнением (дообучением) параметров сети в процессе ее работы. Кроме того, нейросеть всякий раз следует переучивать даже для однотипных объектов, функционирующих в различных условиях и с различными историческими данными об эксплуатации.

Нейросеть можно обучить только тем «диагностическим событиям» (исправным, работоспособным, предотказным и неработоспособным состояниям), которые изначально известны и использованы в процессе обучения. Если не указать нейросети множество допустимых состояний, включая неисправные, то она не различит их, не сможет найти непредусмотренное состояние.

Использование нейронных сетей позволяет на практике решать задачу самодиагностирования, поскольку нейросеть устойчива к аномалиям в данных при правильном обучении и сама распознает диагностическое событие даже при неточности в исходных данных¹. Этого не произойдет в случае применения «жестких» алгоритмов диагностирования.

Формальной целью обучения является минимизация функции потерь (невязки) E [14], зависящей от вектора весовых коэффициентов w , и она может быть реализована в двух различных режимах:

- инкрементный режим (каждая пара вход – цель независимо генерирует обновление весовых коэффициентов сети);
- пакетный режим (все входные данные и функция потерь применяются к сети одновременно).

Хотя второй режим обучения, по сравнению с инкрементным, требует большего объема памяти, он характеризуется более быстрой сходимостью и приводит к меньшим ошибкам.

При наличии шаблонов в количестве S , соответствующих как отсутствию, так и наличию различных неисправностей в системе, для каждой из $P = C_p^2 = S! / 2(S-2)!$ возможных пар вход – выход можно записать вектор ошибок

$$e_p = y_p - \hat{y}_p = [e_{p,1} \quad \dots \quad e_{p,M}]^T, \quad p = \overline{1, P}, \quad (2)$$

где M – число выходов, т. е. искомым невязок диагностируемой системы. Общая же функция потерь, зависящая в соответствии с формулой (1) и рис. 1 от выбора весов $w = (w_1, \dots, w_N)$, принимает вид:

$$E(w) = \frac{1}{2} \sum_{p=1}^P \sum_{m=1}^M (y_{p,m} - \hat{y}_{p,m}(w))^2 = \frac{1}{2} \sum_{p=1}^P \sum_{m=1}^M e_{p,m}^2(w).$$

Любой стандартный алгоритм численной оптимизации [14–16] может быть использован для обновления значений параметров w_i с целью минимизации невязки $E(w)$. Среди них наиболее рас-

¹ Это тезис одного из рецензентов статьи. Авторы приняли с благодарностью.

пространены итеративные, использующие такие характеристики, как:

– градиент функции потерь

$$\text{grad}E = \frac{\partial E}{\partial w} = \left[\frac{\partial E}{\partial w_1} \quad \dots \quad \frac{\partial E}{\partial w_N} \right]^T,$$

– гессиан функции потерь

$$H(E) = \begin{bmatrix} \partial^2 E / \partial w_1^2 & \dots & \partial^2 E / \partial w_1 \partial w_N \\ \vdots & \ddots & \vdots \\ \partial^2 E / \partial w_N \partial w_1 & \dots & \partial^2 E / \partial w_N^2 \end{bmatrix},$$

– якобиан ошибок оценивания

$$J(e) = \begin{bmatrix} \partial e_{1,1} / \partial w_1 & \dots & \partial e_{1,1} / \partial w_N \\ \partial e_{1,2} / \partial w_1 & \dots & \partial e_{1,2} / \partial w_N \\ \vdots & \ddots & \vdots \\ \partial e_{p,M} / \partial w_1 & \dots & \partial e_{p,M} / \partial w_N \end{bmatrix}.$$

Последовательные итерации этих алгоритмов состоят из обновления значений параметров и вычисления нового значения функции потерь до тех пор, пока не будет выполнено условие останова.

Достижимая чувствительность диагностирования к отдельным отказам является сложным вопросом, вряд ли имеющим обобщенную оценку.

В качестве итога отметим, что нейросети аппроксимируют любую нелинейную и динамическую функцию при наличии подходящей структуры весовых коэффициентов. Более того, онлайн-обучение позволяет легко изменять систему диагностирования в тех случаях, когда вносятся изменения в физический процесс или в систему управления. Нейросети могут выполнять функцию обобщения, когда представлены входные данные, отсутствующие в обучающих данных, и принимать разумные решения в случаях зашумленных или поврежденных данных. Они также легко применимы к многопараметрическим системам и имеют высокий уровень параллельности структуры. Нейросеть может одновременно работать с качественными и количественными данными. Кроме того, нейросеть может быть очень полезна, когда отсутствует математическая модель системы, т. е. аналитические модели по какой-либо причине не могут быть применены. Все это вместе, как ожидается, может обеспечить более высокую степень отказоустойчивости.

С другой стороны, принципиальной особенностью нейросети является то, что она работает как «черный ящик», в котором отсутствует качественная и количественная информация о представляемой ею модели. Ввиду этого обстоятельства разработчики и заказчики пока воздерживаются от применения соответствующих решений, прежде всего для задач высокой ответственности (например, в гражданской авиации – для задач, связанных с без-

опасностью). Признано, что существует двусмысленность в отношении работы нейросети в случае непредвиденных ситуаций [17].

Описанный подход таит в себе проблему структурного характера. Заранее практически невозможно подобрать такое сочетание измеряемых параметров и структуры сети, чтобы гарантированно получать уверенное различение отказов из заданного перечня. Обычно это достигается путем проб и ошибок. Кроме того, совокупность обучающих примеров должна быть представительной, т. е. содержать достаточное количество комбинаций условий решаемой задачи.

2. ДИАГНОСТИРОВАНИЕ С НЕЧЕТКИМИ МОДЕЛЯМИ

Подход с нечеткими моделями также помогает обойти проблему точного знания модели диагностируемой системы. При этом такой подход может применяться на двух уровнях: сначала нечеткие описания используются для генерирования невязок, а затем снова с помощью нечеткой логики достигается обнаружение отказов [18–24].

Системы нечеткой логики предлагают лингвистическую модель динамики системы, которую можно легко понять с помощью определенных правил. Им также присущи способности работать с неточными или зашумленными данными.

2.1. Формирование нечеткой модели

Эффективный подход к проектированию системы нечеткой логики начинается с разбиения доступных данных на подмножества (кластеры) [19], характеризующиеся более простым (линейным или аффинным) поведением. Кластер можно определить как группу данных, которые больше похожи друг на друга, чем на данные другого кластера. Сходство между данными обычно выражается в терминах их удаленности от определенного элемента, располагаемого внутри кластера и используемого в качестве прототипа этого кластера. Нечеткая кластеризация обеспечивает эффективный инструмент для разделения данных, при котором переходы между подмножествами являются плавными, определяемыми так называемыми функциями принадлежности, а не резкими.

В общем случае аппроксимация нелинейных MISO-систем может быть достигнута с помощью нечеткого рассуждения [25, 26]. Однако согласно подходу, предложенному в работе [27], следствия становятся четкими функциями входных данных, и каждое нечеткое правило принимает форму

$$R_i : \text{IF } x \in X_i \text{ THEN } \hat{y}_i = a_i^T x + b_i, \quad i = \overline{1, n_c}, \quad (3)$$

где R_i – i -е правило; X_i – i -й кластер; n_c – общее количество кластеров; a_i – вектор параметров; b_i – скалярное смещение; $\hat{y}_i$ – вывод i -го правила, в то время как y_i – выход диагностируемой системы. Правило (3) для одномерной переменной x при $n_c = 5$ иллюстрирует рис. 4.

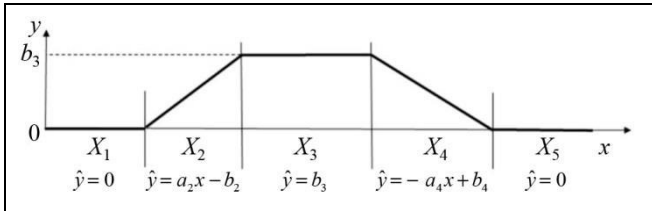


Рис. 4. Графическая иллюстрация правила (3)

Введение функции принадлежности $\lambda_i(x)$, учитывающей логику выбора кластеров (IF ... THEN ...), позволяет записать результирующую формулу для всего множества правил (3) в виде взвешенной суммы

$$\hat{y}_t = \frac{\sum_{i=1}^{n_c} \lambda_i(x_t) y_t^i}{\sum_{i=1}^{n_c} \lambda_i(x_t)}, \quad (4)$$

где x_t – подходящая комбинация входного и выходного сигналов; y_t^i – выходной сигнал локальной линейной (или аффинной) i -й модели, заданный следующим образом:

$$y_t^i = \sum_{k=1}^n a_{i,k} y_{t-k} + \sum_{k=0}^{n-1} b_{i,k} u_{t-k} + c_i, \quad (5)$$

где $a_{i,k}$, $b_{i,k}$ и c_i – параметры i -й модели в подпространстве (кластере) X_i .

Формула (5) содержит как последние полученные u_t , y_{t-1} , так и предыдущие u_{t-1} , y_{t-2} , ... выборки входных и выходных данных, что отражает динамическое поведение системы. Это соответствует рассмотрению наблюдения как линейной авторегрессионной модели с экзогенными входными данными ARX [9] порядка s , в которой вектор регрессора принимает вид

$$x_t = \left[\underbrace{y_{t-1} \cdots y_{t-s}}_{\text{выходная выборка}} \underbrace{u_t \cdots u_{t-s+1}}_{\text{входная выборка}} \right]^T,$$

где t – текущий момент дискретного времени.

Нечеткая кластеризация [20, 28, 29], в отличие от классической кластеризации, позволяет любому элементу принадлежать нескольким кластерам одновременно.

В качестве иного подхода нечеткой модели можно указать на модель нечеткого конечного автомата², которая используется для описания дискретно-событийных систем [22].

2.2. Диагностирование

В качестве невязки может быть принята разность фактически измеренных и получаемых посредством неявной модели (2) оценок выходов системы:

$$r_t = y_t - \hat{y}_t.$$

Укрупненная схема формирования невязки на основе адаптивной нечеткой модели показана на рис. 5. Нечеткий прототип и рекурсивный вычислитель параметров, реализующий какой-либо алгоритм идентификации, используют независимо одни и те же входные и выходные данные.

При необходимости обнаружения нескольких различных неисправностей схема их определения приводится к виду с мультиплицированием прототипов (моделей). Такая схема будет рассмотрена в ч. 4 обзора.

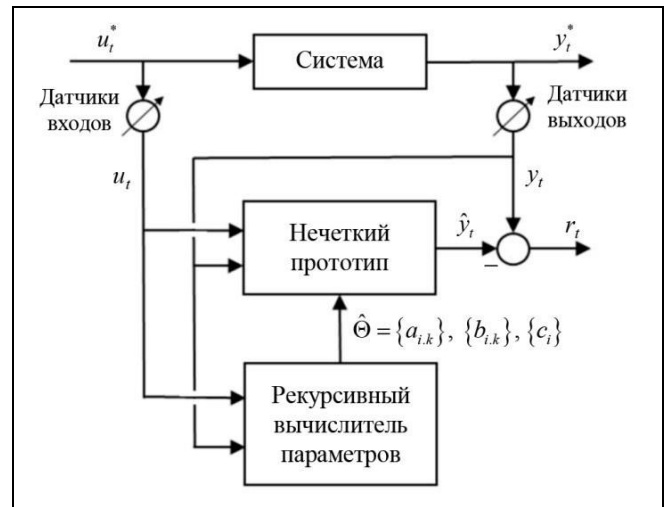


Рис. 5. Схема вычисления невязки на основе нечеткой модели

Отдельной, но весьма важной проблемой является нечеткая кластеризация пространства состояний системы [31]. Большинство алгоритмов нечеткой кластеризации основаны на оптимизации целевой функции $J(Z, U, V)$ c -средних, выполняемой следующим образом:

² При изложении далее этого подхода авторы отошли от принятой в ч. 2 обзора установки рассматривать только динамические системы с моделями (2) и (3) [30], исходя из того, что описываемый способ формирования симптомов применим и к указанным системам.

- представление данных в матричной форме

$$Z = \begin{bmatrix} z_{11} & \cdots & z_{1N} \\ \vdots & \ddots & \vdots \\ z_{n1} & \cdots & z_{nN} \end{bmatrix},$$

где n – размерность данных (число измерений, выполняемых за одно наблюдение), N – количество доступных наблюдений;

- определение матрицы нечетких разбиений $U = [\mu_{ik}]$, содержащей значения функции принадлежности, где i – номер измерения, k – номер кластера;

- определение центров так называемых прототипов $V = [v_1 \cdots v_{n_c}]$, т. е. точек, по расстоянию до которых оценивается удаленность кластера от текущего состояния системы.

Формула минимизации широко распространенной целевой функции c -средних, сформулированная автором монографии [32], выглядит следующим образом:

$$v_k = \arg \min J(Z, U, V) = \\ = \arg \min \sum_{k=1}^{n_c} \sum_{i=1}^N (\mu_{ik})^m (z_i - v_k)^T A (z_i - v_k)$$

с весовым показателем $m > 1$, где матрица A определяет форму кластера.

Для диагностирования отказов на фоне помех измерителей и неопределенностей моделирования требуется соблюдение принципа разделения помех, при котором на генератор невязок не влияют неопределенность моделирования и входные эффекты. Это может быть достигнуто с помощью обобщенных схем наблюдения, условий согласованности [33] или конкретных подходов к проектированию [34].

В статье [22] для дискретно-событийных систем, ведущих себя, как правило, случайным образом [35, 36] и описываемых моделью нечеткого конечного автомата, предлагается метод функционального диагностирования на основе использования математического аппарата нечеткой логики. Ключевым звеном в работе [22] является использование детерминизатора нечеткого конечного автомата, т. е. конечного автомата, одновременно описывающего как исправное, так и неисправное поведение исходного автомата.

По сравнению с алгебраическими подходами [37, 38] к детерминизации нечетких конечных автоматов, сохраняющими размерность исходных автоматов, рассматриваемый подход позволяет использовать при диагностировании дополнительную полезную информацию о степени уверенности

в осуществлении каждого из возможных переходов, что предоставляет потенциальную возможность для увеличения (при необходимости) глубины поиска неисправностей. Плата за это – значительная размерность таблицы переходов детерминизатора.

Схема диагностирования в статье [22] содержит ряд каналов (на единицу больше числа возможных неисправностей в системе). Каждый канал включает наблюдатель в виде детерминизатора нечеткого конечного автомата. Блок принятия решения на основе сравнения выходов дискретно-событийной системы и наблюдателей генерирует решение об исправности или неисправности системы.

Отличительной особенностью метода является то, что для его реализации не требуется предварительное формирование табличного описания средств диагностирования; все вычисления осуществляются непосредственно в процессе диагностирования с использованием компактных аналитических соотношений, что весьма важно для осуществления мониторинга.

В работе [39] представлено исследование по использованию нечеткой логики для диагностирования специфической системы с тремя резервуарами. Необходимо отметить, что этому объекту как тестовому посвящено множество публикаций, где рассматривались различные наблюдатели [40–42] (см. ч. 2 обзора [30]), уравнения паритетности с нейронечеткой идентификацией [43], невязки из физических нелинейных уравнений [44] (как в ч. 2 обзора [30], но уравнения нелинейные), оценка физических параметров с использованием нечетких нейронных сетей [45] и нечеткая модель, основанная на В-сплайне³ сети [46].

Автор работы [39] определяет систему локализации⁴ неисправностей (англ. *Fault Isolation System*, FIS) как четверку

$$FIS = \langle F, S, V, \phi \rangle,$$

где F – множество неисправностей; S – множество симптомов с двухуровневой $\{0, 1\}$ или трехуровневой $\{-1, 0, +1\}$ шкалой; V – множество значений симптомов из интервала $[0, 1]$ или $[-1 \dots +1]$ соответственно; ϕ – функция, определяемая как декартово произведение множеств F и S , т. е. множество всех пар (f_i, s_j) . Особенности вычисления значения симптома иллюстрирует

³ Сокращение от «базисный сплайн», т. е. сплайн с наименьшим носителем для задания степени.

⁴ Под локализацией здесь и далее подразумевается определение места отказа, возникновения неисправного состояния.

рис. 6, где r_j – невязка; $s_j \in S$ – симптом; $v_{ij} \in V$ – непрерывное значение j -го симптома для i -й неисправности; P – позитивная и N –, $N+$ – негативные области оценки симптома.

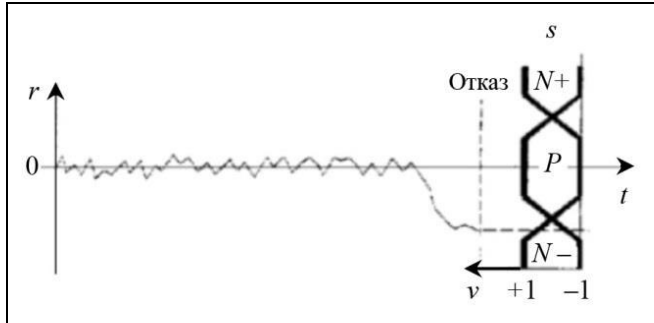


Рис. 6. Трехуровневая оценка симптома

Диагностирование выполняется по набору сформированных таким образом симптомов с использованием пороговых элементов и правил типа (3) для каждой неисправности. Например, для трехзначных симптомов правила по числу неисправностей имеют вид:

$$R_i : \text{IF } s_1 = 0, s_2 = +1 \text{ or } -1, \\ s_3 = +1 \text{ or } -1, \dots \text{ THEN fault } f_i.$$

В итоге FIS представляет собой таблицу, которая присваивает значение или подмножество значений симптомов s_j каждой неисправности f_i .

Для локализации неисправностей дополнительно вычисляются непрерывные значения симптомов v_{ij} по формулам, аналогичным (4), что позволяет учитывать значимость (степень принадлежности) симптомов.

Выполняемая таким образом интерпретация невязок позволяет учесть основную неопределенность, возникающую в процессе диагностирования, т. е. неопределенность симптомов s_j . Утверждается [40], что такой метод обеспечивает гораздо лучшую устойчивость сгенерированных диагнозов к шумам измерений по сравнению с алгоритмами, использующими проверку порогового значения невязки и логический вывод.

3. СТРУКТУРНЫЕ МЕТОДЫ ДИАГНОСТИРОВАНИЯ

Структурный анализ использует структурное представление модели, которое является грубым описанием, учитывающим только то, какие переменные входят в каждое уравнение, что позволяет эффективно и без численных трудностей анализи-

ровать задачи большой размерности. Однако ценой этого является то, что результаты являются слишком общими. Подход к диагностированию на основе структурного анализа был предложен в многочисленных работах [47–50].

Структурный анализ хорошо поддается компьютерной поддержке, доступной как в среде MATLAB, так и на языке Python.

Графо-теоретический алгоритм, называемый декомпозицией Дулмаджа – Мендельсона, является основополагающим для анализа диагностических свойств моделей [47, 49] и заключается в перестановке строк и столбцов так называемой структурной матрицы системы, содержащей специальные символы, отражающие вхождение переменных в различные уравнения математической модели этой системы.

3.1. Структурные модели

Пусть E – множество ограничений или уравнений в модели, а V – множество переменных, тогда структурная модель может быть представлена двудольным графом $G = (E \cup V, A)$, где A – множество ребер между узлами в двух множествах узлов E и V . Ребро $(e_i, v_j) \in A$ тогда и только тогда, когда переменная $v_j \in V$ встречается в модельном отношении $e_i \in E$. Распространенным способом визуализации структурной модели является так называемая матрица биадьюнктности графа, т. е. матрица, имеющая строки и столбцы, соответствующие ограничениям и переменным. Элемент в позиции (i, j) матрицы биадьюнктности пуст, если переменная v_j не входит в ограничение (уравнение) e_i . Матрица биадьюнктности, представляющая структуру модели, также называется просто структурной матрицей.

Набор переменных V может быть классифицирован как неизвестные $X \subset V$, известные $Z \subset V$ или ошибочные $F \subset V$ при выполнении условия $V = X \cup Z \cup F$. Для целей анализа наиболее важной будет часть структуры, связанная с неизвестными переменными, а подграф, включающий только ограничения (уравнения) и неизвестные переменные, называется сокращенным структурным графом.

Для пояснения и иллюстрации структурного анализа обычно применяется демонстрационный пример. В качестве такового далее используется заимствованная из монографии [49] математическая модель электродвигателя. Уравнения этой модели имеют вид:

$$\begin{aligned}
 e_1 : U &= h(R + f_R) + Lh' + K_a h\omega, \\
 e_2 : h &= \sqrt{\frac{T_m}{K_a \omega}}, \\
 e_3 : J\omega' &= T - b\omega, \\
 e_4 : T &= T_m - T_l - f_{\text{load}}, \\
 e_5 : y_h &= h + f_h, \\
 e_6 : y_\omega &= \omega + f_\omega, \\
 e_7 : y_T &= T + f_T, \\
 e_8 : \frac{dh}{dt} &= h', \\
 e_9 : \frac{d\omega}{dt} &= \omega',
 \end{aligned} \tag{6}$$

где использованы обозначения: U – напряжение питания; T_m – вращающий момент; T_l – момент расчетной нагрузки; h – ток обмотки двигателя; ω – угловая скорость вращения двигателя; R – номинальное сопротивление обмотки двигателя; L – индуктивность обмотки двигателя; K_a – коэффициент электромеханической связи; J – суммарная инерция вращения ротора и нагрузки; b – коэффициент трения.

Двигатель оборудован датчиками, измеряющими силу тока y_h , момент на валу y_T и угловую скорость y_ω . В качестве неисправностей рассматриваются неисправности всех датчиков f_h , f_ω и f_T , неисправность сопротивления обмотки f_R и неучтенная дополнительная нагрузка f_{load} на двигатель.

В данном случае и с учетом введения обобщенных обозначений наборы переменных V включают:

$$X = \{h, h', \omega, \omega', T, T_m, T_l\} = \{x_i\}_{i=1,7},$$

$$Z = \{y_h, y_\omega, y_T, y_U\} = \{y_j\}_{j=1,4},$$

$$F = \{f_R, f_h, f_\omega, f_T, f_{\text{load}}\} = \{f_k\}_{k=1,5}.$$

Кроме того, известны значения таких параметров: R, L, K_a, b, J .

Структурная матрица представляется в виде таблицы (e_i – i -я строка, v_j – j -й столбец, здесь $v_j = \{x_{k=1,7}, y_{l=1,4}, f_{m=1,5}\}$), ячейки которой заполняются следующими знаками: • – если переменная v_j входит в e_i непосредственно; D – если переменная v_j представлена ее производной; I – если переменная v_j представлена ее интегралом.

На рис. 7 представлена исходная структурная матрица, соответствующая модели (6).

Результат декомпозиции Дулмаджа – Мендельсона показан на рис. 8, где путем перестановки ограничений (уравнений) $e_1, \dots, e_5$ и переменных $x_1, \dots, x_4$ матрица биадьюнктности редуцированного структурного графа преобразована к блочной треугольной форме. На диагонали выделены три блока: первые два представляют собой точно определенную M^0 часть (число «внутренних» переменных равно числу уравнений), а третий – переопределенную M^+ часть (число «внутренних» переменных меньше числа уравнений) модели системы.

e_1	•	•	•					•								•
e_2	•					•										
e_3			•	•	•											
e_4					•	•	•					•				
e_5	•							•					•			
e_6			•						•					•		
e_7					•					•					•	
e_8	I	D														
e_9			I	D												
	x_1	x_2	x_3	x_4	x_5	x_6	x_7	f_1	f_2	f_3	f_4	f_5	y_1	y_2	y_3	y_4

Рис. 7. Структурная матрица системы (6)

e_4	•	•					•
e_2		•	•				
e_1			•	•	•		
e_8			I	D			•
e_3					•	•	•
e_9					I	D	
e_7							•
e_5			•				
e_6					•		
	x_7	x_6	x_1	x_2	x_3	x_4	x_5

Рис. 8. Сокращенная структурная матрица (диагностическая модель) системы (6)

В соответствии с рис. 8 недоопределенной части M^- у модели нет. Определенная часть позволяет однозначно определить входящие неизвестные, а переопределенная составляет основу для диагностирования.

3.2. Анализ диагностируемости отказов

Под диагностируемостью отказа понимают совокупность двух свойств:

- обнаруживаемость отказа – потенциальная возможность установить факт его наличия,
- локализуемость⁵ отказа – потенциальная возможность определения вида, места и времени возникновения обнаруженного отказа.

Набор обнаруживаемых и локализуемых (различаемых друг от друга) отказов в системе зависит от физических особенностей системы и ее взаимосвязей, а также от имеющихся измерений. Обнаруживаемость и локализуемость отказов – это свойства системы, накладывающие ограничения на производительность диагностирования, которая может быть достигнута любой системой диагностирования.

Без потери общности предполагается, что один отказ f_i может нарушить одно уравнение e_j модели. Если это не так, то путем введения дополнительных уравнений запись приводят к соответствующему виду.

⁵ Иногда этот термин используют с другим толкованием, а именно как предотвращение распространения влияния отказа [51]. Встречается также синоним «изолируемость» [52]. В данном случае понимание термина «локализуемость» увязано со сноской 4.

Пусть для модели M множество O_f обозначает множество всех наблюдений ее функционирования при наличии отказа f , в то время как O_{NF} – множество наблюдений при отсутствии отказов. Тогда со всей очевидностью отказ f в модели M является обнаруживаемым, если выполняется условие⁶ $O_f \setminus O_{NF} \neq \emptyset$.

Из этого вытекает утверждение, что структурно обнаруживаемый отказ f влияет на уравнение e_f , расположенное в переопределенной части M^+ модели

$$e_f \in M^+.$$

Применительно к модели (6) фрагмент декомпозиции Дулмаджа – Мендельсона принимает вид, показанный на рис. 9, где видно, что в переопределенную часть M^+ (выделена серым цветом) модели входят отказы $f_1, \dots, f_4$, в то время как отказ f_5 (f_{load}) не является обнаруживаемым ($e_{f_5} \in M^0$).

Считается, что отказ f_i локализуем (и отличим от отказа f_j), если существует наблюдение, соответствующее отказу f_i (позволяющее его обнаружить) и не связанное с отказом f_j . Сказанное записывается формулой $O_i \setminus O_j \neq \emptyset$ или

$$e_{f_i} \in (M \setminus \{e_{f_j}\})^+.$$

Обнаружение является необходимым условием локализации, т. е. из локализуемости отказа всегда следует его обнаруживаемость, но не наоборот:

$$e_{f_i} \in (M \setminus \{e_{f_j}\})^+ \Rightarrow e_f \in M^+.$$

Для иллюстрации локализуемости на примере (6) следует продолжить декомпозицию Дулмаджа – Мендельсона до результата, показанного на рис. 10.

Анализ рис. 10 показывает, что обнаруживаемые отказы f_1 и f_2 (f_R и f_h) входят в один переопределенный блок (обозначим его M_1^+), а отказы f_3 и f_4 (f_ω и f_T) – в другой переопределенный блок M_2^+ . Таким образом, указанные группы отказов изолируются друг от друга, но внутри групп они не изолируются.

⁶ Читается: множество наблюдений за вычетом наблюдений, соответствующих безотказному функционированию, не пусто.

e_4	•	•					•						•
e_2		•	•										
e_1			•	•	•			•					
e_8			I	D			•						
e_3					•	•	•						
e_9					I	D							
e_7							•					•	
e_5			•						•				
e_6					•					•			
	x_7	x_6	x_1	x_2	x_3	x_4	x_5	f_1	f_2	f_3	f_4	f_5	

Рис. 9. Структурная матрица (обнаруживаемость отказов) системы (6)

e_4	•	•				•							•
e_2		•	•										
e_1			•	•			•	•					
e_5			•						•				
e_8			I	D									
e_3					•	•	•						
e_7						•					•		
e_9					D		I						
e_6							•			•			
	x_7	x_6	x_1	x_2	x_4	x_5	x_3	f_1	f_2	f_3	f_4	f_5	

Рис. 10. Структурная матрица (локализуемость отказов) системы (6)

В целом же локализуемость одиночного отказа может быть вычислена путем анализа упорядоченных пар отказов по формуле

$$\bigcap_{f_j \in F} \left\{ (f_i, f_j) \mid e_{f_i} \in (M \setminus \{e_{f_j}\})^+ \right\}$$

с выполнением по одной декомпозиции Дулмаджа – Мендельсона для каждого отказа в модели. Таким образом, для модели в целом требуется выполнить столько декомпозиций Дулмаджа – Мендельсона, сколько анализируется одиночных отказов системы.

Для реализации потенциальных обнаруживаемости и локализуемости требуются соответствующие наблюдения. Методы структурного анализа позволяют проанализировать располагаемые наблюдения, например, показанные в модели (6) и на рис. 6. Кроме того, имеется возможность решать задачи размещения датчиков с достижением заданного или возможного уровня обнаруживаемости и локализуемости отказов, а также форми-

рования рекомендаций для построения диагностических тестов.

3.3. Обнаружение неисправностей на основе структурного анализа

Известны различные модификации структурного анализа [53] и описания алгоритмов их реализации [54]. В основном используются механизмы формирования и анализа структурно переопределенных множеств (СПМ) M^+ или, точнее, минимального СПМ, когда любое его подмножество не является СПМ. Последнее может быть достигнуто путем разбиения множества M на классы эквивалентности [16] по фигурирующим уравнениям e_i для исключения поиска одного и того же СПМ различными способами.

В результате каждый СПМ содержит набор уравнений, где по меньшей мере одно из них может использоваться для аналитической записи выражения для невязки. Набор уравнений с соответ-

ствующим разрешением относительно неизвестных используется для проектирования генератора невязки.

Для иллюстрации рассмотрим СПМ $M = \{e_1, e_3, e_5, e_7, e_8, e_9\}$ на примере электродвигателя, описываемого моделью (6). Этот набор имеет пять переменных в шести уравнениях, и любое из уравнений может быть использовано для формирования невязки. На рис. 11 показаны графы вычислений при выборе e_5 , e_1 или e_3 соответственно в качестве избыточного уравнения. Так, на рис. 11, а показан вычислительный поток от дифференцированной переменной ω' к недифференцированной переменной ω , т. е. интегрирование. На рис. 11, в – противоположный случай, от h к h' , т. е. дифференцирование. И, наконец, на рис. 11, б – оба случая, интегрирование и дифференцирование.

На схемах рис. 11 использованы обозначения: на левом краю входные переменные, вертикальная черта – разрешение указанного уравнения относительно последующей переменной.

Каждая из показанных схем обладает своими вычислительными особенностями и получаемой чувствительностью невязки к неисправным состояниям.

Получим итоговую расчетную формулу диагностирования на основе, например, схемы рис. 11, в. Эта формула определяет содержание любого реализуемого варианта алгоритма. Выполняя соответствующие схеме подстановки, вводя оператор дифференцирования по времени

$p = d/dt$ и выполняя простые преобразования, получаем «точное» выражение для невязки

$$r = \frac{Jp+b}{K_a} \left(\frac{U}{y_h - f_h} - Lp \ln(y_h - f_h) - R - f_R \right) - y_T + f_T.$$

Однако отказы f_h , f_R и f_T по определению недоступны для непосредственного измерения, и, таким образом, реализуемая формула невязки в окрестности точки $f_h = 0$, $f_R = 0$ и $f_T = 0$, приведенной к сигналу датчика момента на валу двигателя, принимает вид

$$r(f_h, f_R, f_T) = \frac{Jp+b}{K_a} \left(\frac{U}{y_h} - Lp \ln(y_h) - R \right) - y_T, \quad (7)$$

в соответствии с которым для обнаружения по крайней мере отдельных отказов f_h , f_R и f_T используются измеряемые сигналы U , y_h и y_T . При отсутствии отказов равенство (7) принимает нулевое значение. Значения параметров системы J , b , K_a , L и номинальное (безотказное) значение R должны быть известны.

Чувствительность к «постоянным» отказам ($pf_h = 0$, $pf_R = 0$, $pf_T = 0$) характеризуется значениями

$$\frac{\partial r}{\partial f_h} = \frac{Jp+b}{K_a} \left(Lp \frac{1}{y_h} - U \frac{1}{y_h^2} \right), \quad \frac{\partial r}{\partial f_R} = -\frac{b}{K_a}, \quad \frac{\partial r}{\partial f_T} = 1,$$

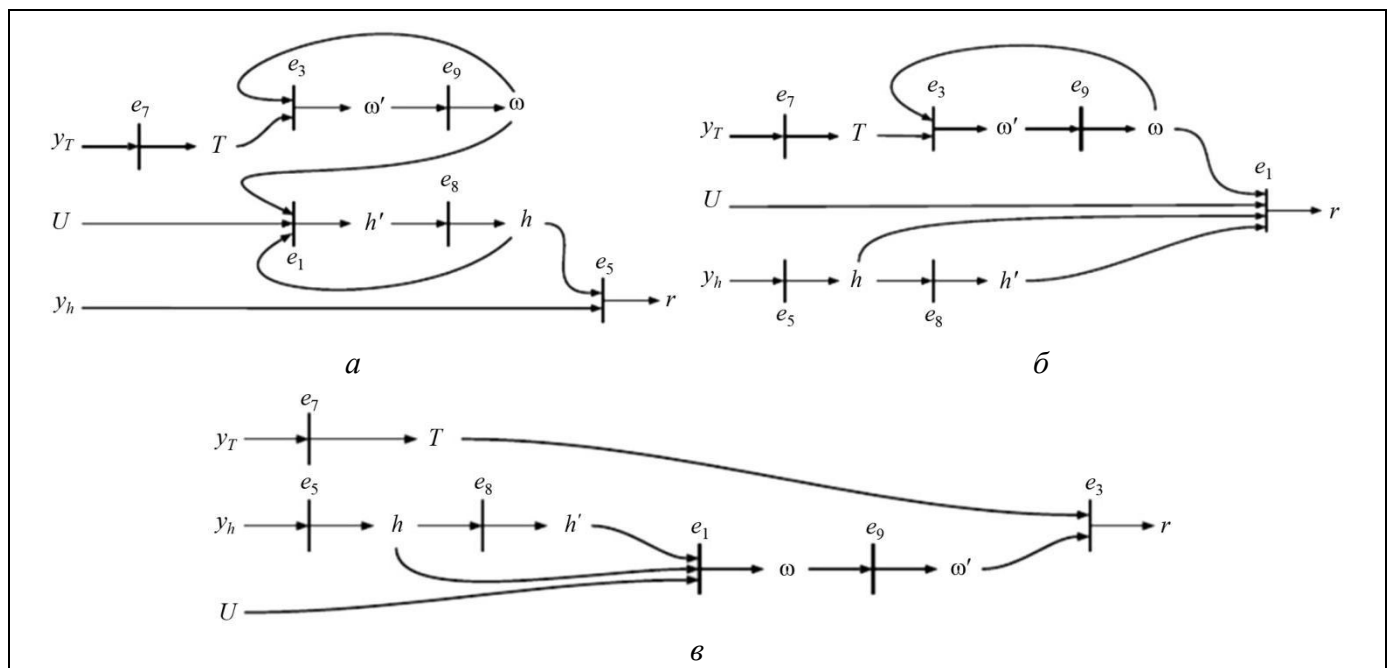


Рис. 11. Схемы СПМ с вариантами формирования невязок: а – интегральным, б – смешанным, в – дифференциальным

откуда следует, что в случае изменения сигнала датчика тока y_h обмотки двигателя в соответствии с нелинейным дифференциальным уравнением

$$JLp^2 \frac{1}{y_h} + bLp \frac{1}{y_h} - JUp \frac{1}{y_h^2} - bU \frac{1}{y_h^2} = 0$$

достаточное условие $\partial r / \partial f_h \neq 0$ чувствительности (7) к отказу f_h не будет удовлетворено.

Последнее обстоятельство указывает на то, что на определенных фазовых траекториях движения системы диагностирование может не выполняться⁷.

4. МЕТОДЫ ДИАГНОСТИРОВАНИЯ НА ОСНОВЕ АППАРАТА ТЕОРИИ МНОЖЕСТВ

При использовании моделей динамических систем в целях их мониторинга всегда существует некоторое несоответствие между результатами моделирования и реальным поведением этих систем. Такое явление обусловлено как пренебрежением некоторыми известными «несущественными» связями, так и наличием неизвестных или неточно известных связей в моделируемом объекте. Возникающие ошибки моделирования вносят неопределенность, которая чаще всего не выходит за пределы оцениваемых диапазонов.

Существует несколько способов учета неопределенности, связанной с моделью, в зависимости от того, находится ли она в параметрах (структурированная) или в структуре модели (неструктурированная). Группа наиболее развитых подходов, известных как активные, основана на генерировании невязок, нечувствительных к оговоренным неопределенностям, но чувствительных к неисправностям системы. В то же время существует другое семейство подходов, известных как пассивные, которые повышают надежность системы обнаружения неисправных состояний путем распространения неопределенности на значения невязок с созданием соответствующих адаптивных порогов⁸. Основополагающими работами, предлагающими второй из подходов, являются [55] во временной области и [27] в частотной области.

4.1. Исходные положения и постановка задачи

Идея подхода, основанного на принадлежности отказа к множеству, заключается в использовании

⁷ Это относится не только к структурным методам.

⁸ И хотя подход использует традиционные модели типа (1) [30], представленного в ч. 2 обзора, не они составляют ядро подхода.

геометрического набора для привязки неопределенных состояний. Известные геометрические наборы включают интервалы, эллипсоиды, многогранники, зонотопы и т. д. Из всех подходов с множествами зонотопы отличаются наименьшей вычислительной сложностью при реализации.

Излагаемый подход оперирует специфическим понятием и обозначениями зонотопа – математической конструкцией⁹ вида

$$Z = \langle c, H \rangle = \{c + Hz, \|z\|_\infty \leq 1\}, \quad (8)$$

где c и H – центр и сегментная матрица зонотопа. Запись (8) используется для формального обозначения интервала в многомерном пространстве. На плоскости это прямоугольник со сторонами $2h_1$ и $2h_2$, в трехмерном пространстве – параллелепипед с ребрами $2h_1$, $2h_2$, $2h_3$ и т. д.

Для зонотопов характерна линейность в смысле

$$\langle c_1, H_1 \rangle \oplus \langle c_2, H_2 \rangle = \langle c_1 + c_2, [H_1, H_2] \rangle,$$

$$L \langle c, H \rangle = \langle Lc, LH \rangle,$$

где $\oplus$ – сумма Минковского¹⁰; L – произвольная матрица соответствующего размера. Здесь $[H_1, H_2]$ обозначает формирование новой сегментной матрицы из исходных H_1 и H_2 , которое, как правило, приводит к «разрастанию» границ зонотопа.

В работе [56] для зонотопа $Z = \langle c, H \rangle$ предлагается оператор взвешенной редукции, обозначаемый $\downarrow_{q,w}^H$ и удовлетворяющий следующему свойству:

$$\langle c, \downarrow_{q,w}^H \rangle \subseteq \langle c, H \rangle,$$

где $q \geq n$ обозначает максимальное число столбцов $\downarrow_{q,w}^H$, а W – весовая матрица соответствующего размера. Использование такой редукции позволяет сокращать сегментную матрицу, определяющую «разброс» возможных значений в окрестности центра зонотопа.

Рассматривается линейная система с дискретным временем $t = 0, 1, 2, \dots$

$$x_{t+1} = Ax_t + Bu_t + D_w w_t, \quad y_t = Cx_t + D_v v_t, \quad (9)$$

⁹ Многогранник, представляющий собой сумму Минковского множества векторов.

¹⁰ Суммой Минковского множеств A и B линейного пространства V называется множество C , состоящее из сумм всевозможных векторов из A и B : $C = \{c | c = a + b, a \in A, b \in B\}$, см. статью [57].



где $x_t \in \mathbb{R}^{n_x}$, $u_t \in \mathbb{R}^{m_u}$ и $y_t \in \mathbb{R}^{n_y}$ – векторы состояний системы, известные входы и выходы соответственно; $w_t \in \mathbb{R}^{m_w}$ и $v_t \in \mathbb{R}^{m_v}$ – векторы возмущений и шума измерения соответственно.

Значения векторов w_t , v_t и начальное состояние x_0 ограничены зонотопами¹¹

$$w_t \in \langle 0, I_{m_w} \rangle, v_t \in \langle 0, I_{m_v} \rangle, x_0 \in X_0 = \langle c_0, H_0 \rangle \quad (10)$$

и на каждом такте полагаются неизвестными.

4.2. Зонотопический наблюдатель

Здесь кратко излагается один из подходов к диагностированию с использованием зонотопов [16, 49].

Для системы (9) вводится наблюдатель с коэффициентом усиления $G(k) \in \mathbb{R}^{n_x \times n_y}$:

$$\hat{x}_{t+1} = A\hat{x}_t + Bu_t + D_w w_t + G_t(y_t - C\hat{x}_t - D_v v_t). \quad (11)$$

Утверждается [58], что для наблюдателя (11) при $\forall t+1 \in N$ справедливо включение

$$\hat{x}_{t+1} \in \mathbb{S}_{t+1} = \langle c_{t+1}, H_{t+1} \rangle, \quad (12)$$

где центр и матрица сегментов определяются формулами

$$\begin{aligned} c_{t+1} &= (A - G_t C)c_t + Bu_t + G_t y_t, \\ H_{t+1} &= \left[(A - G_t C) \downarrow_{q,w}^H, D_w, G_t D_v \right]. \end{aligned} \quad (13)$$

Согласно формулам (13) включение (12) раскрывается следующим образом:

$$\begin{aligned} \hat{x}_{t+1} \in \langle c_{t+1}, H_{t+1} \rangle &= \\ &= \left((A - G_t C) \langle c_t, \downarrow_{q,w}^H \rangle \right) \oplus \\ &\oplus (B \langle u_t, 0 \rangle) \oplus (C_t \langle y_t, 0 \rangle) \oplus \\ &\oplus (D_w \langle 0, I_{m_w} \rangle) \oplus (-G_t D_v \langle 0, I_{m_v} \rangle). \end{aligned} \quad (14)$$

Формула (14) определяет потактовую трансформацию зонотопа, ограничивающего последующие значения вектора $\hat{x}_t$, $\forall t+1 \in N$.

4.3. Обнаружение неисправностей

Для обнаружения наличия неисправностей в системе, когда вместо формул (9) действуют уравнения

$$\begin{aligned} x_{t+1} &= Ax_t + Bu_t + D_w w_t + Q^u f_t^u, \\ y_t &= Cx_t + D_v v_t + Q^y f_t^y, \end{aligned} \quad (15)$$

где $f_t \in \mathbb{R}^{m_f}$ и $Q \in \mathbb{R}^{n_y \times m_f}$ – вектор и матрица влияния неисправностей актюатора (верхний индекс u) и датчика (верхний индекс y) соответственно, выполняются следующие действия:

- получение и хранение данных ввода – вывода $U_t = \{u_t\}$ и $Y_t = \{y_t\}$,
- вычисление множества оцененных состояний $Y_t = \{y_t\}$ с использованием уравнения (14) и интервальных ограничений (10)

$$X_0, u_t, y_t, w_t, v_t \rightarrow X_t^e,$$

- вычисление множества оцененных выходов Y_t^e с использованием второго уравнения (9) и интервальных ограничений (10)

$$X_t^e, v_t \rightarrow Y_t^e,$$

- вычисление пересечения множеств оцененных и измеренных выходов

$$X_t^{f=0} = Y_t^e \cap Y_t. \quad (16)$$

Если для каждого t пересечение (16) не пусто, то считается, что реальное значение выхода системы совпадает с его прогнозом, поэтому делается заключение об отсутствии неисправностей. Пустота такого пересечения является признаком наличия неисправностей в системе.

При этом условие чувствительности теста (16) к неисправностям датчика в уравнениях (15) выглядит как нестрогое неравенство

$$\sum_{t=0}^{\infty} (c_t^{rf})^T c_t^{rf} \geq \beta_s^2 \sum_{t=0}^{\infty} f_t^T f_t \quad (17)$$

при существовании таких матриц¹² $P \in S_{>0}^{n_x}$, $K \in \mathbb{R}^{n_x \times n_y}$, $N \in \mathbb{R}^{n_y \times n_y}$ и скаляров $\alpha > 0$, $\beta > 0$, что имеет место положительная определенность матрицы¹³

$$\begin{bmatrix} \alpha P & * & * \\ \alpha A^T P - \alpha C^T K^T & \alpha P + C^T N C & * \\ -\alpha Q^T K^T & Q^T N^T C & Q^T N Q - \beta^2 I_{m_f} \end{bmatrix} \succ 0.$$

Здесь c_t^{rf} – центр зонотопа $R_t = \langle c_t^r, H_t^r \rangle$, $\forall t \in N$, невязки $r_t^f = y_t - Cx_t$ с параметрами

$$c_t^r = M(y_t - Cc_t), \quad H_t^r = [MCH_t \quad MD_v],$$

¹² Первая из них положительно определенная симметрическая, $P \succ 0$.

¹³ Здесь звездочки использованы для упрощения записи симметрической матрицы.

¹¹ Первые два зонотопа называются унитарными.

где $M \in \mathcal{R}^{n_s \times n_s}$ – матрица усиления зонотопа невязки, в качестве варианта можно использовать равенства $G = P^{-1}K$, $M = N^2$.

Соответствующие доказательства и примеры использования зонотопов можно найти в [51, 59].

5. СТАТИСТИЧЕСКОЕ ДИАГНОСТИРОВАНИЕ

В основе подхода диагностирования по типу распознавания образов лежит метод статистических решений, заключающийся в следующем.

Рассматривается система с дискретным пространством признаков (измерений) размерности k с осями $\xi_1, \xi_2, \dots, \xi_k$ и числом градаций n по осям $\pi_1, \pi_2, \dots, \pi_n$ соответственно. Градации предусматривают как исправное техническое состояние системы, так и наличие различных неисправностей. Далее предполагается, что таких систем достаточно много¹⁴, и в зависимости от соответствующих признаков они разделяются на классы K_{ij} , $i = \overline{1, k}$, $j = \overline{1, n}$, включая варианты без неисправностей $K_{\text{испр}}$ или с различными неисправностями $K_{m\text{-я неисправ}}$, где $m \leq kn - 1$.

Основным параметром является эмпирическая (статистическая) оценка отношения правдоподобия, в одной из редакций (упрощенной, дающей «локальную» оценку) метода вычисляемая по формуле [60]

$$L(\xi) = \frac{f_l(\xi)}{f_h(\xi)} \approx \frac{r_l + 1}{r_h + 1} \frac{N_h + 2}{N_l + 2}, \quad (18)$$

где $f_l(\xi)$ и $f_h(\xi)$ – байесовы оценки вероятностей попадания системы из l -го или h -го классов соответственно в некоторую точку ξ (или окрестность ее); r_l и r_h – число систем l -го и h -го классов, попавших в точку ξ по априорной информации; N_l и N_h – число систем l -го или h -го классов обучающей выборки. Более точные, но сложные формулы с использованием плотностей вероятности можно найти в работах [60, 61].

Для принятия решения о принадлежности системы к тому или иному классу используется пороговое правило типа

- система исправна, если $L(\xi) \in K_{\text{испр}}$,
- у системы m -я неисправность, если $L(\xi) \in K_{m\text{-я неисправ}}$.

¹⁴ С учетом гипотезы об эргодичности случайных процессов [62], для априорных данных это число используемых для обучения систем, а для апостериорных данных – число разнесенных во времени наблюдений диагностируемой системы.

Этап обучения заключается в запоминании систем из обучающей выборки, а вычисление (18) производится при текущем осуществлении распознавания (диагностировании).

Таким образом, формально решение задачи диагностирования (прогнозирования) с помощью математического аппарата статистической классификации не имеет принципиальных отличий от решения других технических задач, требующих распознавания, но существует ряд специфических особенностей, которые необходимо учитывать. Основные этапы решения задачи включают:

- выбор модели распознавания, которая может быть детерминированной или вероятностной и зависит от того, насколько перемешаны множества, принадлежащие к разным классам;
- описание эталонов классов R_λ (исправен и разные неисправности) на основе априорной информации, оценка информативности (значимости для распознавания) параметров позволяет оптимизировать описание эталонов;
- сопоставление текущей информации о контролируемой системе с априорно заданными эталонами классов R_λ ;
- принятие решения о степени работоспособности системы или ее частей по данным контроля.

Основная особенность статистического подхода заключается в том, что для приемлемой результативности его применения требуется значительный объем как априорных (обучающих), так и апостериорных (текущих контрольных измерений) данных. Тем не менее, использование «локальной» оценки (18) позволяет решать задачи с числом параметров (используемых измерений) $k \leq 50$ и числом градаций (связано с количеством возможных неисправностей) по каждому параметру $\pi \leq 8$, т. е. речь идет о не слишком больших выборках, что принципиально для задач мониторинга.

Статистическое диагностирование на основе распознавания образов получило широкое распространение и развитие в отечественной научной литературе и технике, см., например, работы [63–65].

ЗАКЛЮЧЕНИЕ

Привлечение средств представления или описания динамических систем неклассического типа в целом значительно расширило возможности их диагностирования, прежде всего в смысле обхода традиционных для технических приложений трудностей: зашумленности реальных измерений и искажений, неизбежно вносимых в формальные модели. В результате были созданы и получили распространение системы диагностирования и мони-



торинга с принципиально новыми возможностями и областями практического применения.

В то же время этот путь таит в себе дополнительные проблемы: помимо усложнения применяемого аппарата получаемый результат может обладать непрозрачностью, расплывчатостью, неоднозначностью или чрезмерной общностью. Недостаточная конкретика может оказаться серьезным ограничением применения соответствующих подходов.

Четвертая часть обзора будет посвящена анализу новых подходов к диагностированию и объединению различных моделей и методов.

ЛИТЕРАТУРА

1. *Forssell, U., Ljung, L.* Closed-Loop Identification Revisited // *Automatica*. – 1999. – Vol. 35, no. 7. – P. 1215–1241.
2. *Haber, R. and Unbehauen, H.* Structure Identification of Non-linear Dynamic Systems – A Survey on Input/Output Approaches // *Automatica*. – 1990. – Vol. 26, no. 7. – P. 651–677.
3. *Isermann, R.* Fault-Diagnosis Systems. An Introduction from Fault Detection to Fault Tolerance. – Berlin, Heidelberg: Springer, 2006.
4. *Калявин В.П., Давыдов Н.А.* Надежность и диагностика автотранспортных средств. – СПб.: Элмор, 2014. [*Kalyavin, V.P., Davydov, N.A.* Reliability and Diagnostics of Motor Vehicles. – St. Petersburg: Elmore, 2014. (In Russian)]
5. *Палюх Б.В., Богатиков В.Н., Шпрекхер Д.М.* Диагностирование электромеханических систем на основе нейросетевых технологий // Программные продукты и системы. – 2015. – № 3. – С. 5–11. [*Palyukh, B.V., Shprekher, D.M., and Bogatikov, V.N.* Electromechanical Systems Diagnosis Based on Neural Network Technologies // *Software & Systems*. – 2015. – No. 3. – P. 5–11. (In Russian)]
6. *Пушкарев Д.О.* Применение нейросетевых экспертных систем для контроля, диагностики и прогнозирования технического состояния авиационных двигателей // Информационно-технологический вестник. – 2020. – № 3 (25). – С. 51–59. [*Pushkarev, D.O.* Application of Neural Network Expert Systems for Monitoring, Diagnostics and Forecasting the Technical Condition of Aircraft Engines // *Information Technology Bulletin*. – 2020. – No. 3 (25). – P. 51–59. (In Russian)]
7. *Кулида Е.Л., Лебедев В.Г.* Современные подходы к диагностике и прогнозированию технического состояния электромеханического привода летательного аппарата // Проблемы управления. – 2024. – № 3. – С. 3–19. [*Kulida, E.L., Lebedev, V.G.* Modern Approaches to Prognostics and Health Management of an Aircraft Electromechanical Actuator // *Control Sciences*. – 2024. – No. 3. – P. 3–15.]
8. *Vasilescu, M.-V., Epikhin, A.I., Toriya, T.G.* Diagnostics of the Technical Condition of a Boiler Plant Using Bayesian Networks // *Эксплуатация морского транспорта*. – 2023. – № 3. – С. 113–120. [*Vasilescu, M.-V., Epikhin, A.I., Toriya, T.G.* Diagnostics of the Technical Condition of a Boiler Plant Using Bayesian Networks // *Ekspluatatsiya morskogo transporta*. – 2023. – No. 3. – P. 113–120.]
9. *Лоскутов А. И., Столяров А. В., Ряхова Е.А.* Способ контроля технического состояния бортовой радиоэлектронной аппаратуры космических аппаратов на основе взаимодействия элементов распределенной системы технического диагностирования // Контроль. Диагностика. – 2021. – Т. 24, № 1 (271). – С. 44–50. [*Loskutov, A.I., Stolyarov, A.V., Ryakhova, E.A.* Method for Control of Technical State of Onboard Radioelectronic Equipment of Spacecrafts Based On Interaction of Elements of Distributed Technical Diagnosis System // *Testing. Control. Diagnostics*. – 2021. – Vol. 24, no. 1 (271). – P. 44–50. (In Russian)]
10. *Haykin, S.S.* Neural Networks and Learning Machines. Vol. 3. – Upper Saddle River, NJ: Pearson Education, 2009.
11. *Diaconescu, E.* The Use of NARX Neural Networks to Predict Chaotic Time Series // *WSEAS Trans. Comp. Research*. – 2008. – Vol. 3. – P. 182–191.
12. *Liu, G.P., Patton, R.J.* Eigenstructure Assignment for Control System Design. – New York: John Wiley & Sons, 1998.
13. *Medsker, L., Jain, L.C.* Recurrent Neural Networks: Design and Applications. – Cleveland: CRC Press, 1999.
14. *Круглов В.В., Борисов В.В.* Искусственные нейронные сети. Теория и практика. – М.: Горячая линия – Телеком, 2002. [*Kruglov, V.V., Borisov, V.V.* Iskusstvennyye neironnyye seti. Teoriya i praktika. – M.: Goryachaya liniya – Telekom, 2002. (In Russian)]
15. *Рейзлин В.И.* Численные методы оптимизации: учебное пособие. – Томск: Изд-во НИТПУ, 2013. [*Reizlin, V.I.* Chislennyye metody optimizatsii: uchebnoye posobie. – Tomsk: Izd-vo NITPU, 2013. (In Russian)]
16. *Krysander, M., Eslund, J., Nyberg, M.* An Efficient Algorithm for Finding Minimal Overconstrained Subsystems for Model-Based Diagnosis // *IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans*. – 2008. – Vol. 38, no. 1. – P. 197–206.
17. *Korbicz, J., Patan, K., Obuchowicz, A.* Dynamic Neural Network for Process Modelling in Fault Detection and Isolation Systems // *Applied Mathematics and Computer Science*. – 1999. – Vol. 9, no. 2. – P. 519–546.
18. *Dexter, A.L., Benouarets, M.* Model-Based Fault Diagnosis Using Fuzzy Matching // *IEEE Trans. on Sys. Man. and Cyber. Part A: Sys. & Humans*. – 1997. – Vol. 27, no. 5. – P. 673–682.
19. *Isermann, R.* On Fuzzy Logic Applications for Automatic Control, Supervision and Fault Diagnosis // *IEEE Trans. on Sys. Man. and Cyber. Part A: Sys. & Humans*. – 1998. – Vol. 28, no. 2. – P. 221–235.
20. *Полковникова Н.А., Курейчик В.М.* Нейросетевые технологии, нечёткая кластеризация и генетические алгоритмы в экспертной системе // Известия ЮФУ. Технические науки. – 2014. – № 7 (156). – С. 7–15. [*Polkovnikova, N.A., Kureichik, V.M.* Neural Network Technologies, Fuzzy Clustering and Genetic Algorithms in an Expert System // *Bulletin of SFedU. Technical sciences*. – 2014. – No. 7 (156). – P. 7–15. (In Russian)]
21. *Ахметханов Р.С., Дубинин Е.Ф., Куксова В.И.* Использование методов и моделей нечеткой логики в системах технической диагностики // Приводы и компоненты машин. – 2018. – № 1–2 (27). – С. 6–11. [*Akhmetkhanov, R.S., Dubinin, E.F., Kuksova, V.I.* Use of Fuzzy Logic Methods and Models in Technical Diagnostics Systems // *Drives and Machine Components*. – 2018. – No. 1–2 (27). – P. 6–11. (In Russian)]
22. *Шумский А.Е., Жирабок А.Н.* Метод функционального диагностирования дискретно-событийных систем на основе модели нечеткого конечного автомата // Проблемы управления. – 2025. – № 2. – С. 38–49. [*Shumsky, A.E., Zhirabok, A.N.* A Fault Diagnosis Method for Discrete-Event Systems Based on the Fuzzy Finite State Automaton Model // *Control Sciences*. – 2025. – No. 2. – P. 31–41.]

23. Ключкин А. К., Хаматов А. А. Синтез диагностической модели на основе аппарата нечеткой логики // *Авиакосмическое приборостроение*. – 2023. – № 7. – С. 34–41. [Klyuchkin A. K., Khamatov A. A. Sintez diagnosticheskoi modeli na osnove apparata nechetkoi logiki // *Aviakosmicheskoe priborostroenie*. – 2023. – № 7. – S. 34–41. (In Russian)]
24. Аблязов Э.К., Аблязов К.А. Оценка технического состояния портового подъемно-транспортного оборудования (ПТО) с использованием нечеткой логики // *Морские интеллектуальные технологии*. – 2024. – № 2-1 (64). – С. 242–251. [Ablyazov E.K., Ablyazov K.A. Otsenka tekhnicheskogo sostoyaniya portovogo pod'emno-transportnogo oborudovaniya (PTO) s ispol'zovaniem nechetkoi logiki // *Morskie intellektual'nye tekhnologii*. – 2024. – No. 2-1 (64). – S. 242–251. (In Russian)]
25. Fantuzzi, C., Rovatti, R. On the Approximation Capabilities of the Homogeneous Takagi-Sugeno Model // *Proceedings of the 5th IEEE International Conference on Fuzzy Systems*. – Vol. 2. – NJ: IEEE, Piscataway, 1996. – P. 1067–1072.
26. Rovatti, R. Takagi-Sugeno Models as Approximators in Sobolev Norms: The SISO Case // *Proceedings of the 5th IEEE International Conference on Fuzzy Systems*. – Vol. 2. – NJ: IEEE, Piscataway, 1996. – P. 1060–1066.
27. Takagi, T., Sugeno, M. Fuzzy Identification of Systems and Its Applications to Modeling and Control // *IEEE Transactions on Systems, Man and Cybernetics*. – 1985. – No. 1. – P. 116–132.
28. Jun, W., Shitong, W., Chung, F.-L. Positive and Negative Fuzzy Rule System, Extreme Learning Machine and Image Classification // *Int. J. of Machine Learning and Cybernetics*. – 2011. – No. 2 (4). – P. 261–271.
29. Graaff, A.J., Engelbrecht, A.P. Clustering Data in Stationary Environments with a Local Network Neighborhood Artificial Immune System // *Int. J. of Machine Learning and Cybernetics*. – 2012. – Vol. 3, no. 1. – P. 1–26.
30. Буков В.Н., Бронников А.М., Попов А.С., Шурман В.А. Методы мониторинга технического состояния систем в целях управления их избыточностью. Ч. 2. Классические модели // *Проблемы управления*. – 2025. – № 3. – С. 3–14. [Bukov, V.N., Bronnikov, A.M., Popov, A.S., Shurman, V.A. Technical Condition Monitoring Methods to Manage the Redundancy of Systems. Part II: Classical Models // *Control Sciences*. – 2025. – No. 3. – P. 2–11.]
31. *Diagnosis and Fault-Tolerant Control 2: From Fault Diagnosis to Fault-Tolerant Control* / Coord. by V. Puig and S. Simani. London: ISTE Ltd, Hoboken: John Wiley & Sons Inc, 2021.
32. Bezdek, J.C. *Pattern Recognition with Fuzzy Objective Function Algorithms*. – Berlin: Springer Science & Business Media, 2013. –
33. Chow, E.Y., Willsky, A.S. Analytical Redundancy and the Design of Robust Detection Systems // *IEEE Trans. Automatic Control*. – 1984. – Vol. 29, no. 7. – P. 603–614.
34. Duan, G., How, D., Patton, R. Robust Fault Detection in Descriptor Systems via Generalised Unknown Input Observers // *Int. J. Systems Science*. – 2002. – Vol. 32, no. 2. – P. 7724–7729.
35. Zaytoon, J., Lafortune, S. Overview of Fault Diagnosis Methods for Discrete Event Systems // *Annual Reviews in Control*. – 2013. – Vol. 37, no. 2. – P. 308–320.
36. Cassandra, Ch., Lafortune, S. *Introduction to Discrete Event Systems*. 2nd edition. – New York: Springer Science+Business Media, LLC, 2008. – 770 p.
37. Журабков А.Н., Калинина Н.А., Шумский А.Е. Метод мониторинга поведения человека-оператора в человеко-машинных системах // *Известия РАН. Теория и системы управления*. – 2018. – № 3. – С. 98–107. [Zhirabok, A.N., Kalinina, N.A., Shumskii, A.E. Technique of Monitoring a Human Operator's Behavior in Man-Machine Systems // *Journal of Computer and Systems Sciences International*. – 2018. – Vol. 57, no. 3. – P. 443–452.]
38. Журабков А.Н., Калинина Н.А., Шумский А.Е. Метод функционального диагностирования недетерминированных конечных автоматов // *Известия РАН. Теория и системы управления*. – 2020. – № 4. – С. 62–72. [Zhirabok, A.N., Kalinina, N.A., Shumskii, A.E. Method for the Functional Diagnosis of Nondeterministic Finite State Machines // *Journal of Computer and Systems Sciences International*. – 2020. – Vol. 59, no. 4. – P. 565–574.]
39. Kościelny, J.M. Application of Fuzzy Logic for Fault Isolation in a Three-Tank System // *Proc. 14th IFAC World Congress*. Beijing. – 1999. – Vol. 7. – P. 73–78.
40. Hengy, D., Frank, P.M. Component Failure Detection Using Local Second-Order Observers // *Fault Detection & Reliability* / Ed. by M.G. Singh. – Oxford: Pergamon Press, 1987. – P. 147–154.
41. Frank, P.M. Fault Diagnosis in Dynamic Systems via State Estimation – A Survey // In: *System Fault Diagnosis, Reliability and Related Knowledge-Based Approaches* / Ed. by E.A. Tzafestas. – Dordrecht: Dr. Reidel Publ. Comp., 1987. – Vol. 1. – P. 35–98.
42. Koenig, D., Nowakowski, S., Cecchin, T. An Original Approach for Actuator and Component Fault Detection and Isolation // *Proceedings of the CAA Symposium on Fault Detection, Supervision and Safety for Technical Processes SAFEPROCESS '97*. – Kingston upon Hull, 1997. – Vol. 1. – P. 95–105.
43. Garcia, F.L., Izquierdo, V., de Miguel, L., Peran, J. Fuzzy Identification of Systems and Its Applications to Fault Diagnosis Systems // *Proceedings of the CAA Symposium on Fault Detection, Supervision and Safety for Technical Processes SAFEPROCESS '97*. – Kingston upon Hull, 1997. – Vol. 2. – P. 705–712.
44. Koscielny, J.M., Sedziak, D., Sikora, A. The “DIAG” System for Fault Detection and Isolation in Industrial Processes // *Proceedings of the CAA Symposium on Fault Detection, Supervision and Safety for Technical Processes SAFEPROCESS '94*. – Helsinki, 1994. – Vol. 2. – P. 790–795.
45. Han, Z., Frank, P.M. Physical Parameter Estimation Based FDI with Neural Networks // *Proceedings of the CAA Symposium on Fault Detection, Supervision and Safety for Technical Processes SAFEPROCESS '97*. – Kingston upon Hull, 1997. – Vol. 1. – P. 294–299.
46. Benkbedda, H., Patton, R. Information Fusion in Fault Diagnosis Based on B-spline Networks // *Proceedings of the CAA Symposium on Fault Detection, Supervision and Safety for Technical Processes “SAFEPROCESS'97”*. – Kingston upon Hull, 1997. – Vol. 2. – P. 681–686.
47. Cassar, J.P., Staroswiecki, M. A Structural Approach for the Design of Failure Detection and Identification Systems // *IFAC Proceedings Volumes*. – 1997. – Vol. 30, no. 6. – P. 841–846.
48. Ploix, S., Désinde, M., Touaf, S. Automatic Design of Detection Tests in Complex Dynamic Systems // *IFAC Proceedings volumes*. – 2005. – Vol. 38, no. 1. – P. 478–483.
49. *Diagnosis and Fault-tolerant Control 1: Data-Driven and Model-Based Fault Diagnosis Techniques* / Coord. by V. Puig and S. Simani. – London: ISTE Ltd, Hoboken: John Wiley & Sons Inc, 2021.
50. Копкин Е.В., Иваню А.Ю. Структурно-стохастическая вычислительная модель контроля технического состояния бортовых систем космических аппаратов // *Изв. Тульского*



- государственного университета. Технические науки. – 2022. – № 10. – С. 9–14. [Kopkin, E.V., Ivanju, A.Ju. Strukturno-stohasticheskaja vychislitel'naja model' kontrolya tehničeskogo sostojanija bortovyh sistem kosmicheskikh apparatov // Izv. Tul'skogo gosudarstvennogo universiteta. Tehničeskije nauki. – 2022. – No. 10. – P. 9–14. (In Russian)]
51. *Рекомендации по стандартизации Р 50.1.048-2004*: Информационно-телекоммуникационные игровые системы. Термины и определения. – Москва: Госстандарт России, 2004. – 19 с. [Rekomendatsii po standartizatsii R 50.1.048-2004: Informatsionno-telekommunikatsionnye igrovye sistemy. Terminy i opredeleniya. – Moskva: Gosstandart Rossii, 2004. – 19 s. (In Russian)]
 52. Брускин С.Н., Дружаев А.А., Марон А.И., Марон М.А. Эффективные методы построения алгоритмов поиска неисправностей в информационных системах // Интеллектуальные системы в производстве. – 2017. – № 15 (3). – С. 88–93. [Bruskin, S.N., Druzhaev, A.A., Maron, A.I., Maron, M.A. Jeffektivnye metody postroenija algoritmov poiska neispravnostej v informacionnyh sistemah // Intel'ktual'nye sistemy v proizvodstve. – 2017. – No. 15 (3). – P. 88–93. (In Russian)]
 53. Armengol, J., Bregyn, A., Escobet, T., et al. Minimal Structurally Overdetermined Sets for Residual Generation: A Comparison of Alternative Approaches // IFAC Proceedings Volumes. – 2009. – Vol. 42, no. 8. – P. 1480–1485.
 54. Krysanter, M., Eslund, J., Frisk, E. A Structural Algorithm for Finding Testable Sub-models and Multiple Fault Isolability Analysis // Proceedings of 21st International Workshop on Principles of Diagnosis (DX-10). – Portland, USA, 2010. – DOI: <https://doi.org/10.36001/phmconf.2010.v2i1.1940>
 55. Horak, D. Failure Detection in Dynamic Systems with Modeling Errors // AIAA J. of Guidance, Control and Dynamics. – 1988. – Vol. 11, no. 6. – P. 508–516.
 56. Song, Y., Sai, M.P.D., Yu, H. Zonotope-Based Nonlinear Model Order Reduction for Fast Performance Bound Analysis of Analog Circuits with Multiple-Interval-Valued Parameter Variations // Proceedings of IEEE Design, Automation & Test in Europe: Conference & Exhibition. – Dresden, Germany, 2014. – DOI: 10.7873/DATE2014.024
 57. Панюков А.В. Представление суммы Минковского для двух полиэдров системой линейных неравенств // Вестник ЮУрГУ. Серия: Математическое моделирование и программирование. – 2012. – № 14. – С. 108–119. [Panjukov, A.V. Predstavlenie summy Minkovskogo dlja dvuh polijedrov sistemoy linejnyh neravenstv // Vestnik JuUrGU. Serija: Matematicheskoe modeli-rovanie i programmirovanie. – 2012. – No. 14. – P. 108–119. (In Russian)]
 58. Combastel, C. Zonotopes and Kalman Observers: Gain Optimality under Distinct Uncertainty Paradigms and Robust Convergence // Automatica. – 2015. – No. 55. – P. 265–273.
 59. Himmelblau, D.M. Fault Diagnosis in Chemical and Petrochemical Processes. – Amsterdam: Elsevier, 1978.
 60. Гаскаров Д.В., Голинкевич Т.А., Мозгалеvский А.В. Прогнозирование технического состояния и надежности радиоэлектронной аппаратуры / Под ред. Голинкевича Т.А. – М.: Советское радио, 1974. [Gaskarov, D.V., Golinkevich, T.A., Mozgalevskij, A.V. Prognozirovanie tehničeskogo sostojanija i nadezhnosti radiojelektronnoj apparatury / Pod red. Golinkevicha T.A. – M.: Sovetskoe radio, 1974. (In Russian)]
 61. Мозгалеvский А.В., Койда А.Н. Вопросы проектирования систем диагностирования. – Л.: Энергоатомиздат, Ленингр. отделение, 1985. [Mozgalevskij, A.V., Kojda, A.N. Voprosy proektirovanija sistem diagnostirovaniya. – L.: Jenergoatomizdat, Leningr. otdelenie, 1985. (In Russian)]
 62. Вентцель Е.С. Теория вероятностей. – М.: КноРус, 2010. [Ventcel', E.S. Teorija verojatnostej. – M.: KnoRus, 2010. (In Russian)]
 63. Обозов А.А., Таричко В.И. Совершенствование технического диагностирования автомобильных двигателей на основе применения методов статистической теории распознавания образов // Двигателестроение. – 2013. – № 3 (253). – С. 16–22. [Obozov, A.A., Tarichko, V.I. Sovershenstvovanie tehničeskogo diagnostirovanija avtomobil'nyh dvigatelej na osnove primenija metodov statističeskoj teorii raspoznavanija obrazov // Dvigatelistroenie. – 2013. – No. 3 (253). – P. 16–22. (In Russian)]
 64. Усик Б.А., Лучин А.В., Обыночный И.А. Выбор параметров для диагностирования преобразовательных устройств с резервированными элементами методами распознавания образов // Известия Института инженерной физики. – 2022. – № 3 (65). – С. 14–16. [Usik, B.A., Luchin, A.V., Obynochnyj, I.A. Vybory parametrov dlja diagnostirovaniya preobrazovatel'nyh ustroystv s rezervirovannymi jelementami metodami raspoznavanija obrazov // Izvestija Instituta inzhenernoj fiziki. – 2022. – No. 3 (65). – P. 14–16. (In Russian)]
 65. Шамолин М.В. Обобщенная задача контроля в задачах диагностики // Итоги науки и техники. Современная математика и ее приложения. Тематические обзоры. – 2022. – Т. 209. – С. 117–126. [Shamolin, M.V. Obobshhennaja zadacha kontrolya v zadachah diagnostiki // Itogi nauki i tehniki. Sovremennaja matematika i ee prilozhenija. Tematicheskie obzory. – 2022. – Vol. 209. – P. 117–126. (In Russian)]

Статья представлена к публикации членом редколлегии
В. Г. Лебедевым.

Поступила в редакцию 13.11.2024,
после доработки 21.08.2025.
Принята к публикации 27.08.2025.

Буков Валентин Николаевич – д-р техн. наук, АО «Бортовые аэронавигационные системы», г. Москва,
✉ v_bukov@mail.ru,
ORCID iD: <https://orcid.org/0000-0002-5194-8251>

Бронников Андрей Михайлович – д-р техн. наук, МГТУ им. Н.Э. Баумана, ✉ bronnikov_a_m@mail.ru,
ORCID iD: <https://orcid.org/0009-0009-1216-3521>

Попов Александр Сергеевич – канд. техн. наук, ВУНЦ ВВС «Военно-воздушная академия им. проф. Н. Е. Жуковского и Ю. А. Гагарина», г. Воронеж, ✉ saga30@yandex.ru,

Шурман Владимир Александрович – АО «НИИ авиационного оборудования», г. Жуковский, ✉ shurman@niiao.ru

© 2025 г. Буков В. Н., Бронников А. М., Попов А. С., Шурман В. А.



Эта статья доступна по лицензии Creative Commons «Attribution» («Атрибуция») 4.0 Всемирная.

TECHNICAL CONDITION MONITORING METHODS TO MANAGE THE REDUNDANCY OF SYSTEMS.

PART III: Nonclassical Models in Fault Diagnosis

V. N. Bukov*, A. M. Bronnikov**, A. S. Popov***, and V. A. Shurman****

*Airborne Aeronautical Systems, Moscow, Russia

**Bauman Moscow State Technical University, Moscow, Russia

***Zhukovsky–Gagarin Air Force Academy, Voronezh, Russia

****Institute of Aircraft Equipment, Zhukovsky, Russia

*✉ v_bukov@mail.ru, **✉ bronnikov_a_m@mail.ru, ***✉ saga30@yandex.ru, ****✉ shurman@niiio.ru

Abstract. Redundancy management in a technical system involves a monitoring procedure to reconfigure the system as needed. The four-part survey presents modern diagnosis methods for dynamic systems as an integral function of monitoring. Part III is devoted to diagnosis methods employing neural networks, fuzzy models, structural models, set-based models, and a statistical approach. The fundamentals of creating and training neural networks to perform diagnostic functions are considered. The approach with fuzzy models is described, including general modeling rules and the features of their use in diagnosis tasks. The approach with structural models is demonstrated, including its features in failure detection. The fundamentals of set-theory methods, particularly the formalism of zonotopes, are presented. Finally, the approach based on statistical pattern recognition is briefly discussed.

Keywords: artificial neuron, neural network, fuzzy models, membership function, fuzzy clustering, structural models, Dulmage–Mendelsohn decomposition, diagnosability, zonotopes, gradations in discrete feature space, empirical likelihood ratio estimate.

КЛАССИФИКАЦИЯ ПРОЦЕССОВ ДЕЯТЕЛЬНОСТИ

Г. Н. Калянов*, Д. А. Новиков**

***Институт проблем управления им. В. А. Трапезникова РАН, г. Москва

*✉ kalyanov@ipu.ru, **✉ novikov@ipu.ru

Аннотация. В настоящее время продолжает развиваться методология – теория об общих закономерностях организации деятельности и управления ею. Статья посвящена введению системы классификаций процессов, позволяющей для любого процесса определять его место в этой системе и устанавливать наличие взаимосвязей с другими процессами. Выделен ряд классификационных признаков и предложена классификационная схема элементарных процессов на их основе. Рассмотрено понятие комплексного процесса и продемонстрировано его разбиение на элементарные процессы в соответствии с классификационной схемой. Приведен пример комплексного процесса ремонта автотранспортного средства, представлены его функциональная и поведенческая схемы, а также детально описаны соответствующие элементарные процессы различных типов, а именно: процессы его жизненного цикла (поддержки), организационно-управляющий процесс, а также информационные и вычислительные процессы.

Ключевые слова: процесс, деятельность, методология, классификация, жизненный цикл.

ВВЕДЕНИЕ

В последние годы в рамках междисциплинарных исследований сформировалось новое научное направление, связанное с унификацией описания и формализацией различных видов человеческой деятельности (*деятельность* – целенаправленная активность человека). Данное направление, посвященное разработке единой теории деятельности в виде совокупности общих и универсальных моделей, представлено трудами А. М. Новикова и Д. А. Новикова и их коллег [1–5]. Соответствующее учение (теория) об общих закономерностях организации деятельности получило название *методологии*.

Методология включает общую методологию (методологию элементарной деятельности) [1], методологии различных видов человеческой деятельности (частные методологии: методологию научной [5], практической (в том числе методологию управления [4] как ее разновидность), образовательной, художественной и игровой деятельности) [2, 6], а также методологию комплексной деятельности, обобщающую перечисленные методологии на случай «сложной» деятельности. В монографии [1] определено понятие комплексной дея-

тельности как «целенаправленной активности человека, обладающей нетривиальной внутренней структурой, с множественными и/или изменяющимися субъектом, технологией, ролью предмета деятельности в его целевом контексте». Безусловно, вышеперечисленные виды человеческой деятельности не являются исчерпывающими, в частности авторы работ [1–5] упоминают также философскую и религиозную деятельность. Этот список, наверное, может быть расширен и далее.

Поскольку методология рассматривается как «учение об организации деятельности» [2], необходимо пояснить, что собой представляет организация. В Философском энциклопедическом словаре [7] приведены три значения данного понятия, для целей данной работы воспользуемся следующим: организация – «совокупность процессов или действий, ведущих к образованию и совершенствованию взаимосвязей между частями целого». Другими словами, «организовать деятельность означает упорядочить ее в целостную систему с четко определенными характеристиками, логической структурой и процессом ее осуществления» [2]. В монографии [1] процессная структура деятельности рассматривается как неотъемлемая ее часть (наряду с логической и причинно-следственными структурами).

1. ПРОЦЕССНАЯ СТРУКТУРА ДЕЯТЕЛЬНОСТИ И ТЕОРИЯ ПРОЦЕССОВ

Понятие «процесс» используется для описания природных явлений, а также широко распространено в различных областях человеческой деятельности. Можно говорить о физическом, химическом, техническом, технологическом, вычислительном, организационном и многих других процессах. Семантика процесса в этих контекстах разная, но есть и общие черты: длящийся во времени характер и определенная внутренняя структура. Считается, что само слово «процесс» русский язык заимствовал из немецкого во времена Петра I. Немецкое *prozess* произошло от латинского *procedere* – «выходить; продвигаться» (*pro* – «вперед, для, за, вместо» + *cedere* – «идти, ступать»).

Наиболее общее определение процесса дается в Философском энциклопедическом словаре [7]: «процесс – категория философии, характеризующая совокупность необратимых, взаимосвязанных, длительных изменений, как спонтанных, так и управляемых, как самоорганизованных, так и организуемых, результатом которых является некое новшество или нововведение (новые морфологические формы организмов, новые разновидности, социальные, научные, культурные и пр. инновации)».

Процессы можно разделить на *природные* (естественные) и *искусственные* (целенаправленные). Первые протекают в соответствии с законами природы независимо от чьей-либо воли; примерами являются физические, химические, биологические и т. п. процессы. Целенаправленные процессы являются предметом, результатом или содержанием той или иной человеческой деятельности.

Изучением процессов различных категорий занимаются соответствующие разделы науки, простейшая схема их соотношения приведена на рис. 1.

Далее в настоящей работе процесс рассматривается исключительно в контексте методологии человеческой деятельности, т. е. природные процессы не исследуются.

Национальные и международные стандарты (ГОСТ Р ИСО 9000, ISO/IEC 12207, ISO/IEC 15288) определяют *процесс* как совокупность взаимосвязанных или взаимоувязанных действий, преобразующая входы в выходы [8–10]. Однако данное определение отражает только функциональный и информационный аспекты процесса и никак не касается его поведенческого аспекта.

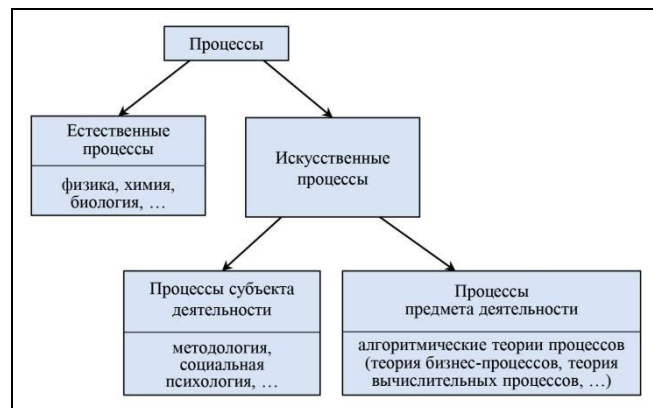


Рис. 1. Процессы как объект исследования

Для целей данной работы будем использовать следующее определение, восходящее к определениям, данным в современном энциклопедическом словаре и Большой советской энциклопедии («последовательная смена явлений, состояний в развитии чего-нибудь» или «совокупность последовательных действий для достижения какого-либо результата»): **процесс** – совокупность последовательных (последовательность) состояний объекта и функций, обеспечивающих переходы между этими состояниями.

Отметим, что функциональная часть процесса представляет собой иерархическую древовидную структуру и разбивается на функции и операции [11, 12].

Функция – набор операций, сгруппированных по определенному признаку и обеспечивающих в совокупности переход объекта в новое состояние.

Операция – элементарное (неделимое) действие, выполняемое над объектом или его элементом (структурная единица действия; способ, которым осуществляется действие в определенных условиях).

Неотъемлемой частью любого процесса являются его входы, выходы и ресурсы.

Входы процесса – объекты, которые поступают внутрь процесса, чтобы претерпеть изменения и преобразования. Природа поступающих объектов достаточно разнообразна. На вход может поступать нечто материальное (сырье, полуфабрикаты) или нематериальное (информация (данные), документы, сообщения, запросы, услуги).

Выходы процесса – то, во что были преобразованы объекты, которые поступили на вход. Иными словами, это результаты процесса. Важно подчеркнуть, что внутри процесса могут рождаться новые объекты или процессы.

Ресурсы процесса – это материальные или нематериальные объекты, необходимые для поддер-



Таблица 1

**Классификация видов деятельности
по целевому объекту**

Вид деятельности	Целевой объект
Игровая	Игра
Образовательная	Человек (образование)
Практическая	Товар, услуга
Научная	Научный результат
Художественная	Произведение искусства
Философская	Научный результат
Религиозная	Человек (вера)

• инжиниринг (проектирование) и реинжиниринг процессов, а также рассмотрение альтернативных решений вышеназванных задач;

- анализ и верификация процессов;
- автоматизация процессов.

При этом первоочередной задачей в рамках теории процессов является создание системы их классификации, позволяющей для любого процесса определять его место в системе и устанавливать наличие связей между классами процессов.

Классификации процессов посвящено много работ специалистов в области управления и организации производства различных отраслей экономики [13–22]. Одной из наиболее полных является классификационная структура бизнес-процессов, предложенная в работе [14] и использующая в качестве оснований следующие шесть признаков: наличие связей с внешней средой (внутренние, внешние), зависимость от предметной области (технологические, организационно-деловые), отношение к основной продукции (основные, вспомогательные, бизнес-процессы управления, бизнес-процессы сети, производственные, административные), масштаб (организация, отдел), стадия производственного цикла товаров (процессы товародвижения, процессы подготовки производства, процессы инфраструктуры), вид деятельности (непосредственно обеспечивающие выпуск продукции процессы, процессы планирования и управления, ресурсные процессы, процессы преобразования, метапроцессы, ориентированные на продукт процессы).

Однако предложенные классификации имеют ряд общих недостатков, а именно:

- отсутствие унификации, узкая направленность, ориентация на конкретные отрасли, области и т. п.;
- ориентация исключительно на экономические категории, затрудняющая использование таких

жания процесса. Как и в случае входов процесса, природа этих объектов крайне разнообразна. В качестве ресурсов могут рассматриваться оборудование и инструменты, персонал, информация, финансы, пространство (например, офис), инфраструктура, информационные системы, алгоритмы, технологии и многое другое. Отметим, что одни и те же объекты в различных процессах могут быть как входами, так и ресурсами или выходами.

В настоящее время вопросам исследования процессов различных категорий посвящено значительное число работ. Однако большинство авторов уделяют свое внимание лишь двум категориям процессов – *вычислительным* [13] и *организационно-управляющим* (или *бизнес-процессам*) [11]. Для данных категорий процессов были разработаны соответствующие теории, для других же категорий процессов в отдельных случаях имеются лишь разрозненные модели и методы их исследований. При этом существующие подходы и полученные в их рамках результаты не позволяют единообразно описывать процессы создания и поведения современных сложных систем. Поэтому актуальной является как задача создания соответствующих теорий для других категорий процессов, так и создание общей теории процессов, интегрирующей и обобщающей перечисленные частные теории.

Если исходить из классификации деятельности по целевой направленности (а именно, игра – учение – труд), то можно говорить:

- о методологии игровой деятельности;
- методологии учебной (образовательной) деятельности;
- методологии трудовой, профессиональной деятельности, которую, в свою очередь, можно подразделить:
 - на практическую деятельность как в сфере материального, так и духовного производства;
 - специфические формы профессиональной деятельности: философскую, научную, художественную, религиозную [2].

В табл. 1 приведены виды деятельности в соответствии с их целевым объектом (ЦО).

В связи с вышесказанным объектом исследования общей теории процессов являются процессы (комплексной) деятельности, а предметом исследования – общие закономерности организации процессов деятельности и управления такими процессами.

Также отметим, что создаваемая общая теория процессов должна обеспечить систематизированный базис для решения таких задач, как:

- моделирование процессов;

классификаций при унификации описания и формализации различных видов человеческой деятельности;

– отсутствие комплексной интеграции с предметом и субъектом деятельности.

2. ОСНОВЫ КЛАССИФИКАЦИИ ПРОЦЕССОВ

Как показано на рис. 1, искусственные процессы (процессы деятельности) разбиваются на две категории – процессы субъекта деятельности и процессы предмета деятельности.

Процессы субъекта деятельности (элементарной или комплексной), общие для всех видов деятельности, приведены в табл. 2 (см. табл. 1 в монографии [1]). Фактически табл. 2 описывает жизненный цикл деятельности как сложной системы, эволюционирующей во времени.

Дальнейшая детализация строится в рамках конкретной деятельности. Процессные модели комплексной деятельности подробно рассматриваются в пятой главе монографии [1] и не являются предметом исследования данной работы.

Соотношение процессов субъекта и предмета деятельности рассматривается во второй главе монографии [1].

В основе предлагаемой классификации **процессов предмета деятельности** лежит их разбиение на следующие четыре группы (рис. 2).

1. Жизненный цикл (ЖЦ) процесса, включающий в себя:

- концепцию, проектирование;
- производство (создание, организацию);
- тестирование;
- применение;
- поддержку, содержание, модернизацию;
- утилизацию.

Дальнейшая детализация процессов рассматриваемой группы строится в рамках конкретной практической деятельности и включает минимально еще два уровня.

2. Вид деятельности в соответствии с частными методологиями деятельности, отражающими специфику каждой из деятельностей в рамках стадий ЖЦ ЦО, а именно:

- игровая деятельность,
- образовательная деятельность,
- практическая деятельность,
- научная деятельность,
- художественная деятельность,
- философская деятельность,
- религиозная деятельность.

Для практической, образовательной, научной и художественной деятельности дальнейшая детализация может осуществляться, например, на основе Общероссийского классификатора видов экономической деятельности [23] и включать минимально еще четыре уровня.

Таблица 2

Фазы, стадии и этапы комплексной деятельности

Фаза	Стадия	Этап
Проектирование	I. Фиксация спроса и осознание потребности	1. Фиксация спроса и осознание потребности
	II. Целеполагание и структурирование целей и задач	2. Создание логической модели
	III. Формирование технологии	3. Проверка готовности технологии и достаточности ресурсов
		4. Создание причинно-следственной модели
		5. Создание технологии нижестоящих элементов
		6. Формирование/модернизация ресурсов
		7. Календарное и ресурсное планирование
		8. Оптимизация
		9. Назначение субъектов и определение ответственности
		10. Назначение ресурсов
Реализация	IV. Выполнение действий и получение результата	11. Выполнение действий и получение результата
Рефлексия	V. Оценивание результата и рефлексия	12. Оценивание результата и рефлексия

Рис. 2. Классификационная схема процессов

Обеспечивающими бизнес-процессами являются процессы, предназначенные для обеспечения основных и сопутствующих процессов и ориенти-

рованные на поддержку их универсальных черт. Так, для любого предприятия подобными процессами являются процесс финансового обеспечения деятельности, процесс обеспечения кадрами, процесс юридического обеспечения и т. п.

Аналогично *информационные процессы* (процессы, связанные с изменением информации или действиями с использованием информации) могут быть разделены на процессы сбора, поиска, обработки, представления, хранения, передачи и защиты информации.

Сбор информации – процесс, включающий в себя нахождение, сбор и извлечение первичных данных, предваряющие решение любой задачи.

Поиск информации – процесс извлечения хранящихся данных, необходимых для достижения конкретной цели.

Обработка информации – целенаправленный процесс изменения содержания или формы подачи информации. Можно выделить два типа обработки информации: обработка, связанная с получением нового содержания, новой информации (преобразование по правилам (в том числе вычисления по формулам), исследование объектов познания по их моделям, логические рассуждения, обобщение и др.); обработка, связанная с изменением формы представления информации, не изменяющая ее содержания (структурирование, кодирование и др.).

Представление информации – процесс приведения информационного материала в более удобный вид в зависимости от ситуации и адресата.

Хранение информации – фиксация ее на некотором носителе тем или иным способом.

Передача информации – процесс, осуществляемый по схеме: источник информации – кодирующее устройство – канал связи – декодирующее устройство – приемник информации.

Защита информации – процесс ограничения доступа посторонних субъектов к информации для обеспечения безопасности.

4. Характеристические свойства процессов, которыми может обладать любой из процессов, представленных в п. 1–3, в частности:

- автоматизированность (ручной, автоматизированный, автоматический),
- параллелизм (синхронный, асинхронный, конвейерный, последовательный),
- состав исполнителей (индивидуальный, групповой, многоуровневый),
- характер (дискретный, непрерывный) и др.

Характеристические свойства процессов не являются основанием для классификации, они лишь

вводят дополнительные требования к языкам описания процессов. Например, для описания параллельных процессов необходимо наличие в языке средств синхронизации ветвей и т. п.

Отметим, что процессы ЖЦ систем (в рамках системной и программной инженерии как наиболее развитых в части ЖЦ дисциплин) вписываются в данную схему. В частности, приведенные в книге [8] технические процессы являются детализацией всех процессов элементарной деятельности в рамках конкретной практической деятельности (например, разработки программных систем). Процессы соглашения, процессы организационного обеспечения проектов и процессы проекта (за исключением процесса управления информацией) являются организационно-управляющими процессами, а процесс управления информацией – совокупностью информационных процессов.

Включенные в вышеприведенную классификацию процессы по сути являются элементарными. В реальной жизни доминируют *комплексные процессы* (интегрирующие):

- один из процессов элементарной деятельности конкретного вида, представленный элементом матрицы на рис. 2 для любого уровня ее детализации;
- один из организационно-управляющих процессов, тип которого соответствует выбранному процессу элементарной деятельности конкретного вида;
- множество информационных процессов, связывающих функции выбранного процесса элементарной деятельности конкретного вида;
- набор вычислительных процессов, детализирующих некоторые функции выбранного процесса элементарной деятельности конкретного вида.

Отметим, что вычислительный процесс является частным случаем управляющего процесса нижнего уровня, при этом он имеет наивысший уровень формализации и, как правило, моделируется системой конечных автоматов или диаграммой переходов состояний специального вида, что позволяет построить комплекс методов его исследования, обеспечивающий прогнозирование его поведения и управление им [24].

3. ПРИМЕР КОМПЛЕКСНОГО ПРОЦЕССА

Пример комплексного процесса приведен на рис. 3 и рис. 4, где представлен процесс ремонта автотранспортного средства. Изображенный сплошными линиями на рис. 3 фрагмент в соответствии с процессами ЖЦ является процессом под-

держки, а в соответствии с видами деятельности относится к процессам практической деятельности. Дальнейшая функциональная детализация позволяет соотнести его с разделом 45 ОКВЭД «Торговля оптовая и розничная автотранспортными средствами и мотоциклами и их ремонт» и, далее, с подразделом 45.2 («Техническое обслуживание и ремонт автотранспортных средств»), включающем (в соответствии с подпунктами 45.20.0 – 45.20.5) следующие виды работ:

- ремонт транспортных средств, в том числе: механический ремонт, ремонт электрических систем, ремонт системы впрыскивания, текущее техническое обслуживание транспортных средств, ремонт кузова, ремонт ходовой части, мойку и полировку, покраску и рисование, ремонт лобового стекла и окон, ремонт автомобильных кресел;
- шиномонтаж и все виды связанных с ним работ;
- антикоррозийную обработку;
- установку дополнительного оборудования (сигнализации, радиоаппаратуры, дополнительных фар и т. п.), запасных частей и принадлежностей,

не относящихся непосредственно к производственному процессу;

- предпродажную подготовку;
- техническую помощь на дорогах;
- транспортирование неисправных автотранспортных средств к месту их ремонта или стоянки.

Изображенный штриховыми линиями на рис. 3 фрагмент, а также детализирующий его процесс, представленный на рис. 4 с помощью диаграммы переходов состояний, является организационно-управляющим процессом, при этом его тип может варьироваться в зависимости от вида предприятия, где он реализован. Например, для сервисного центра он является основным, а для занимающейся перевозками автобазы – вспомогательным.

Фрагменты, изображающие функции, связанные с накопителями данных, представляют собой информационные процессы, типы которых зависят от способа изменения или использования информации. Наконец, фрагмент «Расчет стоимости ремонта» может являться вычислительным процессом в случае его представления в виде формулы, вычисляющей требуемый результат.

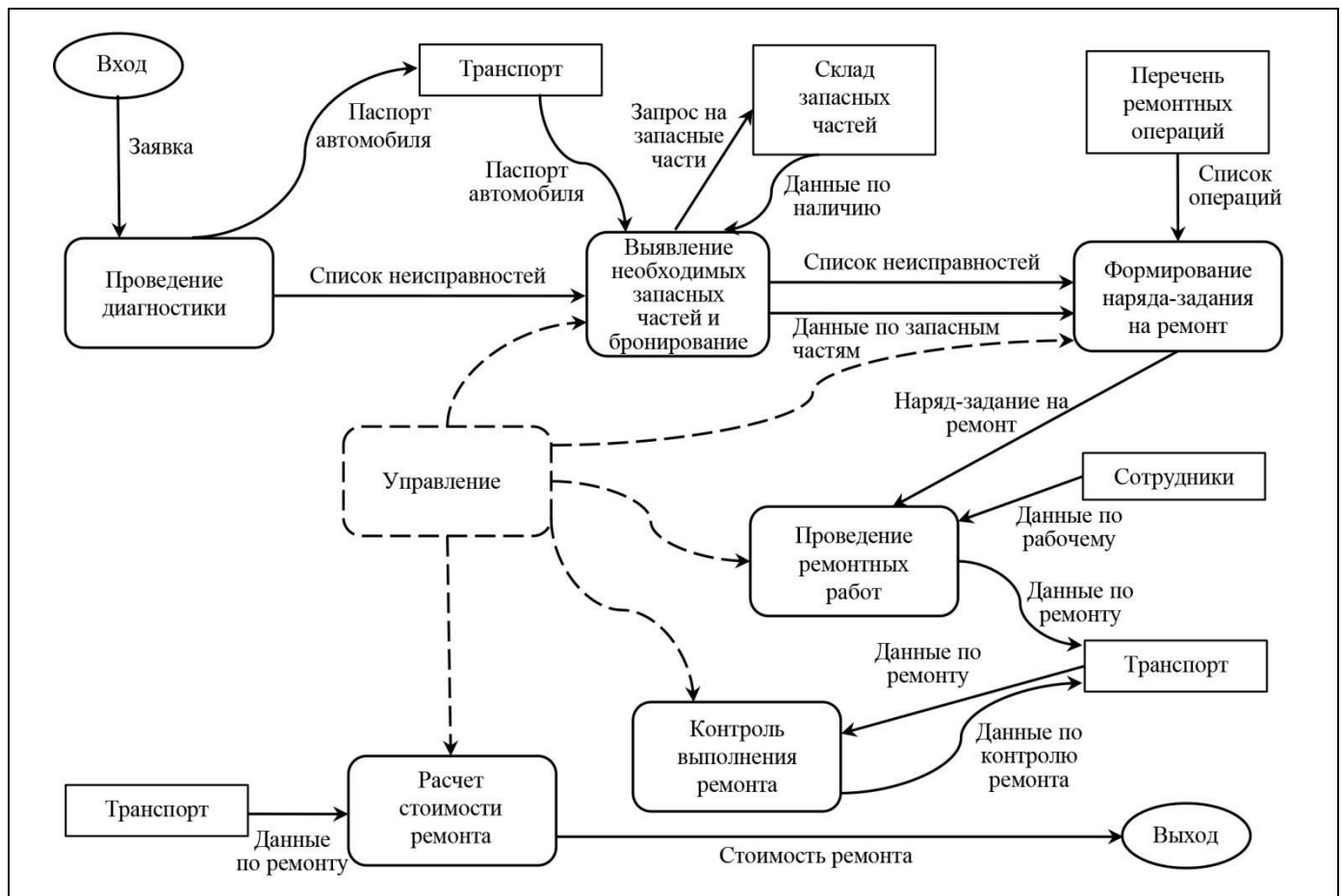


Рис. 3. Функциональная схема процесса ремонта автомобиля



Рис 4. Схема поведения процесса ремонта автомобиля

ЗАКЛЮЧЕНИЕ

Предложенная классификация процессов обеспечивает базис для решения (в рамках теории процессов) перечисленных в § 2 задач, обеспечивающий систематизированное исследование процессов деятельности. Дальнейшими первоочередными задачами в рамках исследуемого направления являются:

- разработка унифицированной модели процесса, интегрирующей его функциональный, информационный и поведенческий аспекты. При этом в качестве основы такой модели могут быть использованы соответствующие модели, применяемые в алгоритмических теориях процессов (прежде всего в теории вычислительных процессов, теории бизнес-процессов);
- разработка комплекса методов исследования процесса на основании стадий и этапов его жизненного цикла.

ЛИТЕРАТУРА

1. Белов М.В., Новиков Д.А. Методология комплексной деятельности. – М.: Ленанд, 2018. – 320 с. [Belov, M.V., Novikov, D.A. Methodology of Complex Activity. Foundations of Understanding and Modelling. – Cham: Springer, 2020. – 229 p.]

2. Новиков А.М., Новиков Д.А. Методология. – М.: СИНТЕГ, 2007. – 668 с. [Novikov, A.M., Novikov, D.A. Metodologiya. – M.: SINTEG, 2007. – 668 s. (In Russian)]
3. Новиков А.М., Новиков Д.А. Методология: словарь системы основных понятий. – М.: Либроком, 2015. – 208 с. [Novikov, A.M., Novikov, D.A. Metodologiya: slovar' sistemy osnovnykh ponyatii. – M.: Librokom, 2015. – 208 s. (In Russian)]
4. Новиков Д.А. Методология управления. – М.: Либроком, 2016. – 128 с. [Novikov, D.A. Control Methodology. – New York: Nova Science, 2013.]
5. Новиков А.М., Новиков Д.А. Методология научного исследования. – М.: Ленанд, 2017. – 272 с. [Novikov, A.M. and Novikov, D.A. Research Methodology: From Philosophy of Science to Research Design. – London: CRC Press, 2013.]
6. Калянов Г.Н. Моделирование, анализ, реорганизация и автоматизация бизнес-процессов. – М.: Финансы и статистика, 2006. – 240 с. [Kalyanov, G.N. Modelirovanie, analiz, reorganizatsiya i avtomatizatsiya biznes-protsessov. – M.: Finansy i statistika, 2006. – 240 s. (In Russian)]
7. Философский энциклопедический словарь. – М.: Советская энциклопедия, 1983. – 840 с. [Filosofskii entsiklopedicheskii slovar'. – M.: Sovetskaya entsiklopediya, 1983. – 840 s. (In Russian)]
8. Батоврин В.К. Системная и программная инженерия. Словарь-справочник. – М.: ДМК Пресс, 2010. – 280 с. [Batorin, V.K. Sistemnaya i programmnaya inzheneriya. Slovar'-spravochnik. – M.: DMK Press, 2010. – 280 s. (In Russian)]
9. ISO/IEC/IEEE 15288: 2023 Systems and Software Engineering – System Life Cycle Processes. – Geneva: International Organization for Standardization, 2023.



10. *ISO/IEC 12207: 2017 Systems and Software Engineering – Software Life Cycle Processes*. – Geneva: International Organization for Standardization, 2017.
 11. *Калянов Г.Н.* Теория бизнес-процессов. – М.: Горячая линия – Телеком, 2023. – 296 с. [*Kalyanov, G.N.* Teoriya biznes-protsessov. – М.: Goryachaya liniya – Telekom, 2023. – 296 s. (In Russian)]
 12. *Калянов Г.Н.* Теория и практика реорганизации бизнес-процессов. – М.: СИНТЕГ, 2000. – 212 с. [*Kalyanov, G.N.* Teoriya i praktika reorganizatsii biznes-protsessov. – М.: SINTEG, 2000. – 212 s. (In Russian)]
 13. *Магомедова Н.Г.* Классификация бизнес-процессов на агропромышленных предприятиях // Вопросы структуризации экономики. – 2012. – № 1. – С. 41–45. [*Magomedova, N.G.* Klassifikatsiya biznes-protsessov na agropromyshlennykh predpriyatiyakh // Voprosy strukturizatsii ekonomiki. – 2012. – No. 1. – S. 41–45. (In Russian)]
 14. *Амирханов К.Г.* Бизнес-процессы промышленного предприятия: понятие и классификация // Вопросы структуризации экономики. – 2005. – № 2. – С. 140–149. [*Amirkhanov, K.G.* Biznes-protsessy promyshlennogo predpriyatiya: ponyatie i klassifikatsiya // Voprosy strukturizatsii ekonomiki. – 2005. – No. 2. – S. 140–149. (In Russian)]
 15. *Остроухова Н.Г.* Бизнес-процессы предприятий ТЭК: понятие, содержание, классификация // Сибирская финансовая школа. – 2012. – № 1 (90). – С. 118–122. [*Ostroukhova, N.G.* Biznes-protsessy predpriyatii TEK: ponyatie, sodержание, klassifikatsiya // Sibirskaya finansovaya shkola. – 2012. – No. 1 (90). – S. 118–122. (In Russian)]
 16. *Патрушева А.М.* Классификация процессов: некоторые аспекты применения // Вестник МГСУ. – 2016. – № 6. – С. 142–151. [*Patruseva, A.M.* Klassifikatsiya protsessov: nekotorye aspekty primeneniya // Vestnik MGSU. – 2016. – No. 6. – S. 142–151. (In Russian)]
 17. *Фролов Т.Ю.* Классификация инновационных процессов на нефтегазовых предприятиях // Известия Санкт-Петербургского экономического государственного университета. – 2016. – № 2 (98). – С. 131–134. [*Frolov, T.Yu.* Klassifikatsiya innovatsionnykh protsessov na neftegazovykh predpriyatiyakh // Izvestiya Sankt-Peterburgskogo ekonomicheskogo gosudarstvennogo universiteta. – 2016. – No. 2 (98). – S. 131–134. (In Russian)]
 18. *Fleaca, E., Fleaca, B.* The Business Process Management Map – An Effective Means for Managing the Enterprise Value Chain // *Procedia Technology*. – 2016. – Vol. 22. – P. 954–968.
 19. *Hassen, M., Gargouri, F.* Multi-dimensional Classification of Sensitive Business Process Modeling Aspects // *Procedia Computer Science*. – 2024. – Vol. 239. – P. 2158–2167. – DOI: <https://doi.org/10.1016/j.procs.2024.06.404>
 20. *Latiffianti, E., Siswanto, N., Wiratno, S., Saputra, Y.* CIMOSA Process Classification for Business Process Mapping in Non-Manufacturing Firms: A Case Study // *Proceedings of 3rd International Materials, Industrial and Manufacturing Engineering Conference (MIMEC2017)*. – Miri, Malaysia, 2017. – AIP Conference Proceedings. – Vol. 1902. – Art. no. 020040. – DOI: <https://doi.org/10.1063/1.5010657>
 21. *Mancioppi, M., Danylevych, O., Karastoyanova, D., Leymann, F.* Towards Classification Criteria for Process Fragmentation Techniques // *Lecture Notes in Business Information Processing*. – 2012. – Vol. 99. – P. 1–12. – DOI: https://doi.org/10.1007/978-3-642-28108-2_1
 22. *Santos, R., Salgado, M., Pereira, V.* Business Process Prioritization Criteria: A Case Study in the Financial Market // *RAUSP Management Journal*. – 2022. – Vol. 57. – No. 1. – P. 35–48.
 23. *OK 029 – 2014* (КДЕС Ред. 2): Общероссийский классификатор видов экономической деятельности (утв. Приказом Росстандарта от 31.01.2014 N 14-ст). [*OK 029 – 2014* (KDES Red. 2): Obshcherossiiskii klassifikator vidov ekonomicheskoi deyatel'nosti (utv. Prikazom Rosstandarta ot 31.01.2014 N 14-st). (In Russian)]
 24. *Миронов А.М.* Теория процессов. – Переславль-Залесский: Университет г. Переславля, 2008. – 345 с. [*Mironov, A.M.* Teoriya protsessov. – Pereslavl'-Zalesskii: Universitet g. Pereslavl'ya, 2008. – 345 s. (In Russian)]
- Статья представлена к публикации руководителем РРС О.В. Логиновским.*
- Поступила в редакцию 14.04.2025,
после доработки 21.05.2025.
Принята к публикации 21.05.2025.*
- Калянов Георгий Николаевич** – д-р техн. наук,
✉ kalyanov@ipu.ru,
ORCID iD: <https://orcid.org/0000-0003-2429-0703>
- Новиков Дмитрий Александрович** – д-р техн. наук, академик РАН,
✉ novikov@ipu.ru,
ORCID iD: <https://orcid.org/0000-0002-9314-3304>
- Институт проблем управления им. В. А. Трапезникова РАН,
г. Москва.
- © 2025 г. Калянов Г. Н., Новиков Д. А.



Эта статья доступна по лицензии Creative Commons «Attribution» («Атрибуция») 4.0 Всемирная.

CLASSIFICATION OF ACTIVITY PROCESSES

G. N. Kalyanov* and D. A. Novikov**

***Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia,

*✉ kalyanov@ipu.ru, **✉ novikov@ipu.ru

Abstract. Nowadays, methodology—the general theory of activity organization and control—continues to develop. This paper introduces a classification system for processes, determining the place of any process and establishing the presence of interrelations with other processes. Several classification attributes are identified, and a corresponding classification scheme of elementary processes is proposed. The concept of a complex process is considered, and its division into elementary processes in accordance with the classification scheme is demonstrated. An example of a complex process of motor vehicle repair is given; its functional and behavioral schemes are presented, and the corresponding elementary processes of various types are described in detail, namely, supporting processes (the processes of its life cycle), organizational and management processes (business processes), as well as informational and computational processes.

Keywords: process, activity, methodology, classification, life cycle.

УПРАВЛЕНИЕ КОМПЛЕКСОМ РАБОТ С НЕОПРЕДЕЛЕННЫМИ МОМЕНТАМИ ПОСТУПЛЕНИЯ ЗАПРОСОВ НА ВЫПОЛНЕНИЕ

Д. А. Кононов*, М. Г. Фуругян**

*Институт проблем управления им. В. А. Трапезникова РАН, г. Москва

**Федеральный исследовательский центр «Информатика и управление» РАН, г. Москва

*✉ dmitrykon52@gmail.com, **✉ rtscas@yandex.ru

Аннотация. Рассматривается задача составления графика выполнения комплекса основных работ при условии, что в некоторые неопределенные моменты времени поступают заявки на выполнение дополнительных, более приоритетных работ. Если при выполнении основного задания поступает заявка на дополнительную работу, то выполнение этого задания завершается и должно быть начато заново в некоторый момент времени после окончания выполнения дополнительной работы. Все работы (основные и дополнительные) выполняются без прерываний. Предполагается, что при выполнении основного задания поступление более чем одной заявки на выполнение дополнительной работы маловероятно; такие случаи не рассматриваются. Кроме того, заявка на выполнение дополнительной работы может прийти только после завершения обслуживания предыдущей дополнительной заявки. Исследуются две формальные постановки задачи. В первой показателем эффективности является момент завершения комплекса основных работ, и задача заключается в его минимизации. Во второй постановке минимизируется вероятность коллизии, т.е. ситуации, когда заявка на выполнение дополнительной работы поступает в момент выполнения основного задания. Решение указанных задач основано на сведении их к бесконечным антагонистическим играм и дискретной аппроксимации последних конечными играми. Рассмотрены модельные примеры. Исследована также постановка с нефиксированными длительностями выполнения основных работ, линейно зависящими от объема выделенных им дополнительных ресурсов. В этом случае поиск допустимого расписания сводится к решению задачи линейного программирования.

Ключевые слова: системы реального времени, планирование работ, коллизия, антагонистическая игра, невозобновляемые ресурсы, смешанная стратегия, оптимальная стратегия.

ВВЕДЕНИЕ

При разработке, эксплуатации и управлении сложными объектами широко применяются специализированные автоматизированные системы управления (АСУ) различного класса и назначения. Такие системы ориентированы на выполнение ряда функций управления в заданных предметных областях и имеют ряд существенных свойств (системных параметров) и характеристик, отражающих эффективность их реализации [1].

В современных условиях важное значение имеют задачи разработки управляющих воздействий в автоматизированных системах реального времени. Их основная характеристика заключается в том, что существенно ограничен сверху интервал времени, за который необходимо произвести обработку входной информации и выдать результат вычислений в виде управляющих воздействий на объект или в виде сообщений пользователю. Еще более сложной становится задача в условиях неопределенности, когда следует найти расписание в меняющейся обстановке – при поступлении новых,

заранее не предусмотренных заявок на выработку управляющих сигналов.

Рассматриваемый класс задач имеет, помимо чисто научной, большую практическую ценность. Потребность в быстрых алгоритмах, составляющих многопроцессорные расписания, часто возникает в задачах распределенных вычислений в реальном времени и в задачах оперативного управления на основе обработки и анализа поступающих в реальном времени данных. В качестве иллюстрации широкой практической распространенности многопроцессорных АСУ реального времени и важности эффективных алгоритмов для выполнения расчетов и управления можно привести следующие примеры.

- Современные системы космического мониторинга представляют собой системы реального времени, непрерывно обрабатывающие поступающие данные о движении объектов в околоземном космическом пространстве.

- Системы управления ядерными реакторами на АЭС получают в режиме реального времени данные с множества датчиков и должны на их основе оперативно осуществлять управляющие воздействия на реактор.

- В аналитических центрах при правительствах развитых стран системы реального времени используются для мониторинга и анализа непрерывно поступающей из различных точек экономической или экологической информации. Такие системы должны эффективно обрабатывать огромные массивы данных и, исходя из них, оперативно оповещать о каких-то замеченных проблемах на ранних стадиях их возникновения.

- При испытаниях летательных аппаратов и других сложных технических объектов большое значение приобретают получение и оперативная обработка вычислительной системой реального времени периодически поступающей информации о состоянии различных узлов объекта.

- В современном аэропорту с большим количеством взлетно-посадочных полос постоянно принимаются решения о назначении тех или иных самолетов на различные полосы и порядке, в котором этим самолетам следует взлетать и садиться.

- В сложных логистических системах решения должны приниматься в режиме реального времени при возникновении критических ситуаций.

- В чрезвычайных ситуациях необходимо экстренно обрабатывать поступающую информацию и в реальном времени рассчитывать силы и средства (составлять графики), направляемые на их ликвидацию.

Необходимо отметить, что корректность АСУ реального времени зависит не только от правильности результатов ее вычислений, но и от времени, за которое эти результаты были получены. Составление расписаний реального времени – важная часть подобных систем, так как проектировщик системы должен быть уверен в том, что все задания будут исполнены в срок.

Наряду с задачей построения допустимого расписания для известной вычислительной системы реального времени актуальной является также и обратная задача построения (синтеза) системы реального времени некоторой минимально возможной конфигурации, в которой всегда может быть найдено допустимое расписание при заданном наборе работ. Особенно актуальной эта проблема является для бортовых вычислительных комплексов, при проектировании которых обычно стремятся минимизировать необходимые вычислительные ресурсы с целью экономии их массы и потребляемой электроэнергии. В работе [2] описан алгоритм синтеза таких систем.

Задачи планирования работ возникают во многих сферах деятельности человека, таких как строительство, экономика, военное дело, экология, разработка полезных ископаемых, управление сложными техническими объектами (самолеты, электростанции, ядерные реакторы), составление расписаний движения транспорта, управление вычислительными процессами, в частности в системах реального времени, в других отраслях. Данной тематике посвящено большое число научных работ. Отметим фундаментальные монографии [3, 4]. В них авторы исследуют различные задачи по теории расписаний и дискретной оптимизации, проводят их классификацию (полиномиально разрешимые и NP-трудные задачи) и приводят алгоритмы их решения. Кроме того, проводится анализ вычислительной сложности предлагаемых алгоритмов. В книге [5] на основе понятия расстояния между задачами разработаны методы решения ряда NP-трудных задач для критерия минимизации максимального запаздывания, а также для задач на быстроедействие.

При решении задач составления расписаний используется разнообразный математический аппарат. Так, в статьях [6, 7] описана методика управления вычислительными процессами с директивными сроками. Методика основана на использовании конечных автоматов с остановкой таймера и построения временных диаграмм. Данный подход особенно актуален при проектировании систем реального времени.



Большой интерес представляют задачи планирования работ в условиях неопределенности и риска. Так, в работах [8–10] исследованы задачи с нефиксированными параметрами. Предполагается, что длительности выполнения работ, а также имеющиеся и требуемые ресурсы задаются вероятностными характеристиками либо интервалами, в пределах которых они могут принимать значения. В последнем случае разработан алгоритм разбиения множества всех возможных значений указанных параметров на так называемые многогранники устойчивости. Каждый такой многогранник характеризуется тем, что для всех принадлежащих ему значений параметров структура оптимального расписания остается неизменной. Это позволяет заранее построить расписание для каждого многогранника, а при проведении вычислений в реальном масштабе времени, когда будут известны значения неопределенных параметров, выбрать нужное решение. Данный подход особенно актуален при проектировании и функционировании систем реального времени, когда на проведение вычислений отводится жестко ограниченное время.

В публикации [11] предлагается методика планирования производства, в которой расписание выполнения работ составляется совместно с анализом возможных изменений производственных мощностей. При этом исходная задача сведена к нелинейной целочисленной задаче математического программирования. В статье [12] исследована задача планирования сроков завершения работ с учетом стохастической изменчивости объемов ресурсов, необходимых для их выполнения.

В настоящей статье продолжены исследования, начатые в работах [13, 14]. Рассматривается задача планирования комплекса работ (основных заданий) в условиях неопределенности. Предполагается, что в некоторые неопределенные моменты времени поступают заявки на выполнение дополнительных (более приоритетных) работ. С применением игрового подхода исследуются две постановки задачи. В первой минимизируется момент завершения выполнения комплекса основных работ. Во второй постановке показателем эффективности является вероятность невозникновения коллизии (т. е. ситуации, при которой заявка на выполнение дополнительной работы поступает в момент выполнения основного задания). Исследована также постановка с нефиксированными длительностями выполнения основных работ, линейно зависящими от объема выделенных для этого дополнительных ресурсов. В этом случае поиск допустимого расписания сводится к решению задачи линейного программирования.

1. ПОСТАНОВКА ЗАДАЧИ

Имеется комплекс основных работ (заданий) $W = \{w_1, w_2, \dots, w_n\}$, для которых известны длительности $t_1, t_2, \dots, t_n$ и последовательность выполнения $w_1 \rightarrow w_2 \rightarrow \dots \rightarrow w_n$.

В некоторые неопределенные моменты времени $0 \leq y_1 \leq y_2 \leq \dots \leq y_m \leq T$ могут поступать заявки на выполнение дополнительных работ $Z = \{z_1, z_2, \dots, z_m\}$, длительности выполнения которых составляют $\tau_1, \tau_2, \dots, \tau_m$ соответственно. Величина T является заданной:

$$T \geq \sum_{i=1}^n t_i.$$

Дополнительные работы являются более приоритетными, чем основные. Если при выполнении основного задания поступает заявка на дополнительную работу, то выполнение этого задания завершается и должно быть заново начато в некоторый момент времени после окончания дополнительной работы. Такая ситуация называется коллизией.

Все работы (основные и дополнительные) выполняются без прерываний. Предполагается, что при выполнении основного задания поступление более чем одной заявки на выполнение дополнительной работы маловероятно и такие случаи не рассматриваются. Кроме того, заявка на выполнение дополнительной работы может прийти только после завершения обслуживания предыдущей дополнительной заявки. В дальнейшем предполагается, что основные и дополнительные работы представляют из себя программные модули, решающие некоторые прикладные задачи, для выполнения которых имеется вычислительное устройство. Оно является возобновляемым ресурсом, т. е. может использоваться многократно. При поступлении заявки на дополнительную работу это устройство немедленно передается для ее выполнения. Предполагается также, что в случае невозникновения коллизий основные работы должны завершиться не позднее срока T . Рассматриваются две постановки задачи.

Задача 1. Требуется выработать такую стратегию выполнения комплекса основных работ, при которой минимизируется момент завершения последнего задания w_n (полного комплекса работ с учетом дополнительных).

Задача 2. Предполагается, что требования на выполнение комплекса основных работ W и заявки на выполнение дополнительных заданий Z посту-

пают многократно. Требуется выработать такую стратегию выполнения основных работ, при которой максимизируется вероятность невозникновения коллизии.

Отметим, что в обеих постановках о поступлении заявки на выполнение дополнительной работы z_j становится известно только в момент времени y_j , $j = \overline{1, m}$. Подобные задачи возникают, например, при проведении летных испытаний. В штатном режиме выполняются вычисления с помощью прикладных модулей w_i , $i = \overline{1, n}$. В неопределенные моменты времени y_j , $j = \overline{1, m}$, может возникнуть нештатная ситуация, заключающаяся, например, в том, что значения некоторых важных параметров вышли за пределы допустимых значений. В этом случае запланированные заранее вычисления прерываются и выполняются дополнительные, более приоритетные работы.

Рассматривается также постановка, в которой длительности основных работ не являются фиксированными и линейно зависят от выделенных им дополнительных (невозобновляемых) ресурсов.

2. РЕШЕНИЕ ЗАДАЧИ 1

Рассмотрим сначала случай $m = 1$. Пусть заявка на выполнение дополнительной работы z приходит в момент времени y , а длительность ее выполнения составляет τ . Определим антагонистическую игру с платежной функцией $F(x_1, x_2, \dots, x_n, y)$. В этой игре стратегия первого игрока определяет моменты времени x_i начала выполнения основных работ $w_i \in W$, $i = \overline{1, n}$, а стратегия второго игрока – момент времени y прихода заявки на выполнение дополнительного задания z . Платежная функция определяется следующим образом:

$$F(x_1, x_2, \dots, x_n, y) = \begin{cases} x_n + t_n, & \text{если } 0 \leq y \leq x_1 - \tau, \\ \text{или если } x_k + t_k \leq y \leq x_{k+1} - \tau \\ \text{при некотором } 1 \leq k \leq n-1, \\ \text{или если } x_n + t_n \leq y \leq T; \\ y + \tau + \sum_{i=k}^n t_i, \\ \text{если } x_k \leq y < x_k + t_k \text{ при некотором } 1 \leq k \leq n \\ \text{или если } x_k - \tau < y \leq x_k \text{ некотором } 1 \leq k \leq n-1. \end{cases}$$

Иными словами, при отсутствии коллизии $F(x_1, x_2, \dots, x_n, y) = x_n + t_n$. В случае, когда коллизия возникает при выполнении некоторой работы $w_k \in W$,

$$F(x_1, x_2, \dots, x_n, y) = y + \tau + \sum_{i=k}^n t_i.$$

Это означает, что оптимальная гарантирующая стратегия первого игрока в планировании выполнения работ W следующая: $x_1^0 = 0$, $x_i^0 = x_{i-1}^0 + t_{i-1}$, $i = \overline{2, n}$. В этом случае при отсутствии коллизии комплекс работ W завершится в момент времени $x_n^0 + t_n$, а при возникновении коллизии – в момент времени $x_n^0 + t_n + \tau$. При любой другой стратегии $x_1, x_2, \dots, x_n$ в худшем случае комплекс работ W завершится в момент времени $x_n + t_n + \tau > x_n^0 + t_n + \tau$, поскольку $x_n > x_n^0$. Таким образом, стратегия $(x_1^0, x_2^0, \dots, x_n^0)$ является оптимальной гарантирующей стратегией при $m = 1$.

Проиллюстрируем это на следующем примере.

Пример 1. Пусть $n = m = 1$. Тогда

$$F(x, y) = \begin{cases} x + t, \\ \text{если } 0 \leq y \leq x - \tau \\ \text{или если } x + t \leq y \leq T, \\ \text{т.е. при отсутствии коллизии;} \\ y + \tau + t, \\ \text{если } x - \tau < y < x + t, \\ \text{т.е. при возникновении коллизии.} \end{cases}$$

Данная платежная функция схематично изображена на рис. 1.

В этом случае $F(0, y) < 2t + \tau$ при всех $y \in [0; T]$; $F(x, y) > 2t + \tau$ при $x \neq 0$ и любом $y \in (x - \tau; x + t)$. Следовательно, $x_1^0 = 0$ – оптимальная гарантирующая стратегия первого игрока. ♦

С помощью аналогичных рассуждений рассматривается случай $n > 1, m > 1$. В качестве плановой стратегии выбираем $(x_1^0, x_2^0, \dots, x_n^0)$. Далее, если в интервале $[x_1^0; x_1^0 + t_1]$ коллизии не возникает, то x_1^0 оставляем без изменения и полагаем $x_2^0 = x_1^0 + t_1$. Если же в интервале $[x_1^0; x_1^0 + t_1]$ возникает коллизия в момент времени y_1 , то полагаем $x_1^0 = y_1 + t_1$. Далее по аналогии в первом случае на возникновение коллизии исследуется интервал



$[x_2^0; x_2^0 + t_2]$, а во втором – интервал $[x_1^0; x_1^0 + t_1]$ и т. д. Таким образом, оптимальная стратегия планирования работ W строится динамически, в зависимости от поступления заявок на выполнение дополнительных заданий.

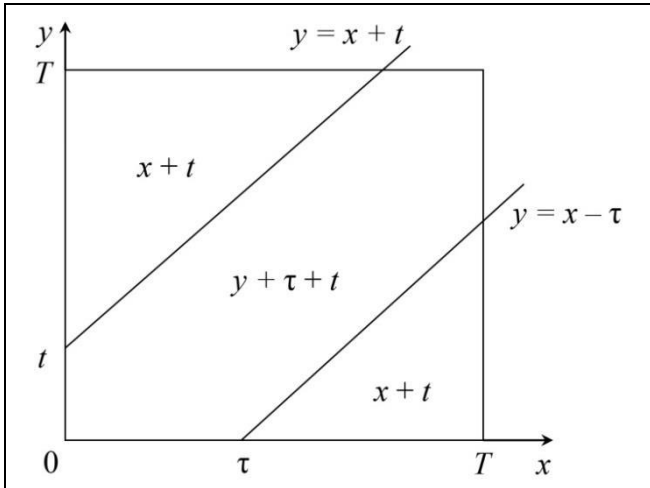


Рис. 1. Платежная функция в задаче 1 при $n = m = 1$

3. РЕШЕНИЕ ЗАДАЧИ 2

Определим платежную функцию антагонистической игры следующим образом:

$$F(x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_m) = \begin{cases} 1, & \text{если при всех } i = \overline{1, n}, j = \overline{1, m} \\ & x_i + t_i \leq y_j \leq x_{i+1} - \tau_j, \\ & \text{или } 0 \leq y_j \leq x_1 - \tau_j, \\ & \text{или } x_n + t_n \leq y_j \leq T; \\ 0, & \text{если при некоторых } 1 \leq i \leq n, 1 \leq j \leq m \\ & x_i \leq y_j \leq x_i + t_i \\ & \text{или } x_i - \tau_j \leq y_j \leq x_i. \end{cases}$$

Пусть $\bar{x} = (x_1, x_2, \dots, x_n)$, $\bar{y} = (y_1, y_2, \dots, y_m)$. Тогда $F(\bar{x}, \bar{y}) = 1$, если коллизия отсутствует, и $F(\bar{x}, \bar{y}) = 0$ при возникновении коллизии. Пусть $X = \{\bar{x} : x_i + t_i \leq x_{i+1}, i = \overline{1, n-1}, x_1 \geq 0, x_n \leq T\}$, $Y = \{\bar{y} : y_j + \tau_j \leq y_{j+1}, j = \overline{1, m-1}, y_1 \geq 0, y_m \leq T\}$, $f(\bar{x})$ – смешанная стратегия первого игрока, т. е. вероятностная мера на множестве X . Тогда из определения платежной функции $F(\bar{x}, \bar{y})$ следует, что

$$E(f, \bar{y}) = \int_X F(\bar{x}, \bar{y}) df(\bar{x})$$

– вероятность отсутствия коллизии при фиксированном $\bar{y}$. Оптимальная смешанная стратегия первого игрока $f^0(\bar{x})$ максимизирует величину

$$\min_{\bar{y} \in Y} E(f, \bar{y}) : \max_f \min_{\bar{y} \in Y} E(f, \bar{y}) = \min_{\bar{y} \in Y} E(f^0, \bar{y}),$$

где максимум берется по всем вероятностным мерам на множестве X , т. е. $f^0(\bar{x})$ является наилучшей гарантирующей стратегией первого игрока.

Пример 2. Рассмотрим случай, когда $n = m = 1$, $X = [0; 1]$, $Y = [0; 1]$, $T = 1$. Вместо x_1, y_1, t_1, τ_1 будем писать x, y, t, τ соответственно. Пусть $t \leq 0,25$, $\tau \leq 0,25$. Тогда

$$F(x, y) = \begin{cases} 1, & \text{если } y \geq x + t \\ & \text{или если } y \leq x - \tau; \\ 0, & \text{если } x - \tau < y < x + t \end{cases},$$

$$x \in [0; 1], y \in [0; 1].$$

На рис. 2 дано схематическое изображение платежной функции F .

Пусть $f^0(x)$ – равномерно распределенная на интервале $[0; 1]$ вероятностная мера. Тогда из определения платежной функции $F(x, y)$ следует, что

$$\int_0^1 F(x, y) df^0(x) \geq 1 - \tau - t$$

при всех $y \in [0; 1]$. Если $f(x)$ – вероятностная мера на интервале на $[0; 1]$, не являющаяся равномерно распределенной, то существует отрезок $[x_0; x_0 + \tau + t] \subseteq [0; 1]$ такой, что

$$\int_{x_0}^{x_0 + \tau + t} df(x) > \tau + t.$$

Следовательно, существует $y \in [0; 1]$, при котором

$$\int_0^1 F(x, y) df(x) < 1 - \tau - t.$$

Поэтому

$$\int_0^1 F(x, y) df^0(x) = \max_f \min_{y \in [0; 1]} \int_0^1 F(x, y) df(x),$$

где максимум берется по всем вероятностным мерам на отрезке $[0; 1]$. Значит, $f^0(x)$ – оптимальная смешанная стратегия первого игрока. ♦

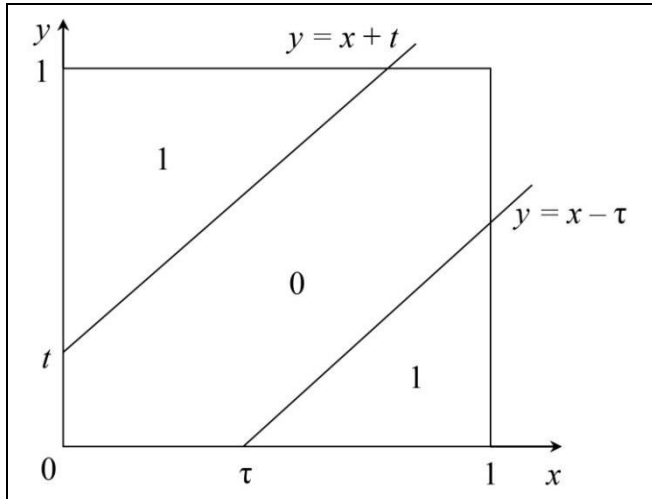


Рис. 2. Платежная функция в задаче 2 при $n = m = 1, t \leq 0,25, \tau \leq 0,25, T = 1$

Далее можно воспользоваться методом дискретной аппроксимации бесконечной игры конечной [15]. В этом случае для любого $\varepsilon > 0$ будет найдена ε -оптимальная смешанная стратегия первого игрока $f_\varepsilon(\bar{x})$, сосредоточенная в конечном числе точек. Пусть $f_\varepsilon(\bar{x})$ сосредоточена в точках $v_1, v_2, \dots, v_p$, в которых имеет скачки

$$q_1, q_2, \dots, q_p, q_j > 0, j = \overline{1, p}, \sum_{j=1}^p q_j = 1.$$

Каждой точке $v_j, j = \overline{1, p}$, соответствует некоторый план выполнения комплекса работ W , который следует применять с вероятностью q_j .

Пример 3. Пусть $n = m = 1, X = [0; 1], Y = [0; 1], T = 1, t = 0,25, \tau = 0,25$. Платежная функция имеет следующий вид:

$$F(x, y) = \begin{cases} 1, & \text{если } y \geq x + 0,25 \\ & \text{или если } y \leq x - 0,25; \\ 0, & \text{если } x - 0,25 < y < x + 0,25 \end{cases},$$

$$x \in [0; 1], y \in [0; 1].$$

Как было показано в примере 2, $f^0(x)$ – оптимальная смешанная стратегия первого игрока и

$$\int_0^1 F(x, y) df^0(x) \geq 1 - 0,25 - 0,25 = 0,5$$

при всех $y \in [0; 1]$. Пусть $\bar{f}(x)$ – вероятностная мера на отрезке $[0; 1]$, сосредоточенная в двух точках, $x = 0$ и $x = 1$, в которых имеет скачки, равные 0,5. Тогда

$$\int_0^1 F(x, y) d\bar{f}(x) \geq 0,5$$

при всех $y \in [0; 1]$. Следовательно, $\bar{f}(x)$ так же, как и $f^0(x)$, является оптимальной смешанной стратегией первого игрока. Таким образом, с вероятностью 0,5 работу w_1 следует начинать в момент времени 0 и с вероятностью 0,5 – в момент времени 1.

На рис. 3 дано схематическое изображение платежной функции F .

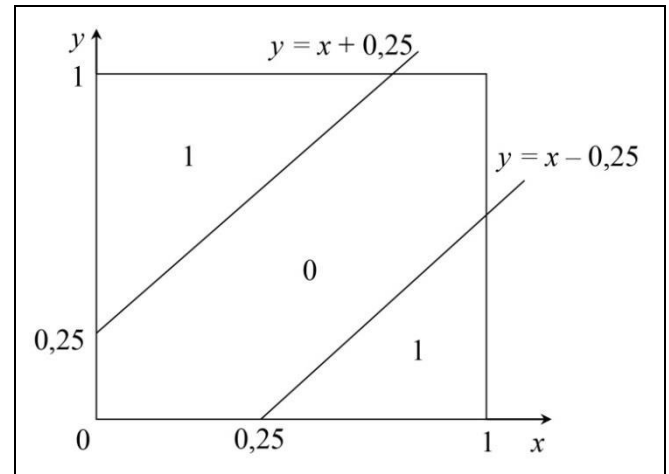


Рис. 3. Платежная функция в задаче 2 при $n = m = 1, t = 0,25, \tau = 0,25, T = 1$

4. ЗАДАЧА С НЕФИКСИРОВАННЫМИ ДЛИТЕЛЬНОСТЯМИ

В настоящем параграфе дополнительно предполагается, что для выполнения основных работ W имеется также L типов дополнительных невозобновляемых ресурсов (т. е. таких ресурсов, которые повторно использоваться не могут), объемы которых составляют $R_1, R_2, \dots, R_L$. Если работе $w_i \in W$ выделено r_{il} ресурсов l -го типа, $l = \overline{1, L}$, то длительность ее выполнения составляет

$$t_i = t_i^0 - \sum_{l=1}^L a_{il} r_{il}, i = \overline{1, n},$$

где a_{il} – заданные неотрицательные числа; t_i^0 – длительность ее выполнения в случае, когда ресурсы ей не выделяются. Предполагается, что выполняются следующие ограничения:

$$0 \leq r_{il} \leq r_{il}^0, i = \overline{1, n}, l = \overline{1, L}, \quad (1)$$

$$\sum_{i=1}^n r_{il} \leq R_l, l = \overline{1, L}, \quad (2)$$

$$t_i^0 - \sum_{l=1}^L a_{il} r_{il}^0 > 0, \quad (3)$$



где r_{il}^0 , $i = \overline{1, n}$, $l = \overline{1, L}$, – заданные положительные числа, означающие предельно допустимые объемы ресурсов, которые могут быть выделены работам. Неравенства (1) задают ограничения на величины каждого типа ресурсов, которые могут быть выделены каждой работе. Неравенства (2) – ограничения на суммарную величину каждого типа ресурса, выделяемую все работам вместе. Неравенства (3) – ограничения на длительности выполнения работ. Распределение ресурсов r_{il} , $i = \overline{1, n}$, $l = \overline{1, L}$, называется допустимым, если выполнены условия (1)–(3).

Цель – определить такое допустимое распределение ресурсов, которое способствовало бы решению задач 1 и 2. Из § 2 и 3 следует, что таким распределением ресурсов будет то, которое минимизирует суммарную длительность выполнения работ W . Таким образом, нужно решить следующую задачу линейного программирования: найти

$$\min_{r_{il}, i=\overline{1, n}, l=\overline{1, L}} \sum_{i=1}^n \left(t_i^0 - \sum_{l=1}^L a_{il} r_{il} \right)$$

при ограничениях (1), (2). Решение этой задачи даст оптимальное допустимое распределение ресурсов.

ЗАКЛЮЧЕНИЕ

Исследована задача планирования комплекса основных работ при условии, что в некоторые неопределенные моменты времени поступают заявки на выполнение дополнительных, более приоритетных работ. Последовательность выполнения основных работ фиксирована. Если при выполнении основного задания поступает заявка на дополнительную работу, то выполнение этого задания прекращается и должно быть заново начато в некоторый момент времени после окончания дополнительной работы. Все работы (основные и дополнительные) выполняются без прерываний. Рассмотрены две постановки задачи. В первой в качестве показателя эффективности выбран момент завершения комплекса основных работ, и задача заключается в его минимизации. Во второй постановке минимизируется вероятность коллизии, т. е. ситуации, когда заявка на выполнение дополнительной работы поступает в момент выполнения основного задания. Решение указанных задач основано на сведении их к бесконечным антагонистическим играм. Работа представленной методики продемонстрирована на модельных примерах. Исследована постановка с нефиксированными длительностями выполнения основных работ, линейно зави-

сящими от объема выделенных им дополнительных ресурсов. В этом случае поиск допустимого расписания сведен к решению задачи линейного программирования.

Полученные результаты могут быть использованы при планировании вычислений, выполняемых во время испытания и функционирования сложных технических объектов (самолеты, ядерные реакторы). В плановом режиме выполняются вычисления с помощью прикладных модулей, а в неопределенные моменты времени может наступить нештатная ситуация, заключающаяся, например, в том, что значения некоторых параметров вышли за пределы допустимой области. В этом случае плановые вычисления прерываются и выполняются дополнительные, более приоритетные работы.

Задачи построения расписаний в условиях неопределенности исследовались в работах [8–10]. В этих публикациях предполагалось, что все используемые ресурсы являются возобновляемыми, а нефиксированными являются некоторые параметры, такие как длительности выполнения работ или объемы имеющихся ресурсов. Эти параметры задаются либо интервально, либо вероятностными характеристиками. Алгоритмы решения рассмотренных задач были основаны на методе ветвей и границ. В отличие от этих работ, в настоящей статье рассмотрена задача с неопределенными моментами поступления запросов. Кроме того, рассмотрен также случай, когда помимо возобновляемых ресурсов имеются также невозобновляемые. Задачи с неоднородным комплексом ресурсов рассмотрены также в работах [13, 14] в детерминированном варианте.

В дальнейшем предполагается рассмотреть более общую постановку, в которой для выполнения основных и дополнительных работ имеется несколько вычислительных устройств.

ЛИТЕРАТУРА

1. Кононов Д.А. Исследование безопасности систем управления на основе анализа их системных параметров / Материалы 28-й Международной конференции «Проблемы управления безопасностью сложных систем» (ПУБСС'2020). – Москва, 2020. – С. 102–108. [Kononov, D.A. Research of Control Systems Security Based on the Analysis of Their System Parameters / Proceedings of the 28th International Conference «Problems of Complex Systems Security Management». – Moscow, 2020. – P. 102–108. (In Russian)]
2. Kononov, D., Furugyan, M. Control of a Complex of Works in Multiprocessor Real-time ACS / Proceedings of the 1st International Conference on Control Systems, Mathematical Modeling, Automation and Energy Efficiency (SUMMA2019). – Lipetsk, 2019. – URL: <https://ieeexplore.ieee.org/document/8947570>.

3. Танаев В.С., Гордон В.С., Шафранский Я.М. Теория расписаний. Одностадийные системы. – М.: Наука, 1984. – 383 с. [Tanaev, V.S., Gordon, V.S., Shafranskii, Ya.M. Theory of Scheduling: Single-Stage Systems. – Moscow: Nauka, 1984. – 383 s. (In Russian)]
 4. Brucker, P. Scheduling Algorithms. – Heidelberg: Springer, 2007. – 378 p.
 5. Лазарев А.А. Теория расписаний. Методы и алгоритмы. – М.: ИПУ РАН, 2019. – 408 с. [Lazarev, A.A. Scheduling Theory. Methods and Algorithms. – M.: ICS RAS, 2019. – 408 s. (In Russian)]
 6. Глонина А.Б., Балашов В.В. О корректности моделирования модульных вычислительных систем реального времени с помощью сетей временных автоматов // Моделирование и анализ информационных систем. – 2018. – Т. 25, № 2. – С. 174–192. [Glonina, A.B., Balashov, V.V. O korrektnosti modelirovaniya modul'nykh vychislitel'nykh sistem real'nogo vremeni s pomoshch'yu setei vremennykh avtomatov // Modelirovanie i analiz informatsionnykh sistem. – 2018. – Vol. 25, no. 2. – P. 174–192. (In Russian)]
 7. Глонина А.Б. Инструментальная система проверки выполнения ограничений реального времени для конфигураций модульных вычислительных систем // Вестн. МГУ. Сер. 15. Вычисл. математика и кибернетика. – 2020. – № 3. – С. 16–29. [Glonina, A.B. Instrumental'naya sistema proverki vypolneniya ogranichenii real'nogo vremeni dlya konfiguratsii modul'nykh vychislitel'nykh sistem // Vestn. MGU. Ser. 15. Vychisl. matematika i kibernetika. – 2020. – No. 3. – P. 16–29. (In Russian)]
 8. Мищенко А.В., Кошелев П.С. Оптимизация управления работами логистического проекта в условиях неопределенности // Изв. РАН. ТиСУ. – 2021. – № 4. – С. 123–134. [Koshchelev, P.S., Mishchenko, A.V. Optimizing Management of Jobs in a Logistic Project Under Conditions of Uncertainty // Journal of Computer and Systems Sciences International. – 2018. – Vol. 60, no. 4. – P. 595–609.]
 9. Горский М.А., Мищенко А.В., Нестерович Л.Г., Халиков М.А. Некоторые модификации целочисленных оптимизационных задач с учетом неопределенности и риска // Изв. РАН. ТиСУ. – 2022. – № 5. – С. 106–117. [Gorskii, M.A., Mishchenko, A.V., Nesterovich, L.G., Khalikov, M.A. Some Modifications of Integer Optimization Problems with Uncertainty and Risk // Journal of Computer and Systems Sciences International. – 2022. – Vol. 61, no. 5. – P. 813–823.]
 10. Косоруков О.А., Лемтюжникова Д.В., Мищенко А.В. Методы и модели управления ресурсами проекта в условиях неопределенности // Изв. РАН. ТиСУ. – 2023. – № 3. – С. 38–56. [Kosorukov, O.A., Lemtyuzhnikova, D.V., Mishchenko, A.V. Methods and Models of Project Resource Management under Uncertainty // Journal of Computer and Systems Sciences International. – 2023. – Vol. 62, no. 2. – P. 305–323.]
 11. Yao, X., Almatooq, N., Askin, R.G., Gruber, G. Capacity Planning and Production Scheduling Integration: Improving Operational Efficiency via Detailed Modelling // Intern. J. Production Research. Published Online. – 2022. – Vol. 60, no. 1. – P. 7239–7261.
 12. Graves, S.C. How to Think About Planned Lead Times // Intern. J. Production Research. – 2022. – Vol. 60, no. 1. – DOI: <http://dx.doi.org/10.2139/ssrn.3485059>
 13. Кононов Д.А., Фуругян М.Г. Распределение неоднородного комплекса ресурсов при региональном планировании в условиях неопределенности // Труды 15 междунар. конф. «Управление развитием крупномасштабных систем» (MLSD'2022). – Москва, 2022. – С. 952–958. [Kononov, D.A., Furugyan, M.G. Planning a Complex of Works with Heterogeneous Resources under Uncertainty // Proceedings of 2022 15th International Conference Management of Large-Scale System Development (MLSD). – Moscow, 2022. – DOI: 10.1109/MLSD55143.2022.9934381]
 14. Кононов Д.А., Фуругян М.Г. Оптимизация использования неоднородного комплекса ресурсов при региональном планировании // Труды 14-й междунар. конф. «Управление развитием крупномасштабных систем» (MLSD'2021). – Москва, 2021. – С. 1231–1237. [Kononov, D.A., Furugyan, M.G. Effective Means of Regional Management: Optimal Use of Heterogeneous Resources // Proceedings of 2021 14th International Conference Management of Large-Scale System Development (MLSD'2021). – Moscow, 2021. – DOI: 10.1109/MLSD52249.2021.9600251]
 15. Фуругян М.Г. О приближении решений некоторого класса бесконечных антагонистических игр // Вестн. МГУ. Сер. 15. Вычисл. математика и кибернетика. – 1978. – № 2. – С. 81–85. [Furugyan, M.G. O priblizhenii reshenii nekotorigo klassa beskonечnykh antagonisticheskikh igr // Vestn. MGU. Ser. 15. Vychisl. matematika i kibernetika. – 1978. – No. 2. – P. 81–85. (In Russian)]
- Статья представлена к публикации членом редколлегии
А.А. Лазаревым.
- Поступила в редакцию 14.04.2025,
после доработки 30.06.2025.
Принята к публикации 16.07.2025.
- Кононов Дмитрий Алексеевич** – д-р техн. наук, Институт проблем управления им. В. А. Трапезникова РАН, г. Москва,
✉ dmitrykon52@gmail.com
ORCID iD: <https://orcid.org/0000-0002-6059-5590>
- Фуругян Меран Габидуллаевич** – канд. физ.-мат. наук, Федеральный исследовательский центр «Информатика и управление» РАН, г. Москва,
✉ rtccas@yandex.ru
ORCID iD: <https://orcid.org/0000-0003-0373-9233>
- © 2025 г. Кононов Д. А., Фуругян М. Г.



Эта статья доступна по [лицензии Creative Commons «Attribution» \(«Атрибуция»\) 4.0 Всемирная](https://creativecommons.org/licenses/by/4.0/).



MANAGING A COMPLEX OF JOBS WITH UNCERTAIN EXECUTION REQUEST ARRIVALS

D. A. Kononov* and M. G. Furugyan**

*Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia

**Federal Research Center "Computer Science and Control," Russian Academy of Sciences, Moscow, Russia

*✉ dmitrykon52@gmail.com, **✉ rtscas@yandex.ru

Abstract. This paper considers the scheduling problem for a complex of basic jobs under the condition that at some uncertain times, execution requests for supplementary higher-priority jobs are received. If a supplementary job request arrives during the execution of a basic job, then the latter is terminated and must be restarted at some time upon the complete service of the former. All jobs (basic and supplementary) are executed without interruption. By assumption, during the execution of a basic job, two or more supplementary job requests are unlikely to arrive, and such cases are not analyzed. Also, a supplementary job request can arrive only after the complete service of the previous supplementary request. Two problem formulations are studied as follows. In the first, the performance criterion is the completion time of the basic job complex, and the problem is to minimize this time. In the second formulation, the probability of a collision is minimized, as a situation where a supplementary job request arrives during the execution of a basic job. The problems are solved via their reduction to infinite zero-sum two-player (antagonistic) games and the discrete approximation of the latter by finite games. Model examples are considered. The problem formulation with non-fixed durations of the basic jobs, linearly dependent on the amount of additional resources allocated, is investigated as well. In this case, job scheduling is reduced to a linear programming problem.

Keywords: job scheduling, collision, zero-sum two-player game, antagonistic game, non-renewable resources, mixed strategy, optimal strategy.

МЕХАНИЗМЫ УСТОЙЧИВОГО РАЗВИТИЯ ТРАНСПОРТНОЙ КОРПОРАЦИИ С КОНТРОЛИРУЕМЫМ ОБУЧЕНИЕМ

В. В. Цыганов

Институт проблем управления им. В. А. Трапезникова РАН, г. Москва

✉ bbc@ipu.ru

Аннотация. Устойчивое развитие транспортной корпорации достигается последовательным сбалансированным улучшением ее экономических и экологических показателей. Этому в решающей степени способствует снижение расхода топлива ее транспортными средствами, экономящее издержки и уменьшающее вредные выбросы в окружающую среду. В связи с этим рассмотрена трехуровневая структура управления расходом топлива в транспортной корпорации в условиях неопределенности, вызываемой как случайными внешними факторами, так и нежелательной активностью сотрудников корпорации, связанной с наличием у них собственных целей, не совпадающих с целью корпорации в целом. На верхнем уровне этой структуры находится Центр, на среднем – Интендант, ответственный за расход топлива в подразделении корпорации, а на нижнем – Водитель принадлежащего корпорации транспортного средства. Моделируется устойчивое поведение сотрудников корпорации (Интенданта и Водителя), ориентированное на снижение расхода топлива в пределах их компетенции. Разработан организационный механизм устойчивого поведения Интенданта, включающий процедуры контролируемого обучения Центра и стимулирования интенданта. Предложен также организационный механизм устойчивого поведения Водителя, основанный на предоставлении Центром возможностей Интенданту для контролируемого обучения, а также на делегировании Центром Интенданту полномочий стимулирования Водителя. Доказано, что комплексирование этих двух механизмов обеспечивает снижение расхода топлива и устойчивое развитие транспортной корпорации. При этом устойчивое поведение ее сотрудников – Интенданта и Водителя – обеспечивает устойчивость корпорации. Полученные результаты проиллюстрированы на примере механизма управления расходом топлива в холдинге «РЖД».

Ключевые слова: транспорт, корпорация, устойчивое развитие, управление, цифровизация, обучение, устойчивое поведение.

ВВЕДЕНИЕ

Устойчивое развитие (англ. *sustainable development*) подразумевает «удовлетворение потребностей сегодняшнего дня, которое не лишает будущие поколения возможности удовлетворить их собственные потребности» [1]. Текущие цели глобального устойчивого развития [2] включают:

- эффективное использования ресурсов;
- уменьшение выбросов, вредных для здоровья людей и окружающей среды;
- создание механизмов планирования и управления, связанных с изменением климата.

В ряду этих целей выделяется создание экономических и экологических транспортных систем [2]. На транспорт приходится 20–25 % мирового потребления энергии и выбросов углекислого газа [3]. Почти 97 % выбросов транспортом вредных веществ в атмосферу приходится на газы, образующиеся в процессе сжигания топлива [4]. В их числе – оксиды азота и твердые частицы, отрицательно влияющие на здоровье людей [5], а также оксиды углерода и углеводороды, относящиеся к парниковым газам [6]. При этом среди всех отраслей экономики наиболее быстрым ростом выбросов парниковых газов отличается транспортная отрасль [7].



Таким образом, для устойчивого развития транспорта важна как его эффективность, так и воздействие на окружающую среду [8]. Для достижения глобальных целей необходимо применение компаниями устойчивых методов хозяйствования [2]. В связи с этим в последние десятилетия получила широкое распространение концепция устойчивого развития компаний (ESG) на принципах ответственного отношения к окружающей среде (E – *environment*), социальной ответственности (S – *social*) и эффективного корпоративного управления (G – *governance*) [9]. Соответственно, цели устойчивого развития транспортной корпорации связаны с улучшением экономических и экологических ее показателей как с помощью улучшения технологий, так и путем активизации, обучения и развития сотрудников (см., например, источники [10, 11]). Изучение связанных с этим психологических аспектов устойчивого развития началось в работе [12]. Дальнейшие исследования привели к формированию представления об устойчивом поведении [13]. Такое поведение подразумевает понимание работником важности повышения эффективности производственной деятельности при сохранении окружающей среды. Следствием устойчивого поведения является бережливое отношение к используемым ресурсам и окружающей среде, которые ведут к чувству удовлетворения и внутренней мотивации сотрудников [13].

При этом необходимо учитывать, что подавляющее большинство корпоративных транспортных средств (КТС) оборудованы двигателями внутреннего сгорания. На эти КТС приходится большая часть расхода топлива корпораций. Например, до 80 % топлива ОАО «РЖД» расходуют локомотивы с дизельными двигателями – тепловозы [10]. Поэтому экономия топлива КТС непосредственно сказывается на снижении затрат ресурсов и издержек корпорации, а также выбросов вредных веществ в выхлопных газах (в том числе парниковых газов). Исходя из этого, транспортные корпорации стремятся минимизировать расход топлива КТС путем обучения и активизации сотрудников (см., например, источники [10, 11]). Привитие сотрудникам психологии устойчивого поведения позволяет беречь топливо КТС, снижая тем самым как эксплуатационные расходы, так и выбросы загрязняющих веществ, а также обеспечивая соблюдение нормативных требований в области охраны окружающей среды.

Тем не менее, регулярно встречаются неподвижные КТС с двигателями, работающими на холостом ходу. Дело в том, что в практике планирования крупной транспортной корпорации будущие лимиты (нормы) расхода топлива ее подраз-

делением и отдельным работником обычно снижаются при снижении текущего расхода топлива. Однако чем ниже лимиты (нормы), тем меньше топлива имеется в распоряжении подразделения и работника корпорации для выполнения поставленных перед ними задач. А поскольку расход топлива зависит от случайных факторов, то при неблагоприятных обстоятельствах подразделение и работник корпорации могут не выполнить эти задачи с вытекающими отсюда негативными последствиями. Поэтому дальновидный менеджер подразделения, как и опытный водитель КТС, могут быть не заинтересованы в снижении расхода топлива ниже нормы. Это типичная проблема адаптивного планирования от достигнутого уровня, являющаяся предметом изучения в теории управления организационными системами [14]. Подход к решению такого рода проблем связан с проектированием организационных механизмов функционирования корпораций [15]. Например, в статье [16] рассмотрен механизм стимулирования внедрения экологических требований к локомотивам в ОАО «РЖД».

В данной статье разрабатываются модели обучения, активизации и устойчивого поведения менеджера и сотрудника подразделения корпорации, ориентированные на снижение расхода топлива КТС, и соответствующие организационные механизмы устойчивого развития транспортной корпорации.

1. УСТОЙЧИВОЕ ПОВЕДЕНИЕ ИНТЕНДАНТА

1.1. Общие предположения

Топливо в корпорации расходуется как на обеспечение потребностей КТС (например, грузовиков или тепловозов), так и на прочие нужды. Эти потребности и нужды зависят от случайных внешних и внутренних факторов, многие из которых становятся известны лишь узкому кругу причастных лиц непосредственно в процессе выполнения ими своих обязанностей. Соответственно, будем предполагать, что лицо, ответственное за расход топлива в корпорации (кратко – Центр) не имеет точной информации о реальных возможностях снижения расхода топлива в ее подразделении и его КТС. Но Центр может воспользоваться советами Консультанта. Оба они получают информацию о фактических расходах топлива подразделением. Кроме того, Центр может устанавливать нормы такого расхода. На основании отклонения фактических расходов топлива от норм Центр может стимулировать лицо, ответственное за расход топлива в подразделении (кратко – Интенданта). При этом Центр не может определить, является ли фак-

тический расход топлива в подразделении минимально достижимым в сложившихся условиях.

Исходя из вышесказанного, рассмотрим трехуровневую систему управления расходом топлива в корпорации, на верхнем уровне которой находится управляющий орган (Центр), на среднем – лицо, ответственное за расход топлива в подразделении (Интендант) и на нижнем – водитель КТС (Водитель). Требуется построить организационный механизм функционирования этой системы, обеспечивающий минимальный расход топлива в условиях неопределенности, обусловленной как случайными внешними факторами, так и нежелательной активностью сотрудников корпорации (Интенданта и Водителя), связанной с наличием у них собственных целей, не совпадающих с целью Центра.

Введем следующие обозначения: t – период времени, $t = 0, 1, \dots$; d_t – расход топлива подразделением в периоде t , складывающийся из расхода топлива КТС c_t и прочего расхода топлива подразделением s_t : $d_t = c_t + s_t$. Значение c_t Водитель сообщает Интенданту в периоде t . При этом предполагается, что $c_t \in C_t = [L(t), \gamma]$, где $L(t)$ – стационарный случайный процесс, определяющий реализацию (значение) l_t случайной величины минимального расхода топлива КТС в периоде t , $L(t) \in \Lambda = [\delta, \varepsilon]$, $\delta > 0$, $\varepsilon \leq \gamma$. Следовательно, $c_t \in C = [\delta, \gamma]$, $t = 0, 1, \dots$

Прочий расход топлива s_t в периоде t определяется самим Интендантом, $s_t \in S_t = [M(t), \alpha]$. Здесь $M(t)$ – стационарный случайный процесс, определяющий реализацию (значение) m_t случайной величины минимального прочего расхода топлива, $M(t) \in M = [\mu, \beta]$, $\mu > 0$, $\beta \leq \alpha$. Следовательно, $s_t \in \Sigma = [\mu, \alpha]$, $t = 0, 1, \dots$

Учитывая, что $d_t = c_t + s_t$, получаем $d_t \in \Psi_t = [A(t), \alpha + \gamma]$, где $A(t) = L(t) + M(t)$ – стационарный случайный процесс, определяющий реализацию (значение) a_t случайной величины минимального расхода топлива подразделением, $a_t = l_t + m_t$, $t = 0, 1, \dots$. При этом, из $c_t \in C = [\delta, \gamma]$ и $s_t \in \Sigma = [\mu, \alpha]$ следует, что $d_t \in \Delta = [\delta + \mu, \alpha + \gamma]$.

Предполагается, что в периоде t Центру известен фактический расход топлива подразделением d_t . Однако ему неизвестна реализация (значение)

a_t случайной величины минимального расхода топлива подразделением. Чтобы стимулировать Интенданта к снижению расхода топлива в условиях неопределенности, Центр присваивает ему одну из двух категорий – первой (обоснованный расход) и второй (необоснованный расход), пользуясь знанием расхода d_t , $t = 0, 1, \dots$. Необходимо сделать это с минимальными потерями.

1.2. Полная и частичная осведомленность Центра

В этом пункте предполагается, что в периоде t Центр знает реализацию a_t случайной величины минимального расхода топлива подразделением, $a_t = l_t + m_t$, $t = 0, 1, \dots$. Эту реализацию характеризует безразмерный относительный показатель $e_t = a_t / (\alpha + \gamma)$, $0 < e_t \leq 1$, $e_t \in D = [(\delta + \mu) / (\alpha + \gamma), 1]$, $t = 0, 1, \dots$. Задача состоит в том, чтобы присвоить Интенданту категорию 1 или 2, отнеся e_t к одному из двух неизвестных подмножеств D_1 и D_2 , составляющих множество D , $D_1 \cup D_2 = D$.

При полной осведомленности Центру известно, что показатель e_t принадлежит, с условной плотностью распределения $\epsilon(e_t | \eta)$ и априорной вероятностью ζ_η , $\eta = \overline{1, 2}$, к одному из двух подмножеств D_1 и D_2 , $D_1 \cup D_2 = D$. Обозначим q_{12} потери, возникающие из-за ошибочного отнесения Центром e к подмножеству D_2 , если на самом деле e принадлежит подмножеству D_1 , а q_{21} – потери, возникающие при ошибочном отнесении Центром e к подмножеству D_1 , если на самом деле e принадлежит подмножеству D_2 .

Введем параметр классификации p , разделяющий подмножества D_1 и D_2 . При этом $e \in D_1 = [(\delta + \mu) / (\alpha + \gamma), p]$, если $e \leq p$, иначе $e \in D_2 = (p, 1]$. Тогда оптимальная дихотомия $\{D_1, D_2\}$ определяется в результате решения задачи оптимизации по p , при которой минимизируются средние потери категоризации:

$$\sum_{\eta=1}^2 \sum_{\theta=1}^2 q_{\eta\theta} \int_{D_\theta} \zeta_\eta \epsilon(e | \eta) de \xrightarrow{\{D_1, D_2\}} \min. \quad (1)$$

Здесь θ – переменная, по которой происходит суммирование, $\theta = \overline{1, 2}$, D_θ , $\theta = \overline{1, 2}$, – подмножества D_1 и D_2 .



При частичной осведомленности Центру неизвестны указанные выше вероятностные характеристики. Следовательно, нельзя определить значение параметра p , решая задачу (1). Но можно, наблюдая случайную реализацию a_t минимального расхода топлива подразделением, $t = 0, 1, \dots$, получать и последовательно уточнять оценки p_t величины p , используя алгоритм контролируемого обучения с помощью Консультанта [17]. При этом Консультант, зная показатель расхода топлива (e_t), сообщает Центру свое мнение $B(e_t)$ о том, избыточен ли этот расход ($B(e_t) = 1$) или нет ($B(e_t) = 0$):

$$B(e_t) = \begin{cases} 1, & \text{если } e_t > p, \\ 0, & \text{если } e_t \leq p, \end{cases} \quad (2)$$

$$t = 0, 1, \dots$$

где $B(\cdot)$ – процедура консультирования. Тогда Центр может использовать алгоритм контролируемого обучения [17]

$$p_{t+1} = P(p_t, e_t) = p_t - \iota_t [p_t - 0,5 -$$

$$-q_{12} + (q_{12} + q_{21})B(e_t)], \quad (3)$$

$$p_0 = p^0, \quad t = 0, 1, \dots$$

где $\iota_t \in I = \left\{ \iota_t > 0 \mid \iota_t > \iota_{t+1}, \sum_{\sigma=1}^{\infty} \iota_{\sigma} < \infty \right\}$. Если $e_t \leq p_t$,

то Центр присваивает Интенданту первую категорию, в противном случае – вторую. Таким образом, категория Интенданта

$$k_t = G(p_t, e_t) = \begin{cases} 1, & \text{если } e_t \leq p_t, \\ 2, & \text{если } e_t > p_t, \end{cases} \quad (4)$$

$$t = 0, 1, \dots$$

1.3. Неосведомленность и контролируемое обучение Центра

Предположим теперь, что Центр и Консультант не знают не только указанные выше вероятностные характеристики, но и минимальный расход топлива a_t . Однако им известен фактический его расход d_t . Введем относительный показатель этого расхода $i_t = d_t / (\alpha + \gamma)$, $0 < i_t \leq 1$, $i_t \in D = [(\delta + \mu) / (\alpha + \gamma), 1]$, $i_t \geq e_t$, $t = 0, 1, \dots$. Тогда Консультант определяет свое мнение, используя в процедуре (2) наблюдаемый показатель i_t вместо

неизвестного e_t . Именно, узнав i_t , Консультант сообщает Центру свое мнение $B(i_t)$ о том, избыточен ли этот расход ($B(i_t) = 1$) или нет ($B(i_t) = 0$):

$$B(i_t) = \begin{cases} 1, & \text{если } i_t > p, \\ 0, & \text{если } i_t \leq p, \end{cases} \quad (5)$$

$$t = 0, 1, \dots$$

Соответственно, Центр рассчитывает оценку b_t параметра p_t с помощью формулы (5), заменив в алгоритме (3) неизвестное e_t на наблюдаемое i_t :

$$b_{t+1} = P(b_t, i_t) = b_t - \iota_t [b_t - 0,5 -$$

$$-q_{12} + (q_{12} + q_{21})B(i_t)], \quad (6)$$

$$b_0 = p^0, \quad t = 0, 1, \dots,$$

где $P(\cdot)$ – процедура обучения. Сопоставляя i_t и b_t по аналогии с формулой (4), Центр определяет категорию g_t расхода топлива в подразделении, возглавляемом Интендантом:

$$g_t = G(b_t, i_t) = \begin{cases} 1, & \text{если } i_t \leq b_t, \\ 2, & \text{если } i_t > b_t, \end{cases} \quad (7)$$

$$t = 0, 1, \dots,$$

где $G(\cdot)$ – процедура категоризации. По сути, значение b_t имеет смысл нормативного значения показателя i_t , в зависимости от которого определяется категория Интенданта. Принимая во внимание, что $i_t = d_t / (\alpha + \gamma)$, получаем, что $h_t = b_t (\alpha + \gamma)$ имеет смысл порогового значения расхода топлива подразделением (d_t). Если этот порог не превышен (т. е. если $d_t \leq h_t$), то Интендант получает первую категорию ($g_t = 1$) и поощрение. В противном случае ($g_t = 2$) Интендант не поощряется (или наказывается).

Процедуры консультирования (5), обучения (6) и категоризации (7) составляют организационный механизм Центра $\Phi = (B, P, G)$, предназначенный для управления расходом топлива в подразделении.

1.4. Цели и решения Интенданта при контролируемом обучении Центра

Используя механизм $\Phi = (B, P, G)$, Центр обучается с подсказкой Консультанта, назначает категорию и стимулирует Интенданта, основываясь на

наблюдении d_t и расчете показателя i_t , $t = 0, 1, \dots$. Однако Интендант лучше осведомлен о расходах топлива, чем Центр, и может использовать это в своих интересах. Предполагается, что, с началом периода t Интенданту становятся известны фактический расход топлива КТС c_t и реализация m_t случайной величины минимального прочего расхода топлива. После этого Интендант выбирает значение d_t , исходя из условия $d_t \geq f_t \equiv c_t + m_t$, так что $d_t \in Y_t = [f_t, \beta + \gamma]$, $f_t \in H = [\delta + \mu, \beta + \gamma]$, $t = 0, 1, \dots$.

Рассмотрим, как Интендант принимает решения при механизме $\Phi = (B, P, G)$. Дальновидный Интендант стремится выбрать значение d_t в периоде t так, чтобы улучшить как текущую, так и будущие категории. Для этого он может управлять величиной s_t прочего расхода топлива подразделением, а также величиной c_t расхода топлива КТС. Формально выбор d_t обусловлен стремлением Интенданта увеличить полезность категорий в периоде t :

$$T_t = T[g_t, \overline{g_{t+1}, \dots, g_{t+\lambda}}], \quad (8)$$

$$T_t \downarrow g_v, v = \overline{1, \lambda}, t = 0, 1, \dots,$$

где $T[\cdot]$ – монотонно убывающая функция своих аргументов (поскольку полезность любой категории для Интенданта убывает с ее номером); λ – дальновидность Интенданта.

Согласно процедуре (7) категория g_t зависит только от выбора Интендантом расхода топлива d_t , $d_t \in Y_t$. Однако для того, чтобы увеличить текущую полезность категорий (8), Интенданту нужно учитывать влияние d_t на категорию g_ξ , которую он получит в том или ином будущем периоде ξ , $\xi = \overline{t+1, t+\lambda}$. Согласно формуле (7) эта категория g_ξ зависит от выбора Интенданта d_ξ и оценки b_ξ в этом периоде. Далее, согласно выражению (6), оценка b_ξ зависит от выборов Интенданта d_ω и оценок b_ω в предыдущих периодах ω , $\omega = \overline{t, \xi-1}$. Рассматривая формулу (6) как рекуррентное соотношение, нетрудно получить, что категория g_ξ зависит от предыдущих выборов Интенданта d_ω , $\omega = \overline{t, \xi-1}$.

Заметим, что в начале периода ξ Интендант будет знать реализацию $f_\xi = c_\xi + m_\xi$, $f_\xi \in H$,

$\xi = \overline{t+1, t+\lambda}$. Предположим, что при этом Интенданту будут также известны будущие расходы топлива КТС c_σ и прочие расходы топлива m_σ , $\sigma = \overline{\xi+1, t+\lambda}$. Тогда Интендант может выбрать значение d_ξ так, чтобы максимизировать функцию (8) при известной реализации f_σ , $\sigma = \overline{\xi+1, t+\lambda}$ с учетом условия $d_\xi \in Y_\xi$, $\xi = \overline{t+1, t+\lambda}$. При этом, как было показано выше, Интенданту нужно учитывать влияние d_ξ на будущие категории g_σ , $\sigma = \overline{\xi+1, t+\lambda}$.

Повторяя эти рассуждения последовательно, начиная с $\xi = t+\lambda$ и заканчивая $\xi = t+1$, получаем, что дальновидному Интенданту целесообразно прогнозировать выборы d_ξ в обратной последовательности. В результате функция (8) преобразуется в прогнозную полезность категорий

$$T_t^o = T_t^o(g_t, f_{t+1}, \dots, f_{t+\lambda}) = \max_{d_{t+1} \in Y_{t+1}} \max_{d_{t+2} \in Y_{t+2}} \dots \max_{d_{t+\lambda} \in Y_{t+\lambda}} T_t, \quad (9)$$

которая зависит от реализаций f_ξ , $\xi = \overline{t+1, t+\lambda}$.

Однако Интенданту неизвестны как будущие расходы топлива КТС c_ξ , так и прочие расходы топлива m_ξ . Поэтому, чтобы выбрать значение d_t , увеличивающее прогнозную полезность категорий (9), Интендант должен иметь прогнозы расходов $f_\xi = c_\xi + m_\xi$, $\xi = \overline{t+1, t+\lambda}$. Предположим, что Интендант, руководствуясь принципом максимального гарантированного результата [14, 15], рассчитывает на самые неблагоприятные прогнозы $f_\xi \in H$, $\xi = \overline{t+1, t+\lambda}$. Тогда функция (9) преобразуется в целевую функцию Интенданта

$$\Gamma_t(d_t) = \min_{f_{t+1} \in H} \max_{d_{t+1} \in Y_{t+1}} \min_{f_{t+2} \in H} \max_{d_{t+2} \in Y_{t+2}} \dots \min_{f_{t+\lambda} \in H} \max_{d_{t+\lambda} \in Y_{t+\lambda}} T_t. \quad (10)$$

При этом множество его оптимальных выборов d_t^* , при которых достигается максимум целевой функции (10) в периоде t , имеет вид:

$$\Xi_t = \{d_t^* \in \Delta \mid \Gamma_t(d_t^*) \geq \Gamma_t(d_t), d_t \in \Delta\}, \quad (11)$$

$$t = 0, 1, \dots$$

Ниже предполагается гипотеза о благожелательности Интенданта по отношению к Центру: если $f_t \in \Xi_t$, то $d_t^* = f_t$, $t = 0, 1, \dots$. Это означает, что подразделение расходует больше топлива, только если это увеличивает целевую функцию Интенданта (10).



1.5. Организационный механизм устойчивого поведения Интенданта

Механизм $\Phi = (B, P, G)$ призван обеспечивать устойчивое поведение Интенданта, направленное на улучшение как экономических, так и экологических показателей корпорации. В соответствии с вышесказанным, такое поведение должно выражаться в стремлении Интенданта уменьшить расход топлива подразделением. Формально это отражает следующее

Определение 1. Поведение Интенданта устойчиво, если, при любом заданном расходе топлива КТС c_t расход топлива в подразделении минимален:

$$d_t^* = f_t, \quad t = 0, 1, \dots \quad (12)$$

С точки зрения практики, как уже указывалось, крупные корпорации часто планируют расход топлива от достигнутого уровня. При этом будущая норма этого расхода снижается при снижении фактического расхода топлива. В рассматриваемой модели это означает, что норматив b_{t+1} , имеющий смысл нормативного значения показателя i_{t+1} в периоде $t+1$, снижается при снижении текущего показателя расхода топлива i_t . Однако, согласно процедуре (6), чем ниже норматив b_{t+1} , тем меньше должен быть показатель расхода топлива i_{t+1} в периоде $t+1$, достаточный для получения Интендантом высшей категории. А поскольку $i_{t+1} \geq e_{t+1}$, где e_{t+1} — случайная величина, то при неблагоприятных обстоятельствах Интендант может получить низшую категорию. Поэтому дальновидный Интендант может быть не заинтересован в снижении показателя расхода топлива i_t ниже норматива b_t (за выполнение которого он получает высшую категорию в периоде t). В связи с этим необходим механизм управления, мотивирующий Интенданта к устойчивому поведению.

Утверждение 1. Поведение Интенданта при механизме $\Phi = (B, P, G)$ устойчиво.

Доказательство. Из формулы (5) следует, что $B(i_\xi)$ не убывает с ростом i_ξ при механизме $\Phi = (B, P, G)$, $\xi = \overline{t+1, t+\lambda}$. А поскольку, согласно формуле (6), $b_{\xi+1}$ не увеличивается с ростом $B(i_\xi)$, то $b_{\xi+1}$ не увеличивается с ростом i_ξ , $\xi = \overline{t, t+\lambda-1}$. Кроме того, согласно выражению (7), $g_\xi = G(b_\xi, i_\xi)$ не увеличивается с ростом b_ξ , $\xi = \overline{t+1, t+\lambda}$. Из вышесказанного

следует, что $g_\xi = G(b_\xi, i_\xi)$ не убывает с ростом i_t , $\xi = \overline{t+1, t+\lambda}$.

Кроме того, из формулы (7) следует, что $g_t = G(b_t, i_t)$ не убывает с ростом i_t . Таким образом, все аргументы функции $T[g_t, g_{t+1}, \dots, g_{t+\lambda}]$ не убывают с ростом i_t . Поэтому, согласно выражению (8), T_t не увеличивается с ростом i_t при любых реализациях $c_{t+1}, \dots, c_{t+\lambda}, m_{t+1}, \dots, m_{t+\lambda}$. Но тогда, согласно формуле (10), максимум функции $\Gamma_t(d_t)$ достигается при минимальном i_t . Отсюда, учитывая, что $i_t = d_t / (\alpha + \gamma)$, получаем, что максимум функции $\Gamma_t(d_t)$ достигается при минимальном d_t . А так как $d_t = c_t + s_t$, то, согласно выражению (11), $d_t = f_t \in \Xi_t$. Но в силу гипотезы о благожелательности Интенданта по отношению к Центру, если $f_t \in \Xi_t$, то $d_t^* = f_t$. Следовательно, по определению (12) поведение Интенданта устойчиво. ♦

По сути, при механизме $\Phi = (B, P, G)$ снижение расхода топлива подразделением, во-первых, не ухудшает текущую категорию Интенданта (7), а во-вторых, не ужесточает нормы расхода топлива в будущем. Тогда при благожелательности к Центру Интендант минимизирует расход топлива подразделением (12).

Однако одного желания Интенданта минимизировать расход топлива подразделения недостаточно. На практике на дорогах часто встречаются стоящие КТС с работающими двигателями. Чтобы этого избежать, нужно, чтобы в минимизации расхода топлива был заинтересован водитель каждого КТС.

2. УСТОЙЧИВОЕ ПОВЕДЕНИЕ ВОДИТЕЛЯ

2.1. Неосведомленность и контролируемое обучение Интенданта

Выбрав механизм $\Phi = (B, P, G)$, Центр обеспечивает устойчивое поведение Интенданта, направленное на минимизацию расхода топлива $f_t = c_t + m_t, t = 0, 1, \dots$. Здесь c_t — расход топлива КТС, сообщаемый Водителем, $c_t \in C_t = [l_t, \gamma]$ (см. п. 1.1). Но Интендант не знает реализации l_t минимального расхода топлива КТС. Для снижения расхода топлива до минимума требуется обеспечить $c_t = l_t, t = 0, 1, \dots$

Предположим, что минимальный расход топлива КТС (l_t) становится известен Водителю в начале периода t . Поскольку Интенданту значение

l_t неизвестно, то Водитель может манипулировать расходом топлива КТС c_t , выбирая $c_t > l_t$, если ему это выгодно. Интендант же должен минимизировать расход топлива КТС c_t , $t = 0, 1, \dots$. Для этого Центр предоставляет Интенданту возможность контролируемого обучения, а также делегирует Интенданту право устанавливать механизм функционирования Водителя, подобный механизму $\Phi = (B, P, G)$ (который сам Центр использует для управления Интендантом, см. п. 1.3).

Введем безразмерные относительные показатели – метрики расхода топлива КТС:

$$k_t = l_t / \gamma, \quad K_t = c_t / \gamma, \quad t = 0, 1, \dots \quad (13)$$

Тогда Интендант должен снизить значение метрики K_t до минимума k_t . Предположим, что для этого Интендант классифицирует работу Водителя в зависимости от расхода топлива КТС как удовлетворительную (класс 1) или неудовлетворительную (класс 0). Неправильная классификация приводит к издержкам. Для повышения обоснованности решений Интендант обучается классификации с помощью Оценщика подобно тому, как Центр обучается категоризации с помощью Консультанта (см. п. 1.3).

Обозначим ρ_{10} издержки, возникающие из-за неправильного присвоения Водителю класса 0 (хотя Водитель заслуживает присвоения класса 1), а ρ_{01} – издержки, возникающие из-за неправильного присвоения Водителю класса 1. Оценщик наблюдает расход топлива КТС (c_t), рассчитывает K_t согласно формуле (13) и сообщает Интенданту свое мнение $R(K_t)$ о том, является ли расход топлива КТС в периоде t избыточным. Формально, если Оценщик считает расход избыточным, то $R(K_t) = 1$, в противном случае $R(K_t) = 0$:

$$R(K_t) = \begin{cases} 1, & \text{если } K_t > \varsigma, \\ 0, & \text{если } K_t \leq \varsigma, \end{cases} \quad t = 0, 1, \dots, \quad (14)$$

где $R(\cdot)$ – процедура оценки; ς – параметр оценки, $\varsigma > 0$. Заметим, что процедура оценки (14) подобна процедуре консультирования (5). Повторяя рассуждения, проведенные в § 1, получаем, что для минимизации средних издержек классификации можно использовать алгоритм контролируемого обучения, подобный процедуре (6). Соответственно, настраиваемый параметр классификации (кратко – норма u_t) вычисляется с помощью рекуррентной процедуры, подобной формуле (6):

$$\begin{aligned} u_{t+1} &= U(u_t, K_t) = u_t - \varrho_t [u_t - 0,5 - \\ &\quad - \rho_{01} + (\rho_{01} + \rho_{10}) R(K_t)], \\ u_0 &= u^0, \quad t = 0, 1, \dots, \end{aligned} \quad (15)$$

где $U(u_t, K_t)$ – процедура нормирования, $\varrho_t \in P = \left\{ \varrho_t > 0 \mid \varrho_t > \varrho_{t+1}, \sum_{\sigma=1}^{\infty} \varrho_{\sigma} < \infty \right\}$, $t = 0, 1, \dots$. Составляя K_t и u_t , по аналогии с процедурой (7), Интендант определяет класс Водителя (v_t):

$$v_t = V(u_t, K_t) = \begin{cases} 1, & \text{если } K_t \leq u_t, \\ 0, & \text{если } K_t > u_t, \end{cases} \quad t = 0, 1, \dots \quad (16)$$

где $V(\cdot)$ – процедура классификации. Величина u_t имеет смысл нормативного значения показателя K_t , в зависимости от которого определяется класс Водителя. Принимая во внимание, что $K_t = c_t / \gamma$, получаем, что величина $n_t = \gamma u_t$ имеет смысл порогового значения расхода топлива Водителем (c_t). Если этот порог не превышен (т. е. если $c_t \leq n_t$), то Водитель получает первый класс ($v_t = 1$) и поощрение. Процедуры оценки (14), нормирования (15) и классификации (16) составляют организационный механизм Интенданта $\Pi = (R, U, V)$, используемый для управления расходом топлива Водителем.

2.2. Цели и решения Водителя при неосведомленности Интенданта

Рассмотрим, как Водитель принимает решения при механизме $\Pi = (R, U, V)$. Норма u_t в этом механизме имеет смысл верхней границы приемлемого для Интенданта показателя расхода топлива КТС. Если показатель расхода топлива K_t не превышает норму u_t ($K_t \leq u_t$), то Водителю присваивается класс 1 ($v_t = 1$), и он поощряется. Поэтому полезность дальновидного Водителя растет с текущим и v будущими классами:

$$\begin{aligned} Z_t &= Z[v_t, v_{t+1}, \dots, v_{t+v}], \\ Z_{\varphi} &\uparrow v_{\varphi}, \quad \varphi = t, t+v, \quad t = 0, 1, \dots \end{aligned} \quad (17)$$

Согласно формулам (13)–(16) полезность (17) зависит от расхода топлива КТС c_{φ} в периоде φ , $\varphi = t, t+v$. При этом Водитель знает текущую

реализацию случайной величины l_t , но не знает будущих реализаций l_χ , $\chi = t+1, t+\nu$.

Предположим, что Водителю известно лишь, что $l_\chi \in \Lambda$ и $c_\chi \in C_\chi$, $\chi = t+1, t+\nu$. При устранении неопределенности в отношении l_χ и c_χ , $\chi = t+1, t+\nu$, Водитель руководствуется принципом максимального гарантированного результата [14, 15]. Тогда, повторяя рассуждения, проведенные в § 1, получаем, что целевая функция дальновидного Водителя $\Theta_t(c_t)$ равна максимальной гарантированной полезности (17):

$$\Theta_t(c_t) = \min_{l_{t+1} \in \Lambda} \max_{c_{t+1} \in C_{t+1}} \min_{l_{t+2} \in \Lambda} \max_{c_{t+2} \in C_{t+2}} \dots \min_{l_{t+\nu} \in \Lambda} \max_{c_{t+\nu} \in C_{t+\nu}} Z_t. \quad (18)$$

Соответственно, в периоде t множество показателей c_t^* , при которых достигается максимум целевой функции (18), имеет вид:

$$\Omega_t = \{c_t^* \in C_t \mid \Theta_t(c_t^*) \geq \Theta_t(c_t), c_t \in C_t\}, \quad t = 0, 1, \dots \quad (19)$$

Примем гипотезу благожелательности Водителя к Интенданту: если $c_t \in \Omega_t$, то $c_t^* = l_t$, $t = 0, 1, \dots$. Это означает, что КТС расходует больше топлива, только если это увеличивает целевую функцию Водителя (18).

2.3. Организационный механизм устойчивого поведения Водителя

Целью Интенданта является снижение расхода топлива c_t до минимума l_t . Этому может воспрепятствовать использование Интендантом планирования от достигнутого уровня, о котором говорилось выше. В рассматриваемом случае это означает, что норма u_{t+1} снижается при снижении значения метрики расхода топлива K_t . Однако, согласно формуле (16), чем ниже норма u_{t+1} , тем меньше топлива должен израсходовать Водитель, чтобы получить класс 1 в периоде $t+1$ ($v_{t+1} = 1$). А поскольку $K_{t+1} \geq k_{t+1}$, а минимально возможная метрика расхода топлива k_{t+1} является случайной величиной, то при неблагоприятных обстоятельствах Водитель может получить класс $v_{t+1} = 0$. По этим причинам дальновидный Водитель может быть не заинтересован в снижении метрики расхода топлива K_t ниже нормы u_t .

Определение 2. Поведение Водителя устойчиво, если расход топлива КТС в каждом периоде минимален:

$$c_t^* = l_t, \quad t = 0, 1, \dots \quad \blacklozenge \quad (20)$$

Утверждение 2. Поведение Водителя при механизме $\Pi = (R, U, V)$ устойчиво.

Доказательство. При механизме $\Pi = (R, U, V)$, согласно процедуре (16), текущий класс $v_t = V(u_t, K_t)$ не снижается с уменьшением K_t . Рассмотрим зависимость будущего класса $v_\varphi = V(u_\varphi, K_\varphi)$, $\varphi = t+1, t+\nu$, от K_t . Согласно выражению (14) оценка $R(K_\varphi)$ в периоде φ не растет с уменьшением K_φ , $\varphi = t+1, t+\nu$. Из формулы (15) следует, что норма u_φ не уменьшается с уменьшением $R(K_{\varphi-1})$. Кроме того, согласно выражению (16), класс $v_\varphi = V(u_\varphi, K_\varphi)$ не снижается с уменьшением u_φ . С учетом этих трех монотонных зависимостей класс $v_\varphi = V(u_\varphi, K_\varphi)$ в периоде φ не уменьшается с уменьшением K_t при $\varphi = t+1, t+\nu$.

Согласно формуле (17) целевая функция Водителя (18) увеличивается при росте текущего и будущих классов v_φ , $\varphi = t, t+\nu$. Тогда из сказанного в предыдущем абзаце следует, что эта целевая функция $\Theta_t(K_t)$ не убывает с уменьшением K_t . А поскольку $K_t \geq k_t$, то $\Theta_t(k_t) \geq \Theta_t(K_t)$, $K_t \in K_t$. Но тогда, по определению (19), $k_t \in \Omega_t$. Отсюда, согласно гипотезе о благожелательности Водителя, $K_t^* = k_t$, и из выражения (13) следует формула (20). $\blacklozenge$

3. МЕХАНИЗМЫ ОБЕСПЕЧЕНИЯ УСТОЙЧИВОСТИ ТРАНСПОРТНОЙ КОРПОРАЦИИ

Определение 3. Транспортная корпорация устойчива, если расход топлива в ее подразделении (включая расход топлива КТС) в каждом периоде минимален:

$$d_t^* = a_t, \quad t = 0, 1, \dots \quad \blacklozenge \quad (21)$$

Утверждение 3. Для устойчивости транспортной корпорации, достаточно использовать комплексный механизм $\Sigma = (\Phi, \Pi)$, который включает механизмы $\Phi = (B, P, G)$ и $\Pi = (R, U, V)$.

Доказательство. При механизме $\Phi = (B, P, G)$, согласно утверждению 1, поведение Интенданта устойчиво. По определению 1, это означает,

что при любом заданном расходе топлива КТС c_i расход топлива в подразделении минимален: $d_i^* = f_i$, $t = 0, 1, \dots$

При механизме $\Pi = (R, U, V)$, согласно утверждению 2, поведение Водителя устойчиво. При этом по определению 2 расход топлива КТС в каждом периоде минимален: $c_t^* = l_t$, $t = 0, 1, \dots$. Подставляя последнее равенство в формулу (12) и учитывая, что $f_t = c_t + m_t$, $a_t = l_t + m_t$, получаем выражение (21). ♦

По сути, доказательство утверждения 3 основано на том, что устойчивое поведение сотрудников корпорации – Интенданта и Водителя – обеспечивает устойчивость корпорации в целом. Следовательно, комплексный механизм $\Sigma = (\Phi, \Pi)$ можно назвать организационным механизмом устойчивости транспортной корпорации. Структуру этого механизма иллюстрирует рисунок.

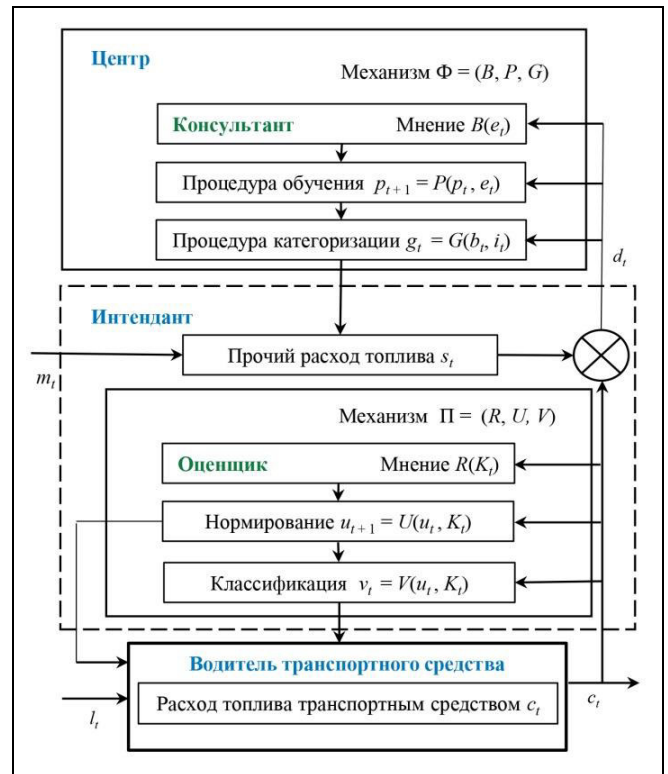
4. ПРИМЕР: МЕХАНИЗМ ОБЕСПЕЧЕНИЯ УСТОЙЧИВОСТИ ЖЕЛЕЗНОДОРОЖНОЙ КОРПОРАЦИИ

Крупномасштабные железнодорожные корпорации разрабатывают собственные стратегии устойчивого экономического и экологического – развития на основе принципов ESG (см. например, источники [10, 11]). Чтобы привлекать пассажиров и грузоотправителей, им необходимо снижать расход топлива тепловозами. Это ведет к значительному сокращению затрат заинтересованных лиц, одновременно снижая нагрузку на окружающую среду.

Рассмотрим такого рода подход к повышению устойчивости железнодорожной корпорации на примере холдинга «РЖД». Поскольку около 80 % топлива в этом холдинге расходуется тепловозами [10], то на них приходится около 80 % затрат на топливо, а также около 80 % ущерба окружающей среде от выбросов вредных веществ в отработавших газах (в том числе парниковых газов). Поэтому важнейшее значение для устойчивого развития холдинга «РЖД» имеет снижение расхода топлива тепловозами.

Разработке организационных механизмов снижения расхода топлива локомотивами холдинга «РЖД» посвящена статья [18]. В эти механизмы включены процедуры контролируемого обучения сотрудников [19]. Разработка моделей и методов управления созданием и внедрением инновационных средств и технологий снижения расхода топлива локомотивами холдинга представлена в отчете [10]. Результатом этих работ стал механизм снижения расхода топлива, обеспечивающий устойчивое развитие холдинга «РЖД» [10, 18, 19]. Этот механизм является аналогом комплексного механизма, представленного на рисунке.

Рассмотрим аналог комплексного механизма, представленного на рисунке, на региональном уровне. В нем



Организационная структура и механизм устойчивости транспортной корпорации

роль Центра играет работник региональной дирекции тяги железной дороги – филиала ОАО «РЖД», ответственный за снижение расхода топлива на этой железной дороге. В роли Интенданта выступает работник эксплуатационного локомотивного депо (ЭЛД) данного филиала ОАО «РЖД», ответственный за расход топлива в ЭЛД (кратко – менеджер), а в роли Водителя – машинист тепловоза, приписанного к ЭЛД. Указанным ответственным работникам помогают советники (назначаемые, как правило, из числа бывших ответственных работников).

В качестве КТС рассмотрим маневровый тепловоз ТЭМ18ДМ производства Брянского машиностроительного завода, входящего в состав ЗАО «Трансмашхолдинг» (далее кратко – тепловоз) [20]. Только на Октябрьской железной дороге (ОЖД) работают около двухсот тепловозов этой марки. Рассмотрим функционирование организационного механизма $\Pi = (R, U, V)$ применительно к взаимодействию менеджера и машиниста ЭЛД ОЖД при эксплуатации тепловоза данной марки (например, в ЭЛД «Санкт-Петербург-Финляндский» – ТЧЭ-12 [20]).

Согласно утверждению 2 поведение машиниста тепловоза (как Водителя) при механизме $\Pi = (R, U, V)$ устойчиво. Поэтому для сокращения расхода топлива менеджеру ЭЛД (как Интенданту) достаточно использовать данный механизм, например, в процессе ежесдельного контроля расхода топлива в ЭЛД. Именно, будем считать, что машинист сообщает менеджеру по-



часовой расход топлива управляемым им тепловозом в среднем за прошлую неделю. Обозначим через t номер недели, $t = 0, 1, \dots$. Указанный почасовой расход топлива маневровым тепловозом ТЭМ18ДМ (c_t) колеблется в пределах 10–12 л/ч [20]. Поэтому, согласно сказанному в п. 1.1, положим $\delta = 10$, $\gamma = 12$, $c_t \in C = [10, 12]$, $t = 0, 1, \dots$

Таким образом, в процессе функционирования механизма $\Pi = (R, U, V)$ менеджер и оценщик используют данные о почасовом расходе топлива в среднем за прошлую неделю (c_t). Следуя формуле (13), вводится безразмерный относительный показатель расхода топлива тепловозом – метрика $K_t = c_t / 12$; $5/6 < K_t \leq 1$, $t = 0, 1, \dots$. Далее, следуя формуле (14), оценщик сообщает менеджеру свое мнение $R(K_t)$ о том, избыточен ли этот расход (и тогда $R(K_t) = 1$) или нет (и тогда $R(K_t) = 0$). После этого менеджер рассчитывает норму u_t расхода топлива тепловозом с помощью процедуры (15). Затем, используя процедуру (16), менеджер определяет класс машиниста v_t .

Процедуры оценки (14), нормирования (15) и классификации (16) и составляют организационный механизм $\Pi = (R, U, V)$, используемый менеджером для снижения расхода топлива машинистом тепловоза. Для проведения модельного расчета предположим, что по окончании недели t в течение квартала (т. е. $t = \overline{0, 11}$) машинист сообщает менеджеру средний почасовой расход топлива управляемым им тепловозом c_t , указанный в верхней строке таблицы. Для расчета величин $R(K_t)$, v_t , u_t по формулам (14)–(16) примем следующие значения параметров механизма $\Pi = (R, U, V)$: $\varsigma = 0,92 \cong 11/12$, $u_0 = 0,90$, $\rho_{01} = 0,5$, $\rho_{10} = 1$, $\varrho_t = 1/(t+10)$, $t = \overline{0, 11}$. При этом формулы (14) и (15) приобретают соответственно вид:

$$R(K_t) = \begin{cases} 1, & \text{если } K_t > 0,92, \\ 0, & \text{если } K_t \leq 0,92, \end{cases} \quad (22)$$

$$t = \overline{0, 11},$$

$$u_{t+1} = \{u_t(t+9) + 1,5[1 - R(K_t)]\} / (t+10), \quad (23)$$

$$u_0 = 0,9, \quad t = \overline{0, 11}.$$

Результаты расчетов величин v_t , $R(K_t)$ и u_t соответственно по формулам (16), (22) и (23) за 12 недель квартала ($t = \overline{0, 11}$) сведены в таблицу.

Рассмотренный пример иллюстрирует простоту и прозрачность разработанного организационного механизма $\Pi = (R, U, V)$, а также применимость доказанного утверждения.

В целях обеспечения устойчивости холдинга «РЖД» в концепцию требований к локомотивам холдинга в области охраны окружающей среды внесено положение об использовании разработанного организационного механизма, в том числе о гибкой настройке нормативов расхода топлива для филиального и деповского уровня иерархии управления тягой, а также для стимулирования причастных работников.

ЗАКЛЮЧЕНИЕ

Чтобы устойчиво развиваться, транспортная корпорация должна снижать как экономические издержки, так и ущерб окружающей среде, наносимый ее деятельностью. Экономия топлива ее транспортными средствами существенно снижает как эксплуатационные издержки корпорации, так и выбросы загрязняющих веществ в окружающую среду. Поэтому устойчивость транспортной корпорации зависит от эффективности управления расходом топлива КТС.

Тем не менее, на дорогах по-прежнему встречаются стоящие без движения КТС с работающими двигателями. Чтобы избежать этого, транспортная корпорация должна обеспечить устойчивое поведение руководителей среднего звена и водителей, улучшающее ее экономические и экологические показатели.

Предложенный в работе комплексный механизм направлен на обеспечение устойчивого поведения сотрудников транспортной корпорации, заинтересовывая их в повсеместном снижении расхода топлива. Полученные в работе результаты

Расход топлива тепловозом и расчетные величины организационного механизма в течение квартала

t	0	1	2	3	4	5	6	7	8	9	10	11
c_t	10,7	11,1	11,0	11,3	10,9	11,0	11,2	10,8	11,2	10,7	10,6	11,1
K_t	0,89	0,93	0,92	0,94	0,91	0,92	0,93	0,90	0,93	0,89	0,88	0,93
v_t	1	1	0	0	0	0	1	0	0	0	1	1
$R(K_t)$	0	1	0	1	0	0	1	0	1	0	0	1
u_t	0,90	0,96	0,87	0,92	0,85	0,90	0,94	0,88	0,92	0,87	0,90	0,93

были использованы при разработке концепции требований к локомотивам ОАО «РЖД» в области охраны окружающей среды. Рекомендуются применение полученных результатов и для КТС, используемых корпорациями автомобильного и водного транспорта.

Дальнейшие исследования и разработки в этой области могут быть связаны:

- с разработкой механизмов устойчивого развития транспортных корпораций с альтернативными процедурами обучения;
- обобщениями полученных результатов для решения других задач устойчивого развития;
- внедрением полученных теоретических результатов на практике.

ЛИТЕРАТУРА

1. *Наше общее будущее: Доклад Международной комиссии по окружающей среде и развитию* / пер. с англ. под ред. С.А. Евтеева, Р.А. Перелета – М.: Прогресс, 1989. – 248 с. [*Our Common Future: Report of the World Commission on Environment and Development*. – New York: United Nations, 1987. – 247 p.]
2. *Transforming our world: the 2030 Agenda for Sustainable Development*. UN General Assembly Resolution A/RES/70/1. – New York: United Nations, 2015. – 35 p. – URL: https://www.un.org/en/development/desa/population/migration/generalassembly/docs/globalcompact/A_RES_70_1_E.pdf (дата обращения 17.04.2025). [Accessed April 17, 2025.]
3. *Caille, A., Al-Moneef, M., Castro, F.B., et al. Transport Technologies and Policy Scenarios to 2050*. – London: World Energy Council, 2007. – 140 p. – URL: <https://web.archive.org/web/20081204051628/http://www.worldenergy.org/publications/809.asp> (дата обращения 17.04.2025). [Accessed April 17, 2025.]
4. *About Transportation & Climate Change: Transportation's Role in Climate Change*. – Washington: US Department of Transportation, 2015. – URL: <https://web.archive.org/web/20151031215015/http://climate.dot.gov/about/transportations-role/overview.html> (дата обращения 17.04.2025). [Accessed April 17, 2025.]
5. *Kim, M.-K., Park, D., Kim, D.Y. Quantification of the Ecological Value of Railroad Development Areas Using Logistic Regression Analysis* // *International Journal of Environmental Research and Public Health*. – 2021. – Vol. 18, no. 22. – Art. no. 11764.
6. *Iliev, S., Stanchev, H., Mitev, E. An Experimental Comparison of Emissions of a Common-Rail Diesel Engine Fueled with Ethanol and Butanol Additives* // *AIP Conference Proceedings*. – 2023. – Vol. 2868. – Art. no. 020018.
7. *Metz, B., Davidson, O.R., Bosch, R., et al. Transport and Its Infrastructure* / In: *Climate Change 2007: Mitigation of Climate Change*. – New York: Cambridge University Press. – 2007. – P. 323–386. – URL: <https://www.ipcc.ch/report/ar4/wg3/transport-and-its-infrastructure/> (дата обращения 17.04.2025). [Accessed April 17, 2025.]
8. *Савосина М.И. Оценка эффективности устойчивого развития транспорта* // *Мир транспорта*. – 2020. – Т. 18, № 2. – С. 50–66. [*Savosina, M.I. Assessment of Effectiveness of Sustainable Transport Development* // *World of Transport and Transportation*. – 2020. – Vol. 18, no. 2. – P. 50–66.]
9. *Pollman, E. The Making and Meaning of ESG* // *Harvard Business Review*. – 2024. – Vol. 14. – P. 403.
10. *Бурков В.Н., Еналеев А.К., Цыганов В.В. и др. Модели и методы управления разработкой и внедрением инновационных энергоэффективных средств и технологий в ОАО «РЖД»* // *Отчет о выполнении проекта РФФИ № 17-20-05216*. – М.: ИПУ РАН, 2018. – 148 с. [*Burkov, V.N., Yenaleyev, A.K., Stroganov, V.I., et al. Modeli i metody upravleniya razrabotkoy i vnedreniyem innovatsionnykh energoeffektivnykh sredstv i tekhnologiy v ОАО «RZHD»* // *Отчет о выполнении проекта РФФИ № 17-20-05216*. – М.: ИПУ РАН, 2018. – 148 с. (In Russian)]
11. *Union Pacific: Sustainability*. – URL: <https://www.up.com/about/esg/index.htm> (дата обращения 27.07.2025). [Accessed July 27, 2025.]
12. *Corral-Verdugo, V., Frías, M., García, C. Introduction to the Psychological Dimensions of Sustainability* / In: *Psychological Approaches to Sustainability*. Ed. by V. Corral, C. García, M. Frías. – New York: Nova Science Publishers, 2010. – P. 3–18.
13. *Tapia-Fonllem, C., Corral-Verdugo, V., Fraijo-Sing, B. Sustainable Behavior and Quality of Life* / In: *Handbook of Environmental Psychology and Quality of Life Research*. Ed. by G. Fleury-Bahi, E. Pol, O. Navarro. – Cham: Springer, 2017. – P. 265–268.
14. *Новиков Д.А. Теория управления организационными системами*. – М.: ЛЕНАНД, 2022. – 500 с. [*Novikov, D.A. Teoriya upravleniya organizatsionnymi sistemami*. – М.: LENAND, 2022. – 500 s. (In Russian)]
15. *Burkov, V., Kondratiev, V., Korgin, N., Novikov, D. Mechanism Design and Management. Mathematical Methods for Smart Organizations*. – New York: Nova Science, 2013. – 187 p.
16. *Enaleev, A. Comprehensive Assessment and Incentive Mechanism to Implement Environmental Requirements for Locomotives* // *Proceedings of the 6th International Conference on Control Systems, Mathematical Modeling, Automation and Energy Efficiency*. – Lipetsk, 2024. – P. 804–809.
17. *Цыганов В.В. Стратегическое адаптивное управление транспортом* // Тр. 17-й международной конференции «Управление развитием крупномасштабных систем». – Москва, 2024. – С. 52–62. [*Tsyganov, V., Strategic Adaptive Transport Management* // *Proceedings of 2024 17th International Conference on Management of Large-Scale System Development (MLSD)*. – Moscow, 2024. – P. 1–5. – DOI: 10.1109/MLSD61779.2024.10739611]
18. *Цыганов В.В., Федянин Д.Н. Механизмы наставничества в организационных системах и повышение энергоэффективности железнодорожного транспорта* // *Вестник Уральского государственного университета путей сообщения*. – 2019. – № 2. – № 3 (43). – С. 55–66. [*Tsyganov, V.V., Fedyanin, D.N. Mekhanizmy nastavnichestva v organizatsionnykh voprosakh i povyshenii energoeffektivnosti zheleznodorozhnogo transporta* // *Vestnik Ural'skogo gosudarstvennogo universiteta putey soobshcheniya*. – 2019. – No. 2. – No. 3 (43). – S. 55–66. (In Russian)]
19. *Цыганов В.В., Басыров С.К., Ковалев С.М. Механизм обучения энергоэффективности на цифровом железнодорожном транспорте* // Тр. VIII научно-технической конференции «Интеллектуальные системы управления на железнодорожном транспорте. Компьютерное и математическое моделирование». – Москва, 2019. – С. 185–190.



[Tsyganov, V.V., Basyrov, S.K., Kovalev, S.M. Mekhanizm obucheniya energoeffektivnosti na tsifrovom zheleznodorozhnom transporte / Tr. VIII nauchno-tekhnicheskoy konferentsii «Intellectual'nyye sistemy upravleniya na zheleznodorozhnom transporte. Komp'yuternoye i matematicheskoye modelirovaniye». – Moscow, 2019. – S. 185–190. (In Russian)]

20. Генкин Д. Новые маневровые // Октябрьская магистраль. – 28.11.2019. – URL: <https://company.rzd.ru/ru/9401/page/78314?accessible=true&id=182659> [Genkin, D. Novye manevrovye // Oktjabr'skaja magistral'. – November 28, 2019. – URL: <https://company.rzd.ru/ru/9401/page/78314?accessible=true&id=182659> (In Russian)]

Статья представлена к публикации членом редколлегии академиком РАН Д. А. Новиковым.

Поступила в редакцию 22.04.2025,
после доработки 20.06.2025.
Принята к публикации 25.06.2025.

Цыганов Владимир Викторович – д-р техн. наук, Институт проблем управления им. В. А. Трапезникова РАН, г. Москва
✉ bbc@ipu.ru
ORCID iD: <https://orcid.org/0000-0001-9274-0548>

© 2025 г. Цыганов В. В.



Эта статья доступна по лицензии Creative Commons «Attribution» («Атрибуция») 4.0 Всемирная.

SUSTAINABLE DEVELOPMENT MECHANISMS FOR A TRANSPORT CORPORATION WITH SUPERVISED LEARNING

V. V. Tsyganov

Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia

✉ bbc@ipu.ru

Abstract. Sustainable development of a transport corporation is achieved through a consistent and balanced improvement of its economic and environmental performance indicators. This is largely facilitated by reducing fuel consumption by corporate vehicles, which saves costs and decreases harmful emissions into the environment. In this regard, a three-level structure for managing fuel consumption in a transport corporation is considered under uncertainty caused by both random external factors and the undesirable activity of corporate employees associated with their private goals not coinciding with the goal of the entire corporation. The top level of this structure is the Principal, the middle level is the Intendant, responsible for fuel consumption in a corporate division, and the bottom level is the Driver of a corporate vehicle. The sustainable behavior of corporate employees—the Intendant and Driver—to reduce fuel consumption within their competence is modeled. An organizational mechanism for the Intendant's sustainable behavior is developed, including supervised learning procedures for the Principal and incentive procedures for the Intendant. An organizational mechanism for the Driver's sustainable behavior is also proposed: the Principal provides the Intendant with supervised learning opportunities and delegates his authority to incentivize the Driver. As proven, the integration of these two mechanisms reduces fuel consumption and ensures sustainable development of the transport corporation. In addition, the sustainable behavior of corporate employees—the Intendant and Driver—ensures the sustainability of the entire corporation. The results are illustrated by an example of the fuel consumption management mechanism in JSC Russian Railways.

Keywords: transport, corporation, sustainable development, control, digitalization, supervised learning, sustainable behavior.

ИССЛЕДОВАНИЕ ВЗАИМОСВЯЗИ АКТИВНОСТИ В СОЦИАЛЬНОЙ СЕТИ И АКАДЕМИЧЕСКОЙ УСПЕВАЕМОСТИ СТУДЕНТОВ ТЕХНИЧЕСКОГО УНИВЕРСИТЕТА

В. В. Сыч

МГТУ им. Н.Э. Баумана;
Институт проблем управления им. В.А. Трапезникова РАН, г. Москва

✉ sycheslavovich@gmail.com

Аннотация. Проанализирована академическая успеваемость студентов МГТУ им. Н. Э. Баумана и её связь с их активностью в социальной сети ВКонтакте. С помощью методов машинного обучения выделены различные траектории успеваемости, отражающие динамику образовательных достижений. На основе анализа подписок студентов идентифицированы маркерные сообщества, характеризующие преобладание студентов, относящихся к определённым категориям успеваемости. С применением теории графов выполнена кластеризация сообществ, выявившая структурные группы интересов студентов. Для каждой траектории построены стохастические векторы долей интересов к кластерным сообществам, позволившие определить кластеры сообществ, статистически значимо связанные с конкретными траекториями успеваемости. Полученные результаты подтвердили наличие связи между цифровым поведением студентов и их академическими результатами, что способствует разработке моделей прогнозирования успеваемости с учётом интересов студентов в социальных сетях.

Ключевые слова: академическая успеваемость, ВКонтакте, машинное обучение, теория графов, прогнозирование успеваемости.

ВВЕДЕНИЕ

В условиях цифровизации образовательного процесса данные, генерируемые студентами в социальных сетях, становятся важным источником информации для анализа их академической и социальной активности. Ранее в ряде исследований была показана возможность дифференцировать студентов по уровню академической успеваемости на основе их активности в социальной сети ВКонтакте [1]. Также изучалась зависимость карьерных достижений от участия в профессиональных социальных сетях [2], а академические успехи рассматривались как распространяющиеся через дружеские и учебные связи [3]. Выявлена инвертированная U-образная зависимость между интенсивностью использования социальных сетей и академической результативностью [4], а также установлено, что психическое здоровье может выступать посредником в этой взаимосвязи [5].

Подавляющее большинство студентов имеет страницы в социальной сети ВКонтакте, на которых накоплен значительный объём информации, отражающий их интересы и социальные связи. Несмотря на то, что соцсети редко используют для поиска серьёзной информации и чаще открывают «для отдыха», долгосрочная структура подписок отражает устойчивые предпочтения и личные ориентиры пользователей, что позволяет связывать их цифровой след с образовательными показателями и использовать его в качестве надёжного предиктора академической успеваемости. Данное исследование направлено на выявление взаимосвязи между академической успеваемостью студентов и их поведением в цифровом пространстве и выполнено с применением методов машинного обучения и теории графов.

Актуальность исследования обусловлена необходимостью разработки эффективных инструментов мониторинга академических достижений и

раннего выявления студентов, находящихся в группе риска. Традиционные методы оценки успеваемости, основанные на ограниченных выборках и опросах, не всегда учитывают многогранные аспекты поведения студентов вне учебных заведений. Анализ цифрового следа в социальных сетях позволяет получить более полное представление о студенте, учитывая его интересы, социальные связи и даже интересы друзей, что может дополнить существующие методы оценки.

Цель исследования состоит в анализе взаимосвязи между траекториями академической успеваемости и цифровыми интересами студентов. Уникальность работы заключается в использовании обширной базы данных: из 76 000 студентов, обучавшихся в МГТУ им. Н.Э. Баумана в период с 2014 по 2023 гг., 48 000 удалось сопоставить страницы в социальной сети ВКонтакте. Данные включают как академические показатели (результаты сессий), так и информацию о подписках и дружеских связях, что позволяет формировать комплексный цифровой портрет каждого студента. Пример графа дружеского коллектива, построенного на основе данных сети ВКонтакте, приведён на рис. 1.

Методологическая база исследования включает алгоритмы кластеризации, стандартизации и анализа главных компонент для выделения типичных траекторий успеваемости, а также методы теории графов для анализа структуры сообществ. Иссле-

дование выполнено на основе данных социальной сети ВКонтакте и электронных записей университета, анонимизированных после сопоставления.

1. ОПИСАНИЕ ДАННЫХ

Анализ проводится на основе двух основных источников информации. Первый источник – академическая база данных, полученная из системы «Электронный университет», охватывающая результаты сессий студентов за период 2014–2023 гг. Для каждого студента зафиксированы оценки по каждой сессии, что позволяет вычислить средний балл по каждому из четырёх курсов обучения. Второй источник – данные из социальной сети ВКонтакте, собранные с помощью VK API. Для каждого пользователя извлекаются списки подписок на сообщества и информация о дружеских связях, что позволяет анализировать интересы студентов и изучать организацию их сетевого окружения. Были выгружены профили, подписанные на сообщества, связанные с университетом, и все их друзья. Подобный подход выгрузки данных использовался в работе [6].

Сопоставление академических данных с профилями ВКонтакте проводилось итеративно. На первом этапе формировалось «ядро» студентов с высокой степенью совпадения данных, где сопоставление осуществлялось на основе строгого со-

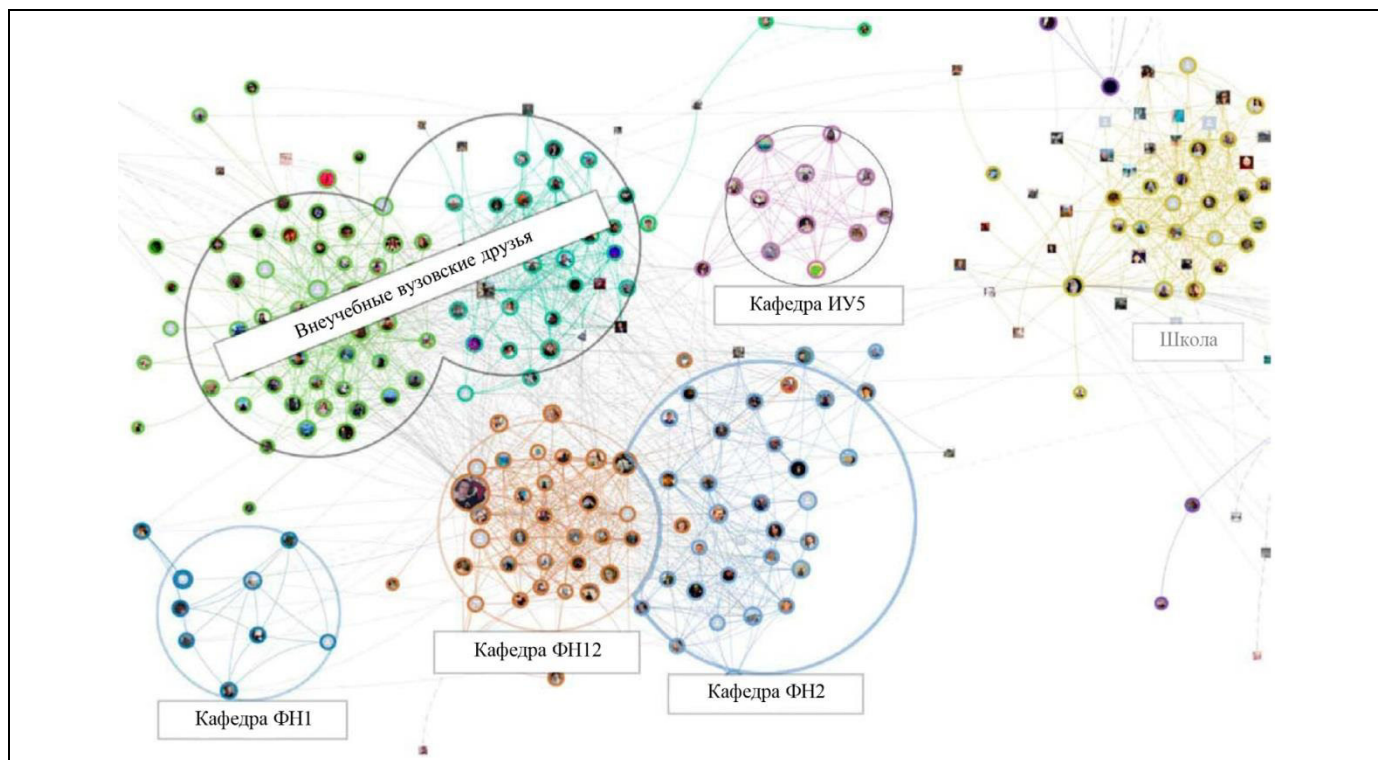


Рис. 1. Граф дружеских коллективов студента по связям из сети ВКонтакте на примере страницы автора (подписаны крупные кластеры)

ответствия Ф. И. О. и даты рождения, дополнительно расчётом расстояния Левенштейна [7] для несовпадающих форм имён. Затем с использованием информации о дружеских связях на втором этапе идентифицировались дополнительные профили: если у кандидата совпадали лишь базовые параметры (например, имя и год рождения), но при этом он был связан с уже идентифицированными студентами из конкретной учебной группы, то его с большей вероятностью относили к этой же группе. Этот процесс основывался на принципе «скажи мне, кто твой друг, и я скажу, кто ты», так как дружеские связи служили дополнительным маркером достоверности. Такой подход позволил включить в исследование тех студентов, у кого исходных совпадений было недостаточно для однозначной идентификации. Итеративное повторение процесса (на первой итерации было получено ядро в 5 000 студентов, на второй около 22 000 студентов, на третьей – 46 000) базируется на принципе: чем больше профилей друзей студента уже подтверждены, тем выше вероятность корректного сопоставления его профиля с академическими данными. Например, в случае, когда множество пользователей имеют одинаковое имя и совпадающую дату рождения, дополнительный признак в виде дружеских связей со студентами, профили которых подтверждены, существенно повышает достоверность идентификации. Такой подход позволяет значительно расширить выборку и обеспечивает её высокую репрезентативность. После проведения процедуры сопоставления данные анонимизировались и только в таком виде использовались для анализа. Подобные методики расширенного сопоставления цифрового следа с образовательными данными ранее демонстрировали высокую диагностическую ценность при исследовании успеваемости старшеклассников и студентов [8, 9].

2. АНАЛИЗ АКАДЕМИЧЕСКОЙ УСПЕВАЕМОСТИ

Для анализа траекторий успеваемости были отобраны студенты, успешно завершившие четыре курса бакалавриата или специалитета без академических задолженностей, в выборку попало 15 495 мужчин и 6 285 женщин. Кластерный подход к сегментации академических результатов ранее применялся при оценке связи между временем, проводимым в социальных сетях, и оценками студентов-биологов [10] и при анализе использования соцмедиа на мобильных устройствах в международной выборке [11], что подтверждает его пригодность для выявления латентных учебных траек-

торий. Для каждого студента i вычислялся вектор средних баллов по курсам

$$x_i = (x_{i,1}, x_{i,2}, x_{i,3}, x_{i,4}) \in \mathbb{R}^4,$$

где $x_{i,c}$ – средний балл на курсе c ($c = 1, 2, 3, 4$).

Далее матрица $X \in \mathbb{R}^{N \times 4}$ подвергалась стандартизации:

$$z_{i,c} = \frac{x_{i,c} - \mu_c}{\sigma_c},$$

где μ_c и σ_c – соответственно выборочное среднее и стандартное отклонение по курсу c .

Для разбиения студентов по траекториям академической успеваемости был применён алгоритм k -средних (k -means). Оптимальное число кластеров k выбиралось по коэффициенту силуэта (*Silhouette Score*) – показателю внутрикластерной компактности и межкластерной разделимости, предложенному П. Роусеу [12]. Максимум метрики (около 0,43) достигался при $k = 2$, однако такое разбиение не отражает разнообразия учебных траекторий. При увеличении k от 6 до 16 коэффициент силуэта стабилизировался в диапазоне 0,21–0,27, что свидетельствует об умеренном, но устойчивом качестве кластеризации.

Для дополнительного подтверждения оптимального числа кластеров был применён метод статистики разрыва (*Gap Statistic*), предложенный Р. Тибширани и соавторами [13]. Этот метод позволяет сравнить качество кластеризации исходных данных с кластеризацией случайно сгенерированных данных, распределённых равномерно в том же диапазоне. Чем выше значение статистики разрыва $Gap(k)$, тем отчётливее структура данных (выделяемые кластеры значимее случайных). Значения статистики разрыва при $k = 6$ и $k = 8$ оказались близкими: $0,994 \pm 0,006$ и $0,911 \pm 0,003$ соответственно. Незначительное уменьшение значения статистики разрыва при переходе от шести к восьми кластерам находится в пределах стандартной ошибки, формируя так называемое «плато». Это свидетельствует о том, что детализация кластеров при увеличении их числа с шести до восьми существенно не ухудшается.

Таким образом, несмотря на умеренные значения метрик качества, выбор $k = 8$ является обоснованным компромиссом между статистической устойчивостью результатов и необходимостью достаточной детализации траекторий успеваемости для последующего анализа её взаимосвязи с цифровыми интересами студентов.

После разбиения каждому студенту была присвоена метка кластера, пронумерованного по воз-

растанию общей средней успеваемости – простого арифметического среднего всех оценок за восемь семестров (табл. 1). Дополнительно отмечается тенденция снижения доли мужчин при переходе от кластеров с низкой успеваемостью к кластерам с высокой успеваемостью.

Траектории представлены на рис. 2 и 3.

Итоговое разбиение характеризуется следующими траекториями:

- **Траектория 0.** Стабильно низкая успеваемость.
- **Траектория 1.** Низкие оценки на начальных курсах с последующим улучшением.
- **Траектория 2.** Первоначально хорошие результаты с последующим снижением.
- **Траектория 3.** Низкая успеваемость с заметным ростом.

– **Траектория 4.** Стабильные хорошие оценки с небольшим проседанием.

– **Траектория 5.** Умеренные оценки, переходящие к устойчивым хорошим результатам.

Таблица 1

Статистика по траекторным кластерам

Кластер	Количество студентов	Средний балл	Доля мужчин, %
0	1910	3,48	90,0
1	2510	3,81	85,7
2	2120	3,92	82,8
3	2870	4,19	76,7
4	1950	4,21	71,6
5	2200	4,43	69,7
6	3500	4,57	52,9
7	4720	4,86	56,5

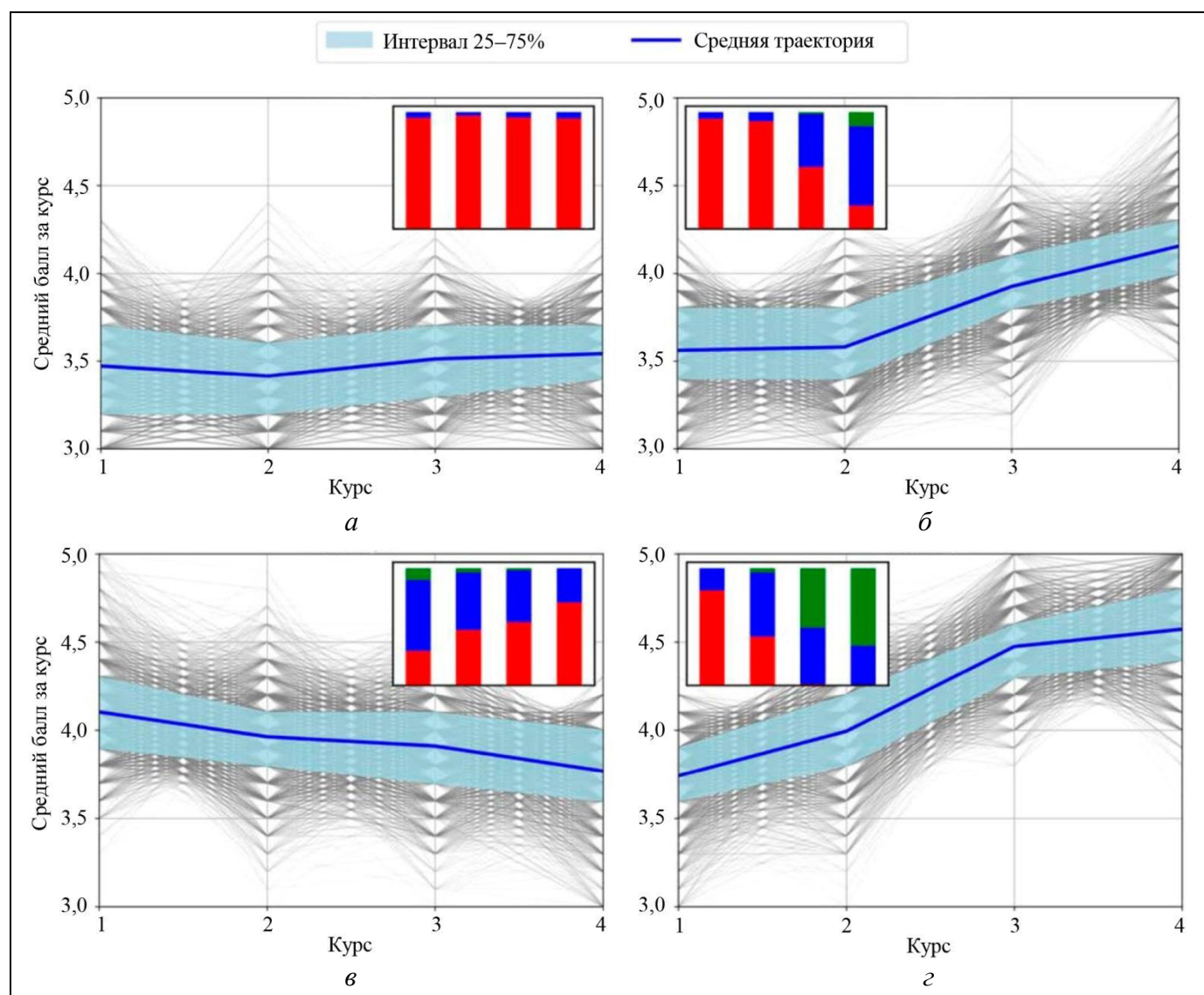


Рис. 2. Кластеры траекторий: а – кластер 0, б – кластер 1, в – кластер 2, г – кластер 3

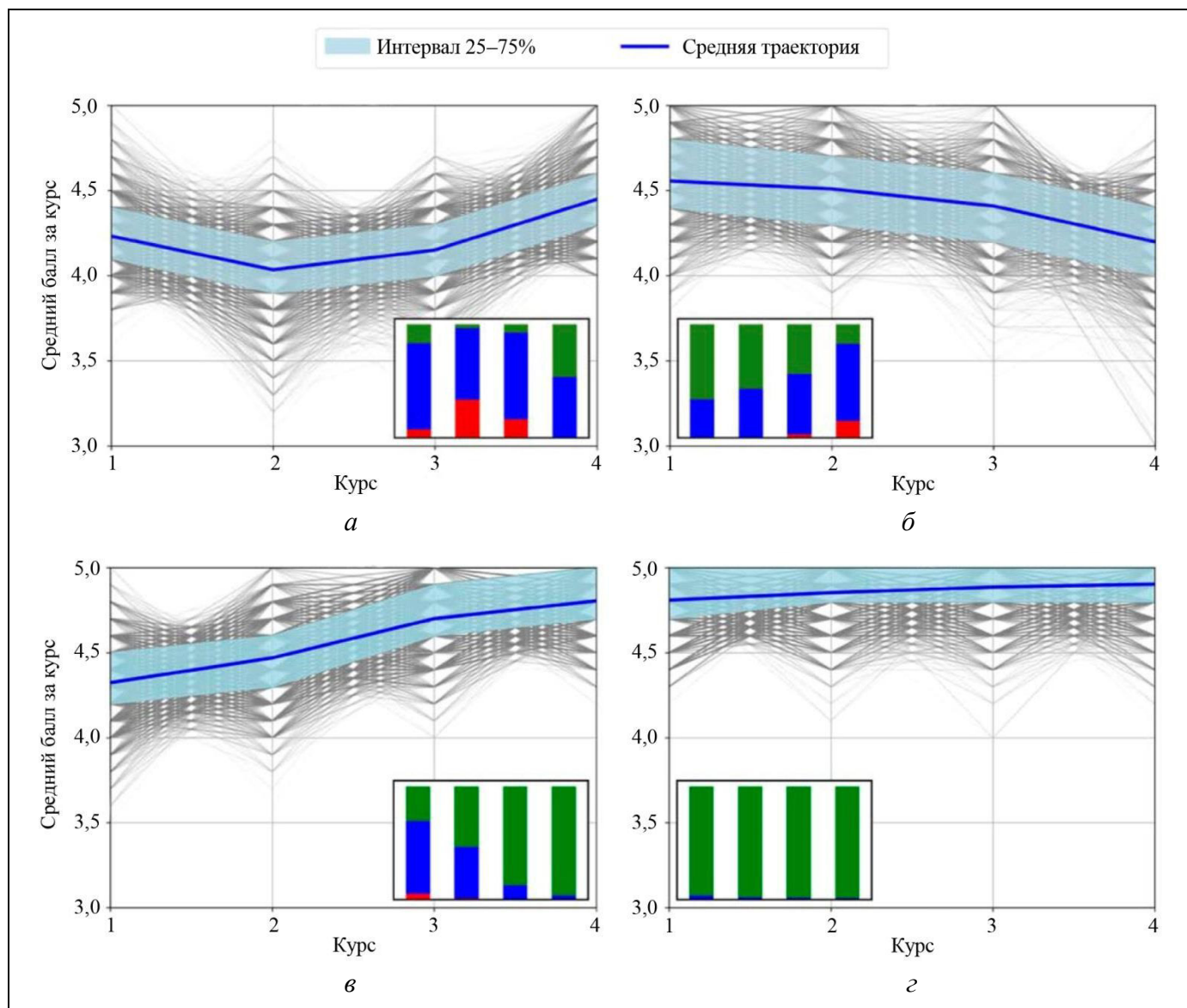


Рис. 3. Кластеры траекторий: а – кластер 4, б – кластер 5, в – кластер 6, г – кластер 7

Для наглядности в каждом кластере представлены диаграммы с пропорциями студентов различных категорий успеваемости: красные столбцы обозначают долю троечников ($\bar{g} < 3,9$), синие – хорошистов ($3,9 \leq \bar{g} < 4,5$), а зелёные – отличников ($\bar{g} \geq 4,5$), где $\bar{g}$ – средний балл студента, вычисленный как арифметическое среднее всех его оценок по сессиям. Эти диаграммы позволяют оценить распределение уровня успеваемости студентов по курсам в каждой траектории.

Для дополнительной визуализации результатов кластеризации применён метод главных компонент (англ. *Principal Component Analysis*, PCA). На рис. 4 точки, соответствующие студентам, отобра-

жены на плоскости, где горизонтальная ось отражает общий уровень успеваемости, а вертикальная – динамику изменений. Центр масс множества отмечен крестиком.

Повторные запуски алгоритма *k-means* подтвердили устойчивость полученного разбиения, а результаты визуализации демонстрируют разделение кластеров, несмотря на частичное наложение их границ в проекции на двумерное пространство. Данная структура обеспечивает надёжную основу для дальнейшего анализа взаимосвязи между цифровыми интересами и академической успеваемостью студентов, так как средние значения траекторий хорошо описывают основные различия между полученными траекторными кластерами.

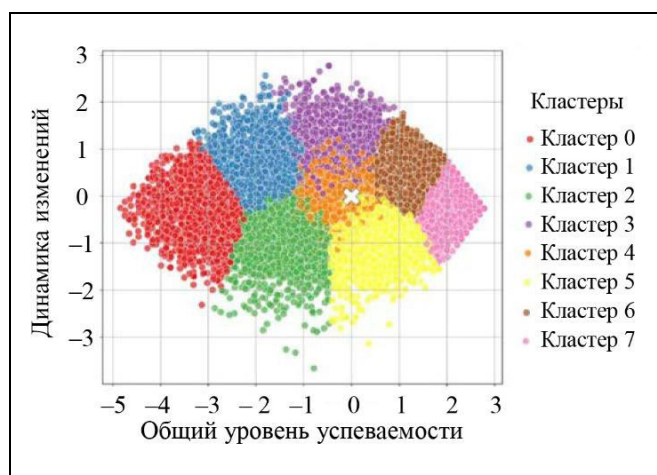


Рис. 4. Кластеры студентов по успеваемости, полученные с помощью метода главных компонент

3. ГРАФОВАЯ МОДЕЛЬ СООБЩЕСТВ В КОНТАКТЕ И ЕЁ РОЛЬ В АНАЛИЗЕ АКАДЕМИЧЕСКИХ ТРАЕКТОРИЙ СТУДЕНТОВ

3.1. Исходные данные

Исходными данными послужили все паблики (сокр. от «публичная страница», также говорят «сообщество» или «группа») ВКонтакте, на которые подписано не менее 50 студентов; таким образом сформирован корпус, содержащий порядка 4500 сообществ. Подход к использованию подписок на сообщества как признаков академической успешности зарекомендовал себя в недавней работе по выявлению лучших студентов средствами продвинутых NLP-алгоритмов [14]. Прежде чем переходить к теоретико-графовому анализу, каждое сообщество проверялось на связь с академическими траекториями посредством однофакторного дисперсионного анализа (англ. *Analysis of Variance*, ANOVA): для 760 пабликов различия между траекториями оказались значимыми при $p < 0,05$, для 350 – при $p < 0,01$.

Пусть $G=(V, E, W)$ – неориентированный граф, где V – множество пабликов, а ребро $(c_1, c_2) \in E$ существует, если найдётся хотя бы один общий подписчик-студент. Сила связи между пабликами оценивалась при помощи взвешенного коэффициента Жаккара (*weighted Jaccard coefficient*), впервые предложенного для оценки сходства флористических составов [15]:

$$w(c_1, c_2) = \ln \left(1 + \frac{|S_{c_1} \cap S_{c_2}|}{|S_{c_1} \cup S_{c_2}|} \right),$$

где S_{c_i} – студенты, подписанные на паблик c_i . Логарифмирование позволяет сгладить крайние значения коэффициента при очень больших или, напротив, почти нулевых пересечениях аудиторий.

Для обнаружения сообществ применялся алгоритм Лувена [16], максимизирующий модулярность графа. При однократном запуске на всём графе при базовом параметре $\text{resolution} = 1$ получаются несколько anomalously крупных и множество маленьких кластеров, поэтому использовалась поэтапная схема. Сначала весь граф кластеризовался при низком значении resolution ; далее слишком большие группы рекурсивно переразбивались тем же методом с $\text{resolution} = 0,7$. Дополнительно были убраны 30 % самых слабых связей, поскольку общее их количество было очень велико. Итеративное объединение результатов обеспечило финальное разбиение на 21 кластер сопоставимого размера.

3.2. Визуализация и интерпретация кластеров

Итоговая структура графа сообществ представлена на рис. 5: слева – полный граф с 21 кластером, справа – кластер 10 в приближении, включающий преимущественно образовательные паблики. Цвета узлов соответствуют преобладающему уровню академической успеваемости подписчиков: красный – троечники, синий – хорошисты, зелёный – отличники. Такое кодирование позволяет сразу оценить однородность составов кластеров по уровню успеваемости.

Отметим, что траекторные кластеры ($n = 8$) отражают динамику изменений успеваемости студентов во времени, тогда как категории «троечники» / «хорошисты» / «отличники» (3/4/5) рассчитываются по индивидуальному среднему баллу за всё время учебы и вводятся отдельно для наглядной интерпретации уровня успеваемости в каждом паблике. Они служат дополнительной «проекцией» данных.

Описание кластеров и распределение студентов по категориям успеваемости приведены в табл. 2 (описания основаны на субъективных обобщениях автора).



Рис. 5. Визуализация графа сообществ ВКонтакте

Таблица 2

Характеристика кластеров пабликов ВКонтакте по результатам кластеризации.

Кластер	Описание пабликов	3/4/5
0	Крупные сообщества с универсальным развлекательным контентом, охватывающие широкую аудиторию	41/32/27
1	Сообщества, специализирующиеся на юмористическом контенте, включая мемы, связанные с киберспортом и видеонграми (CS2, PUBG)	42/33/25
2	Высокопопулярные паблики с мемами, рассчитанные на массовую аудиторию	42/31/27
3	Паблики, посвящённые спортивным трансляциям и киберспорту (футбол, UFC, Dota2)	46/32/22
4	Развлекательные паблики с простыми, широко распространёнными мемами	42/33/25
5	Сообщества с авторскими, уникальными или нишевыми мемами	46/34/20
6	Развлекательные паблики с повторяющимися шаблонными мемами	46/33/21
7	Сообщества, распространяющие постироничные мемы, ориентированные на специфическую аудиторию	40/35/25
8	Юмористические паблики, ориентированные на распространение видеоконтента.	46/34/20
9	Сообщества, представляющие постиронию и абсурдистский юмор	42/34/23
10	Паблики, связанные с образовательными и научными темами: программирование, стажировки, исследования	27/34/39
11	Студенческие паблики, в том числе факультетские сообщества и группы типа «Подслушано»	32/35/33
12	Сообщества, ориентированные на студенческую жизнь, включая мероприятия и активности	27/35/37
13	Паблики с доброжелательными мемами, часто ориентированные на тему домашних животных (особенно котов)	32/34/34
14	Сообщества, посвящённые эстетике, фитнесу, питанию и уходу за собой, преимущественно ориентированные на женскую аудиторию	26/33/41
15	Паблики о моде, индустрии красоты, интернет-магазинах и женских мемах	25/34/41
16	Сообщества, связанные с саморазвитием, литературой, изучением иностранных языков, эстетическими концепциями	30/33/37
17	Зарубежные сообщества, содержащие женские мемы и эстетический контент	29/34/38
18	Сообщества, посвящённые неформальной моде, торговле одеждой и обувью	46/35/19
19	Музыкальные сообщества, ориентированные на рэп-исполнителей и индустрию музыки	45/35/20
20	Игровые паблики (Dota2, киберспорт), а также группы, связанные с торговлей поддержанными автомобилями	46/34/20



4. СВЯЗЬ ЦИФРОВЫХ ИНТЕРЕСОВ СТУДЕНТОВ С ИХ АКАДЕМИЧЕСКИМИ ТРАЕКТОРИЯМИ

4.1. Формализация интересов и вычисление логарифмического отношения шансов

Для каждого студента i построен стохастический вектор интересов

$$P_i = (P_{i,0}, P_{i,1}, \dots, P_{i,20}),$$

обеспечивающий равнозначный учёт всех кластеров – $\left(\sum_k P_{i,k} = 1\right)$, где $P_{i,k}$ – это отношение числа

подписок студента i в кластере k к общему числу его подписок. Этот вектор отражает степень концентрации интересов студента на конкретных кластерах пабликов.

Чтобы выявить различия в цифровых интересах студентов разных академических траекторий, далее использовалась выборка студентов, для которых были известны данные о подписках: после фильтрации осталось 4 500 мужчин и 1 900 женщин. Для оценки применялась метрика логарифмического отношения шансов (*log-odds ratio*) [17], измеряющая отклонение доли подписок на кластер k в академической траектории t от глобального среднего значения:

$$\log_odds_{k,t} = \ln\left(\frac{P_{k,t}}{1 - P_{k,t}}\right) - \ln\left(\frac{P_{k,g}}{1 - P_{k,g}}\right),$$

где

$P_{k,t}$ – усечённая (*trimmed mean*) средняя доля подписок на кластер k среди студентов траектории t . Сначала для каждого студента i из траектории t рассчитывается отношение числа его подписок в кластере k к общему числу его подписок, а затем по полученным отношениям всех студентов траектории t берётся усечённая средняя;

$P_{k,g}$ – глобальная усечённая средняя доля подписок на кластер k по всей выборке студентов. Сначала для каждого студента рассчитывается отношение числа его подписок в кластере k к общему числу его подписок, а затем по всем этим отношениям берётся усечённая средняя.

Чтобы уменьшить искажение из-за нетипично низких или высоких индивидуальных значений, расчёт средних долей для каждого кластера проводился с использованием усечённого среднего: по 10 % самых низких и самых высоких значений долей исключались из анализа. Это позволило полу-

чить более надёжные и информативные оценки, устойчивые к наличию выбросов.

Положительные значения $\log_odds$ говорят о повышенном интересе к кластеру среди студентов конкретной траектории, отрицательные – о пониженной доле подписок по сравнению с глобальной средней $P_{k,g}$, т. е. об «избегании» этого кластера. Для наглядной интерпретации этой метрики можно привести следующие примеры:

- $\log_odds = 0,5$ означает, что шансы встретить подписки на данный кластер среди студентов указанной траектории выше примерно на 65 % по сравнению с общим уровнем, так как $e^{0,5} \approx 1,65$;
- $\log_odds = -1$ означает, что шансы встретить подписки на данный кластер среди студентов указанной траектории примерно в $e^{-1} \approx 0,37$ раза, т. е. в $1/0,37 \approx 2,7$ раза ниже, чем в среднем.

В дальнейшем анализе логарифмические коэффициенты ($\log_odds$) вычислялись отдельно для мужчин и женщин. Причины раздельного расчёта:

• **Неровный гендерный состав траекторий.** В траекторном кластере 0 мужчин около 90 %, в траектории 7 – лишь около 57 %. При агрегации это искажает глобальные средние $P_{k,g}$ и, соответственно, логарифмические коэффициенты.

• **Половая специфика подписок.** Часть пабликов (кластеры 14–17) по определению ориентирована на женскую аудиторию; их «избегание» мужчинами не информативно для академических выводов, но увеличивает дисперсию в общей выборке.

На рис. 6 представлена тепловая карта $\log_odds$, которая показывает выраженные различия интересов студентов-мужчин в зависимости от академической траектории. Под значениями $\log_odds$ указана процентная доля подписок на этот кластер среди студентов данной траектории.

Наблюдения:

• Траектории 0–2 (низкая и падающая успеваемость). Сильное смещение в сторону развлекательных кластеров 0–6 ($\log_odds \approx 0,18$ – $0,51$) и умеренное в сторону уличной моды и рэп-комьюнити. Также «избегание» образовательного кластера 10 и вузовских сообществ 11–12 (до $-0,41$).

• Траектория 7 (стабильные отличники). Картина зеркальная: отрицательные коэффициенты на развлекательный контент и заметный скачок интереса к кластерам 10–12 (образование, вузовские сообщества).

Далее на рис. 7 приведена аналогичная карта для женщин.

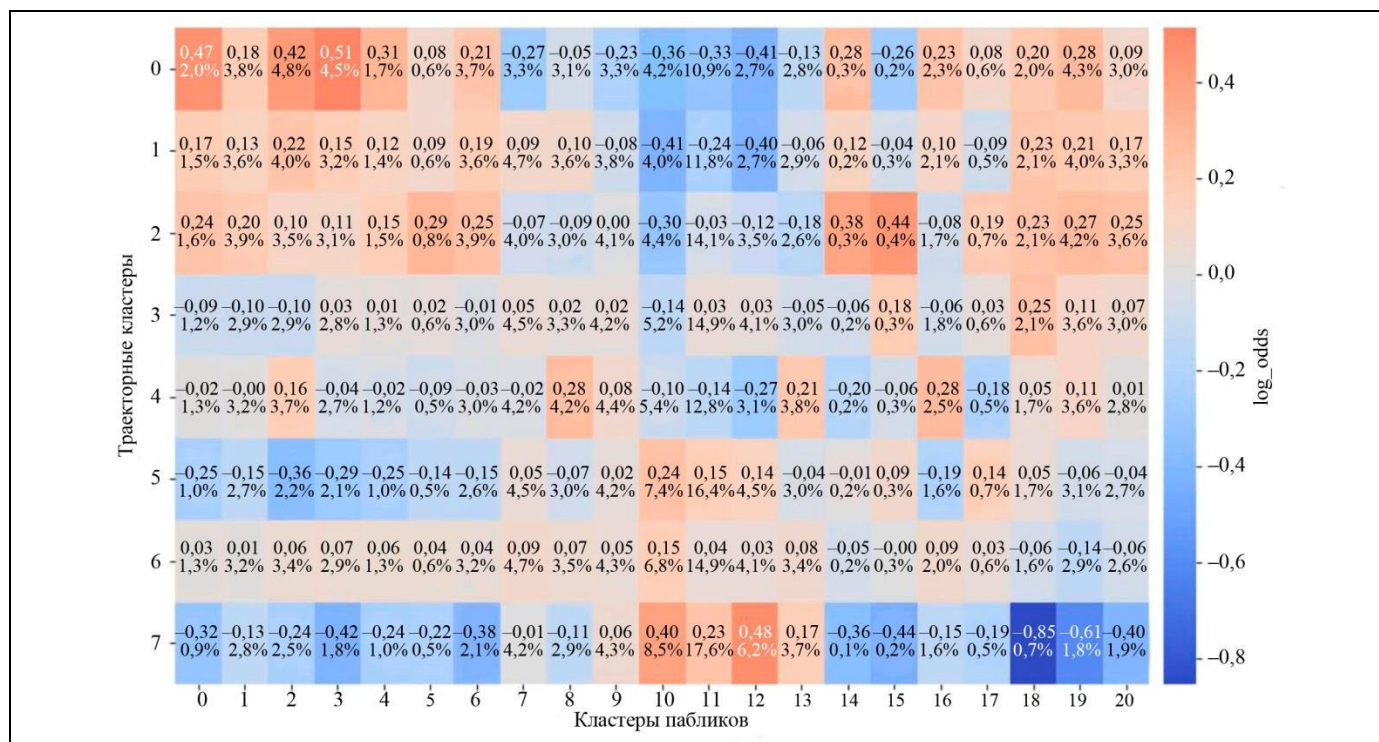


Рис. 6. Тепловая карта по log_odds для мужской выборки студентов

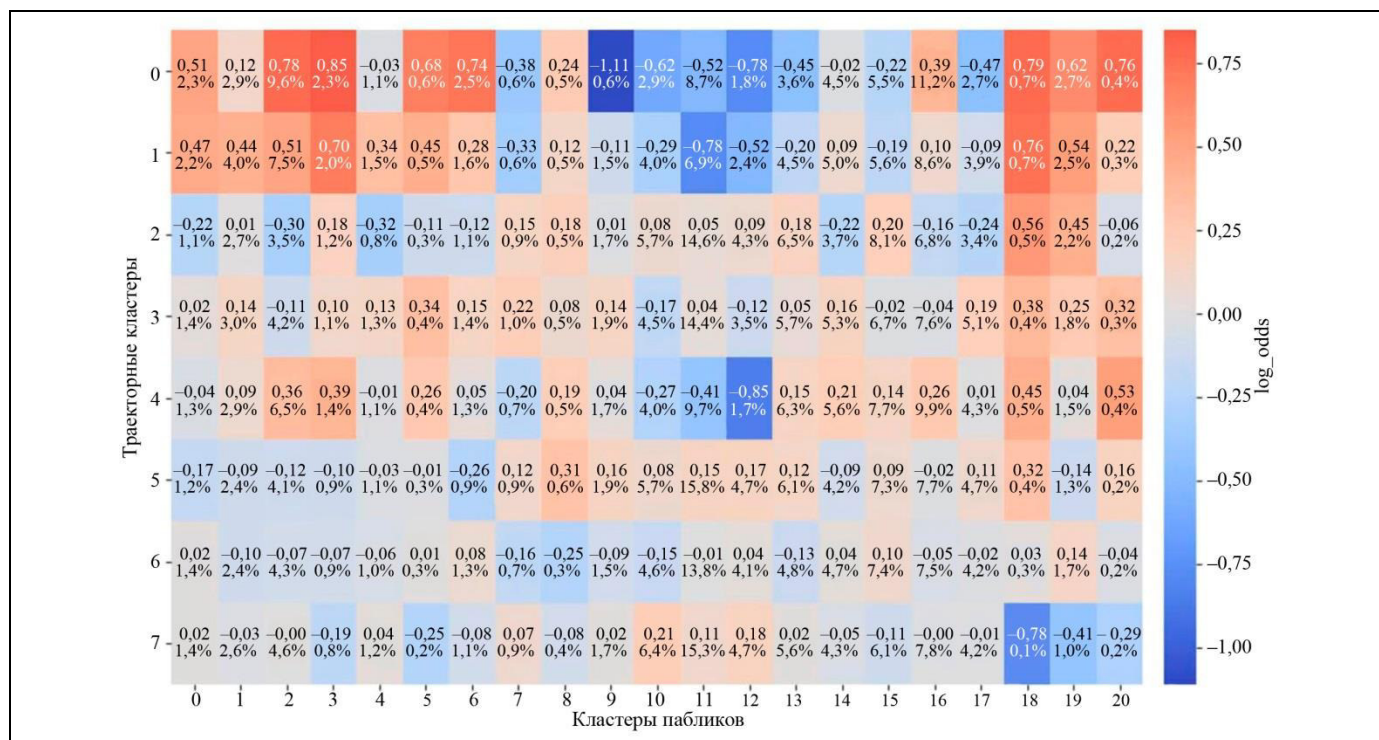


Рис. 7. Тепловая карта по log_odds для женской выборки студентов

По студенткам из слабых траекторий можно сделать аналогичные выводы: высокий интерес к развлекательным кластерам и избегание образовательных. Однако для лучших траекторий (6 и 7) наблюдается более равномерная картина распре-

деления интересов. Несмотря на общую очевидность выводов (развлечения преобладают у слабых студентов, образование – у сильных), внутри каждой траектории наблюдается индивидуальный характер предпочтений. Эти различия по цифровым



интересам могут быть использованы при построении прогностической модели академической успеваемости.

4.2. Проверка гипотезы о независимости академических траекторий и цифровых интересов

Для формальной проверки гипотезы о статистической зависимости между академической успеваемостью студентов и тематикой их цифровых интересов был проведён анализ таблиц сопряжённости с использованием критерия χ^2 Пирсона отдельно для мужской и женской выборок

(табл. 3). Глобальный χ^2 -тест подтвердил наличие статистически значимой связи между принадлежностью к конкретной траектории успеваемости и профилем подписок студентов как среди мужчин, так и среди женщин.

Более детально различия проиллюстрированы на рис. 8. На рис. 8, а, б представлены частные χ^2 -статистики по каждому из 21 кластера интересов: красный цвет обозначает значимые с точки зрения χ^2 различия ($\alpha = 0,05$), серый – незначимые.

Видно, что связи с академическими траекториями стали значимыми в 10 кластерах для мужчин и 15 – для женщин. На рис. 8, в, г изображены оцен-

Таблица 3

Результаты глобального χ^2 -теста независимости академических траекторий и цифровых интересов

Пол	Выборка, чел.	χ^2 -статистика	Число степеней свободы	p-значение	Критическое значение χ^2 ($\alpha = 0,05$)
Мужчины	~ 4500	557,6	140	1e-51	168,613
Женщины	~ 1900	755,8		1e-85	

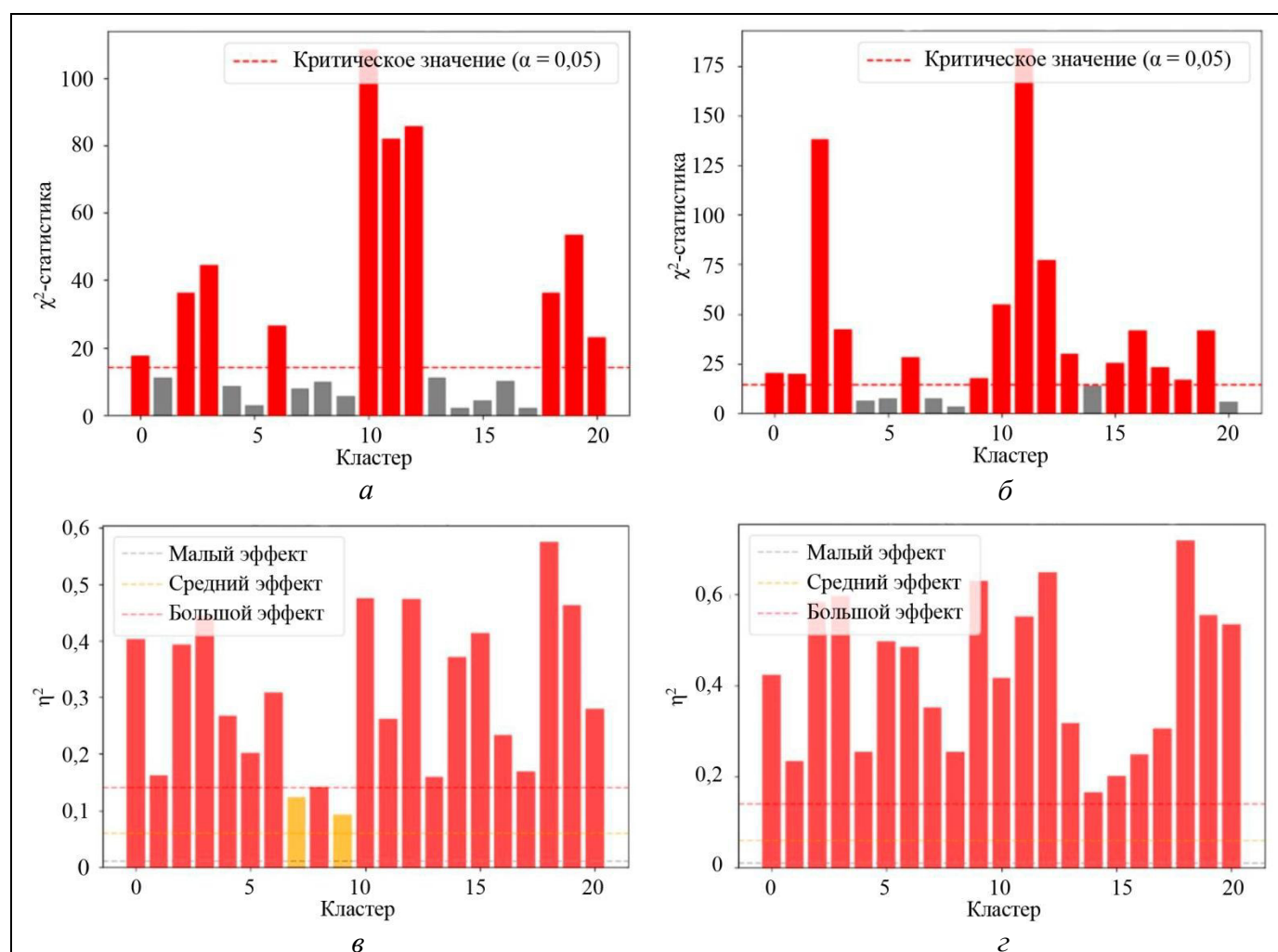


Рис. 8. Частные показатели связи между кластерами цифровых интересов и академическими траекториями: а, б – χ^2 -статистики (красный – $p < 0,05$; пунктир – критическое значение χ^2); в, г – оценки размеров эффекта η^2 (пунктиры – пороги); а, в – мужчины, б, г – женщины

ки размера эффекта η^2 , где пунктирные линии указывают пороги 0,01 (малый эффект), 0,06 (средний) и 0,14 (большой), принятые в работе [18]. Большинство значимых кластеров демонстрирует $\eta^2 > 0,3$, а особенно у женщин около половины кластеров превышают уровень 0,5, что говорит о высокой практической значимости выявленных различий. Наибольшая сила связи с траекторией успеваемости прослеживается в кластерах, связанных с образовательными и студенческими сообществами (номера 10–12), а также с развлекательными и модно-музыкальными пабликами (номера 18–19), что логически коррелирует с ранее наблюдавшимися паттернами подписок. Полученные данные однозначно отвергают нулевую гипотезу о независимости академической траектории и цифровых интересов, подтверждая, что цифровые предпочтения студентов репрезентативны и содержательно связаны с показателями их успеваемости.

ЗАКЛЮЧЕНИЕ

В ходе проведённого исследования зафиксирована статистическая взаимосвязь между тематикой подписок студентов МГТУ им. Н. Э. Баумана в социальной сети ВКонтакте и их академической успеваемостью. Полученные результаты подтверждают, что цифровой след, формируемый через интересы и сетевую активность, может выступать значимым предиктором образовательных траекторий и, возможно, других офлайн-показателей. Формальная проверка гипотезы с использованием критерия хи-квадрат Пирсона показала высокую статистическую значимость ($p < 0,001$) выявленных зависимостей для обеих гендерных выборок, а анализ размера эффекта подтвердил практическую значимость обнаруженных различий.

Исследования других авторов также свидетельствуют о зависимости между активностью в социальных сетях и академической успеваемостью, однако большинство из них ограничивается небольшими выборками и оценкой времени активности в соцсетях.

В перспективе планируется расширение анализа путём включения сетевой структуры дружеских связей студентов, что позволит построить интегрированную модель прогнозирования успеваемости на основе как индивидуальных цифровых интересов, так и их сетево-структурного положения в студенческой сети (метрики центральности, принадлежность к сообществам, брокерские роли). Такой подход откроет возможности для более точной диагностики академических рисков и форми-

рования персонализированных рекомендаций по академической поддержке – от подбора курсов и консультаций до включения в релевантные учебные сообщества.

ЛИТЕРАТУРА

1. *Смирнов И.Б.* Дифференциация учащихся по академической успеваемости в социальной сети: дис.... канд. пед. наук. – М.: НИУ ВШЭ, 2018. – 210 с. [*Smirnov, I.B.* Differentiation of Students by Academic Performance in a Social Network: Cand. Ped. Sci. Diss. – Moscow: NRU HSE, 2018. – 210 p. (In Russian)]
2. *Nikitkov, A., Sainy, B.* The Role of Social Media in Influencing Career Success // *International Journal of Accounting & Information Management*. – 2014. – Vol. 22, no. 4. – P. 273–294. – DOI 10.1108/IJAIM-02-2014-0009
3. *Dokuka, S., Valeeva, D., Yudkevich, M.* How Academic Achievement Spreads: The Role of Distinct Social Networks in Academic Performance Diffusion // *PLOS ONE*. – 2020. – Vol. 15, no. 7. – Art. no. e0236737. – DOI: 10.1371/journal.pone.0236737
4. *Tafesse, W.* Social Networking Sites Use and College Students' Academic Performance: Testing for an Inverted U-shaped Relationship Using Automated Mobile App Usage Data // *International Journal of Educational Technology in Higher Education*. – 2022. – Vol. 19. – Art. no. 16. – DOI: 10.1186/s41239-022-00322-0
5. *Al Mosharrafa, R., Akther, T., Siddique, F.K.* Impact of Social Media Usage on Academic Performance of University Students: Mediating Role of Mental Health under a Cross-sectional Study in Bangladesh // *Health Science Reports*. – 2024. – Vol. 7, no. 1. – Art. no. e1788. – DOI 10.1002/hsr2.1788
6. *Chkhartishvili, A., Gubanov, D.A., Melnichuk, V.S., Sych, V.V.* Exploratory Data Analysis and Natural Language Processing Model for Analysis and Identification of the Dynamics of COVID-19 Vaccine Opinions on Small Datasets // *Advances in Systems Science and Applications*. – 2023. – Vol. 23, no. 3. – P. 108–126. – DOI: 10.25728/assa.2023.23.3.1381
7. *Levenshtein, V.I.* Binary Codes Capable of Correcting Deletions, Insertions and Reversals // *Soviet Physics Doklady*. – 1966. – Vol. 10, no. 8. – P. 707–710.
8. *Каптур В.В., Губанов А.В., Фещенко А.В. и др.* Взаимосвязь образовательных достижений старшеклассников и их цифрового следа в социальной сети // Педагогика и просвещение. – 2020. – № 4. – С. 37–51. [*Kashpur, V.V., Gubanov, A.V., Feshchenko, A.V., et al.* The Relationship between Educational Achievements of High-School Students and Their Digital Footprint in a Social Network // *Pedagogy and Education*. – 2020. – No. 4. – P. 37–51. (In Russian)]
9. *Smirnov, I.B.* Estimating Educational Outcomes from Students' Short Texts on Social Media // *EPJ Data Science*. – 2020. – Vol. 9. – Art. no. 27. – DOI: 10.1140/epjds/s13688-020-00238-1
10. *Дудина В.А., Анисимовская А.И., Горченко А.Л., Данчышак Т.В.* Оценка влияния социальных сетей и других факторов на успеваемость студентов биологического факультета // Молодой учёный. – 2019. – № 48 (286). – С. 438–440 [*Dudina, V.A., Anisimovskaya, A.I., et al.* Assessment of the Influence of Social Networks and Other Factors on Academic



- Performance of Biology Students // Young Scientist. – 2019. – No. 48 (286). – P. 438–440. (In Russian)]
11. *Giunchiglia, F., Zeni, M., Gobbi, E., et al.* Mobile Social Media Usage and Academic Performance // Computers in Human Behavior. – 2018. – Vol. 82. – P. 177–185. – DOI: 10.1016/j.chb.2017.12.041
 12. *Rousseeuw, P.J.* Silhouettes: A Graphical Aid to the Interpretation and Validation of Cluster Analysis // Journal of Computational and Applied Mathematics. – 1987. – Vol. 20. – P. 53–65. – DOI: 10.1016/0377-0427(87)90125-7
 13. *Tibshirani, R., Walther, G., Hastie, T.* Estimating the Number of Clusters via the Gap Statistic // Journal of the Royal Statistical Society: Series B (Statistical Methodology). – 2001. – Vol. 63, no. 2. – P. 411–423. – DOI: 10.1111/1467-9868.00293
 14. *Gorshkov, S., Ignatov, D.I., Chernysheva, A.Yu., et al.* Identifying Top-Performing Students via VKontakte Social Media Communities Using Advanced NLP Techniques // IEEE Access. – 2025. – Vol. 13. – P. 962–979. – DOI: 10.1109/ACCESS.2024.3521857
 15. *Jaccard, P.* Étude Comparative de la distribution florale dans une portion des Alpes et du Jura // Bulletin de la Société Vaudoise des Sciences Naturelles. – 1901. – Vol. 37. – P. 547–579. – DOI: 10.5169/seals-266450
 16. *Blondel, V.D., Guillaume, J.-L., Lambiotte, R., Lefebvre, E.* Fast Unfolding of Communities in Large Networks // Journal of Statistical Mechanics: Theory and Experiment. – 2008. – No. 10. – Art. no. P10008. – DOI: 10.1088/1742-5468/2008/10/P10008
 17. *Barnard, G.A.* Statistical Inference // Journal of the Royal Statistical Society: Series B (Methodological). – 1949. – Vol. 11, no. 2. – P. 115–139. – DOI: 10.1111/j.2517-6161.1949.tb00028.x
 18. *Cohen, J.* Statistical Power Analysis for the Behavioral Sciences: 2nd ed. – Hillsdale, NJ: Lawrence Erlbaum Associates, 1988. – 567 p.
- Статья представлена к публикации членом редколлегии академиком РАН Д. А. Новиковым.*
- Поступила в редакцию 04.05.2025,
после доработки 27.07.2025.
Принята к публикации 28.07.2025.*
- Сыч Владислав Витальевич** – студент, МГТУ им. Н. Э. Баумана; инженер, Институт проблем управления им. В. А. Трапезникова РАН, г. Москва,
✉ sycheslavovich@gmail.com
ORCID iD: <https://orcid.org/0009-0001-6787-8948>
- © 2025 г. Сыч В. В.



Эта статья доступна по лицензии Creative Commons «Attribution» («Атрибуция») 4.0 Всемирная.

THE CORRELATION BETWEEN THE SOCIAL NETWORK ACTIVITY AND ACADEMIC PERFORMANCE OF TECHNICAL UNIVERSITY STUDENTS: A CASE STUDY OF VKontakte

V. V. Sych

Bauman Moscow State Technical University, Moscow, Russia
Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia

✉ sycheslavovich@gmail.com

Abstract. The academic performance of students at Bauman Moscow State Technical University and its correlation to their activity in the VKontakte social network are considered. Machine learning methods are used to identify distinct performance paths reflecting the dynamics of educational achievement. The subscription lists of students are analyzed to identify marker communities characterizing the predominance of specific performance categories. Graph-theoretic clustering is applied to reveal structural groups of student interests. For each path, the stochastic vectors of interest shares across the community clusters are constructed and then used to identify the clusters having a statistically significant relation to particular performance paths. The results indicate a correlation between the digital behavior and academic outcomes of students, which contributes to the development of performance prediction models considering student interests in a social network.

Keywords: academic performance, VKontakte, machine learning, graph theory, performance prediction.

ПРИМЕНЕНИЕ НЕДООПРЕДЕЛЕННЫХ МОДЕЛЕЙ В ЗАДАЧЕ ГЛОБАЛЬНОЙ ЛОКАЛИЗАЦИИ МОБИЛЬНОГО РОБОТА[#]

А. Д. Московский

Национальный исследовательский центр «Курчатовский институт», г. Москва

✉ moscowskyad@yandex.ru

Аннотация. Рассматривается применение недоопределенных моделей для решения задачи локализации мобильного робота. Предлагается комплексная технология, использующая семантические карты и карты в виде облаков точек. Комплексная технология призвана уменьшить время и увеличить точность глобальной локализации на больших, а также симметричных и периодических пространствах. Вместо локализации по данным дальномеров, генерирующих облака точек, предлагается сначала определить соответствия между наблюдаемыми роботом объектами и семантической картой (решить задачу распознавания сцены), а далее с использованием недоопределенных вычислений решить задачу локализации по визуальным ориентирам. С помощью недоопределенных вычислений определяются интервальные ограничения на положения робота, представленные несколькими наборами для каждой гипотезы, полученной при решении задачи распознавания сцен. В полученных интервальных ограничениях предлагается выполнить поиск положения робота с использованием данных дальномеров на основе фильтра частиц, который инициализирован в данных ограничениях. Эксперименты, проведенные на открытом наборе данных KITTI-360, показали, что локализация на основе недоопределенных вычислений позволяет уменьшить пространство поиска до 0,2 % от размера исходной карты. Предложенная комплексная технология демонстрирует значительное преимущество в сравнении с подходами, основанными исключительно на облаках точек или исключительно на визуальных ориентирах, особенно в случаях, когда получено несколько гипотез о соответствии наблюдаемых объектов семантической карте.

Ключевые слова: глобальная локализация, недоопределенные модели, программирование в ограничениях, распознавание сцен, семантические карты, мобильный робот.

ВВЕДЕНИЕ

Задача локализации играет одну из ключевых ролей в современной мобильной робототехнике, поскольку подавляющее большинство подходов к управлению движением мобильного робота базируется на знании его текущего положения. В условиях, когда затруднено использование спутниковой навигации, например, в городах или внутри помещений (в комнатной среде), широко применяются методы локализации, строящие трехмерную карту в виде облака точек. Популярность дан

ного подхода определяется в том числе и его универсальностью: не требуется размещение дополнительных меток (устройств) в окружении робота, которые помогали бы в локализации. Подобная подробная карта может применяться и в задаче навигации, поскольку содержит информацию о препятствиях. Она может быть построена с помощью таких распространенных сенсоров, как сканирующие лазерные дальномеры и камеры глубины. Современное развитие методов SLAM (*Simultaneous Localization and Mapping*) [1] позволяет строить достаточно точные карты больших областей, однако при дальнейшем использовании подобных карт возникает задача глобальной локали-

[#] Работа проведена в рамках государственного задания НИЦ «Курчатовский институт».



зации – определения стартового положения робота без знания каких-либо начальных условий. Данная задача решается при инициализации (старте работы) робота и в случаях, если робот «сбился» (положение робота не соответствует ожидаемому) ввиду ошибок в ходе локализации, или, например, был перемещен третьими лицами. Все эти ситуации не позволяют использовать данные о предыдущем положении робота и его смещении для локализации. Поиск положения робота во всем пространстве (обычно ограниченном размером карты) становится сложнее с ростом размера карты, что отмечают авторы обзора [2] подходов, применяемых для решения данной задачи. Помимо увеличения числа требуемых вычислений, проблемой становится то, что упомянутые городские и комнатные среды зачастую содержат области с достаточно схожей геометрией, особенно когда речь идет о типовых постройках. Подобная «периодичность» и «симметричность» окружающей среды приводит к так называемым локальным минимумам на карте, в которых робот может ошибочно глобально локализоваться. И число таких локальных минимумов растет с размером карт.

К решению данной проблемы исследователи подходят с разных сторон. Одно направление использует только облака точек, его задача – построение качественных дескрипторов, связанных с облаками точек, позволяющих выделить и сопоставить геометрические и другие особенности его элементов, в том числе семантические. В целом в рамках данного подхода можно выделить три основные стадии: выделение дескрипторов, сопоставление дескрипторов, уточнение положения. Среди методов выделения дескрипторов можно отметить такие, как FPFH (*Fast Point Feature Histograms*) [3], NDT (*Normal Distribution Transform*) [4], обучаемые подходы NeRF (*Neural Radiance Fields*) [5], Minkloc3d [6] и др. [7]. Для сопоставления дескрипторов применяются разные подходы: распространены методы на основе RANSAC (*RANdom SAmple Consensus*) [8], графовые подходы [9], оптимизационные алгоритмы [10], обучаемые подходы [11]. Метод ICP (*Iterative Closest Point*) [12] и его дальнейшие расширения (см., например, работу [13]), используются для последующего уточнения положения робота. Также используются такие вероятностные подходы, как гистограммный фильтр или фильтр частиц [14], применяемые к облакам точек [15]. Однако, поскольку данные дескрипторы, в первую очередь, базируются на гео-

метрических особенностях взаиморасположения точек, то проблема локальных минимумов ввиду периодичности и симметричности пространств не может быть устранена. Несмотря на то, что построение дескрипторов сужает пространство поиска, они остаются на уровне облаков точек и в какой-то мере сопряжены с перечисленными ранее проблемами, поэтому другие подходы переходят на более «высокий» уровень представления карты, о котором пойдет речь далее.

Подход Place Recognition [16], подобно построению дескрипторов, кодирует сразу все окружение робота в виде вектора чисел. Обычно для этого применяются нейросетевые модели, которые обучаются так, что одно и то же окружение, снятое с разных ракурсов, в разное время суток и т. п., продуцирует близкие друг к другу векторы. Часто такие модели имеют мультимодальный вход и учитывают как облака точек и «сырые» изображения, так и семантику сцены. Таким образом, карта снабжена набором «ключевых кадров», для которых было выполнено построение кодирующих их векторов, а локализация сводится к поиску кадра, наиболее совпадающего с текущим наблюдением робота, полученным по набору сенсоров. Такой подход хорош тем, что использует всю доступную информацию об окружении, однако требует плотного покрытия пространства ключевыми кадрами. Сильное изменение ракурса также может негативно влиять на качество поиска кадра для различных композиций сенсоров, использующихся на работе. Чем более приближена к полному обзору композиция сенсоров, тем лучше для таких методов, однако не для всех роботов это является возможным и целесообразным.

Еще одно направление, в рамках которого выполнена данная работа, использует так называемые семантические карты. Это карты, где размечены объекты с их семантическими метками (классами). Методы, позволяющие в автоматическом наборе получать такие карты, активно развиваются [17, 18]. Набор объектов тех же классов выделяется в окружении, которое непосредственно наблюдает робот. Далее возникает задача распознавания сцены, а именно сопоставления объектов карты с объектами в окружении (сцене). Поиск такого соответствия на уровне объектов позволяет сузить пространство поиска по сравнению с поиском в трехмерных картах. Семантика придает дополнительную «уникальность» сценам с однообразной геометрией. После того, как сопоставление найде-

но, можно воспользоваться методами локализации по визуальным ориентирам. Однако одной из проблем данного направления является то, что подобные методы распознавания сцен генерируют целый ряд гипотез о сопоставлении объектов. В результате требуется проводить локализацию для каждой из гипотез и оценивать получившееся качество сопоставления объектов. Направление, использующее семантические карты, имеет ряд преимуществ – применяемые методы менее чувствительны к изменениям ракурса и позволяют вручную редактировать семантические карты. Редактирование карт вручную является достоинством метода, поскольку окружающее пространство способно меняться во времени, а проводить полностью процедуру картирования может быть нецелесообразно.

Учитывая вышеизложенное, целью исследования является ускорение работы и увеличение точности глобальной локализации мобильного робота на картах, представленных в виде облаков точек. Выбрано направление, которое, в дополнение к облакам точек, использует семантические карты и требует решения задачи распознавания сцен. Данная работа сконцентрирована на проблеме локализации мобильного робота по результатам применения методов распознавания сцен с учетом такой специфики задачи, как высокие погрешности определения положения объектов, ошибки в сопоставлении объектов и выбор лучшего результата среди набора гипотез.

1. МЕТОДЫ ЛОКАЛИЗАЦИИ ДЛЯ РЕШЕНИЯ ЗАДАЧИ РАСПОЗНАВАНИЯ СЦЕН

В современных работах по локализации с использованием семантических карт, посвященных решению задачи распознавания сцен, одним из самых распространенных методов является оптимизационный подход сопоставления двух наборов трехмерных точек – SVD (*Singular Value Decomposition*) [19]. Локализация этим методом по результатам распознавания сцен проводится в публикациях [20–22]. Данный подход минимизирует среднеквадратичную ошибку положений точек между двумя наборами, однако требует точного поэлементного соответствия одного набора другому и не учитывает погрешности определения положения в расчетах. В работе [23] используется упомянутый подход ICP, примененный не к облакам точек, а к двум наборам объектов: на карте и на сцене. В ря-

де работ (см., к примеру, публикации [24, 25]) для локализации робота используется технология 2D-3D перепроецирования (*reprojection*), позволяющая сопоставить положения плоских объектов на изображении с их трехмерными положениями. Этот подход позволяет обойтись без трехмерной локализации объектов сцены в пространстве, а следовательно, отчасти нивелировать погрешности в определении этого расстояния, которые могут быть значительными.

Перспективным направлением видится сочетание локализации по облакам точек с данными, полученными в результате применения методов распознавания сцен. Интересный подход был предложен в работе [26]: производилась отдельная трехмерная регистрация облаков точек выделенных объектов, дальнейшее уточнение проводилось с помощью ICP.

Вместе с тем не было обнаружено работ, которые использовали бы поисковые вероятностные подходы на основе гистограммного фильтра или фильтра частиц в задаче локализации по визуальным ориентирам в исследуемой постановке. Это объясняется тем, что для этих методов требуется ограничить область поиска, что проблематично для большинства рассмотренных подходов к локализации, так как они выдают в качестве ответа точку в пространстве, а не область. Использование ограничений на положение робота в виде интервалов позволило бы решить данную проблему. Общий подход программирования в ограничениях (англ. *constraint programming*) [27] позволяет получать данные такого рода. Обобщение этого подхода в виде недоопределенных вычислений (Н-вычислений) [28] было применено к задаче локализации по визуальным ориентирам и последующему использованию гистограммного фильтра внутри полученных ограничений [29]. В данной же работе будет рассмотрен этот подход в применении к локализации робота по ряду гипотез с использованием фильтра частиц (ФЧ), определяющего положение робота с помощью облаков точек в полученных ограничениях.

2. ПРЕДЛАГАЕМАЯ КОМПЛЕКСНАЯ ТЕХНОЛОГИЯ ГЛОБАЛЬНОЙ ЛОКАЛИЗАЦИИ

Комплексная технология глобальной локализации, предлагаемая в настоящей работе, выполнена в рамках направления, использующего семантические карты. Такая комплексная технология лока-



лизации требует использования как слоя карты в виде облака точек, так и семантического слоя, на котором размечены положения объектов и указана некоторая семантическая информация о них. Помимо двух слоев карт, входными данными являются изображения и облака точек от сенсоров робота. Решение задачи локализации в рамках предлагаемой комплексной технологии состоит из следующих этапов:

1. Распознавание и локализация объектов в окружении робота (формирование сцены).
2. Решение задачи распознавания сцены – поиск соответствия между распознанными объектами и семантической картой, получение ряда гипотез.
3. Применение подхода недоопределенной локализации (Н-локализации) к полученным гипотезам: отброс противоречивых гипотез и определение интервальных ограничений на положение робота.
4. Определение положения робота с помощью ФЧ в полученных ограничениях для каждой гипотезы, расчет численных оценок качества.
5. Выбор решения с наибольшей численной оценкой.

Первый этап осуществляется с применением либо классического способа распознавания объектов на изображениях и дальнейшей их локализации с использованием карт глубины или облаков точек [30], либо с применением подходов трехмерного распознавания [31].

Второй этап, включающий в себя решение задачи распознавания сцен, подробно изложен в работе [32]. В ней предлагается использование алгоритмов теории графов, позволяющих выделить геометрические особенности взаиморасположения групп объектов. В работе было описано использование базисных визуально-языковых моделей на примере CLIP (*Contrastive Language-Image Pre-training*) [33] для учета визуального сходства объектов, а не только семантической метки, как это делалось ранее большинством исследователей в области локализации робота по семантическим картам. Использование критерия визуального сходства объектов в дополнение к критерию геометрических особенностей позволило значительно увеличить точность распознавания сцен в сравнении с методами, опирающимися только на один из критериев. Однако ряд проблем, связанных с работой с объектами, включая высокие погрешности их локализации, наличие визуально схожих объектов, ошибки систем распознавания, а также множественность решений приводит к тому, что на выходе таких систем генерируется множество гипо-

тез и верное решение не всегда обладает наивысшей численной оценкой.

На третьем этапе применяется подход Н-локализации, которому посвящена данная работа. Данный подход используется для решения задачи локализации робота по визуальным ориентирам, он формирует интервальные ограничения на положения робота и позволяет определить противоречия во входных данных, которые могут возникнуть из-за перечисленных особенностей результатов решения задачи распознавания сцен.

На четвертом этапе применяется ФЧ для определения положения робота с использованием слоя карты в виде облака точек. Применение ФЧ обосновано тем, что он естественным образом может быть инициализирован в интервальных ограничениях, полученных от Н-локализации. Кроме того, ФЧ способен работать с широким набором входных данных, для которых можно получить численную оценку качества любой из частиц, которая, в свою очередь, в задаче локализации представляет собой гипотезу о положении робота.

На пятом этапе с использованием полученной численной оценки определяется лучшее решение для всех полученных на втором этапе гипотез.

Схема комплексной технологии глобальной локализации приведена на рис. 1.

2.1. Локализация по ориентирам на основе недоопределенных моделей

Для применения Н-вычислений требуется создать описание решаемой задачи в виде недоопределенной модели (Н-модели). Такая модель представляет собой описание искомых величин, их области определения и функций интерпретации и присваивания, по которым искомые величины обновляются [28]. Искомые величины, а именно положение робота, задаются в виде так называемых мультиинтервалов – упорядоченному по возрастанию непересекающемуся набору интервалов:

$$^*a = \left[\left[a_{\text{low}}^1, a_{\text{high}}^1 \right], \dots, \left[a_{\text{low}}^N, a_{\text{high}}^N \right], \right.$$

$$\left. a_{\text{low}}^i \text{high}^i \in R, a_{\text{low}}^i \leq a_{\text{high}}^i, a_{\text{high}}^i < a_{\text{low}}^{i+1} \right],$$

где *a – некая недоопределенная переменная (Н-переменная), заданная мультиинтервалом; a_{low}^i – нижние границы интервалов; a_{high}^i – верхние границы интервалов, образующих мультиинтервал; $i = 1, N - 1$. Мультиинтервалы являются частным случаем Н-переменных, которые здесь и далее будут обозначаться символом «*» слева от имени переменной.

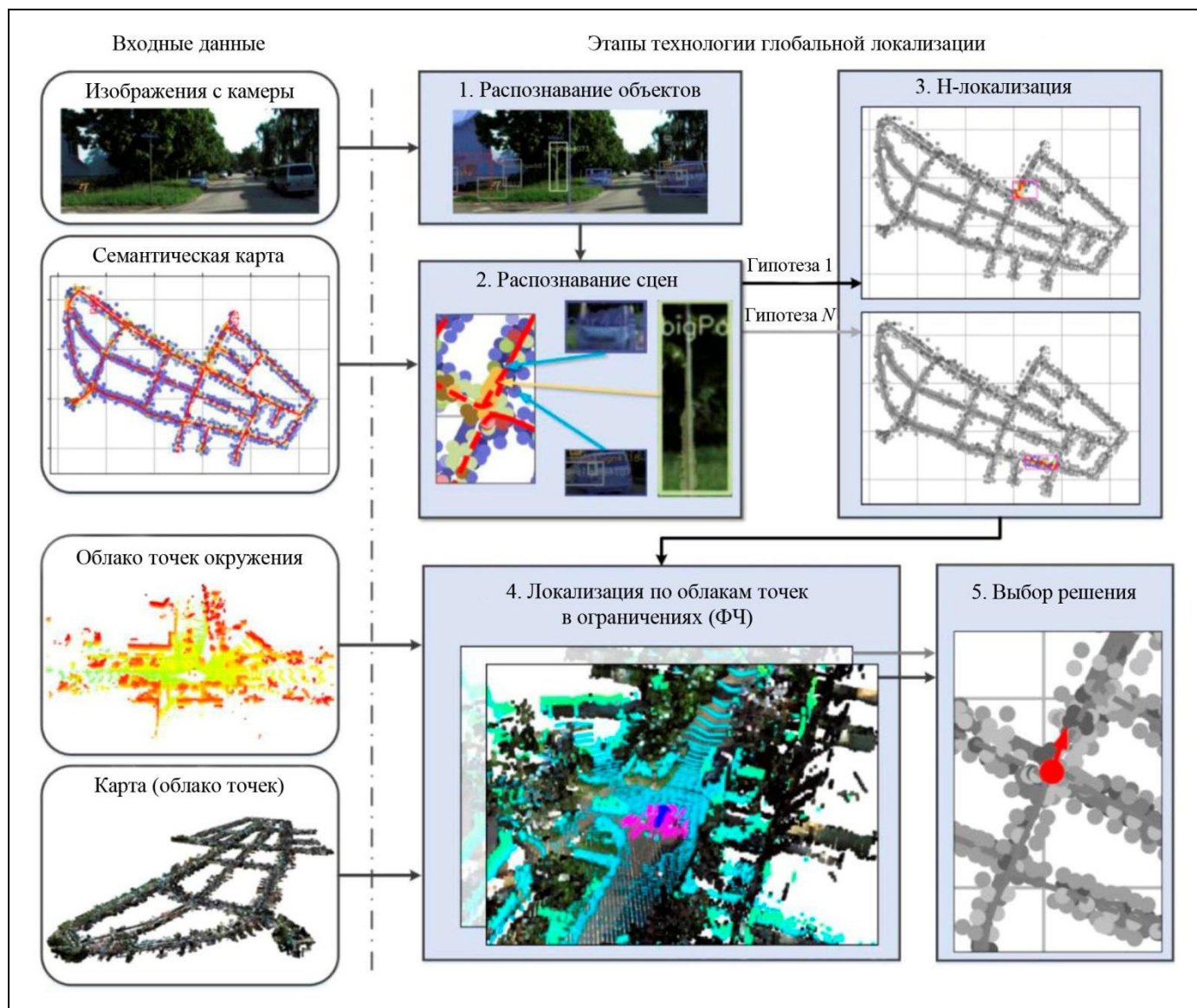


Рис. 1. Схема комплексной технологии решения задачи глобальной локализации

Изначально эти мультиинтервалы инициализированы всеми доступными значениями. Далее определяются так называемые функции интерпретации, применение которых должно уменьшать неопределенность, т. е. сужать границы мультиинтервалов, поскольку новое вычисленное значение на каждом шаге итерации есть пересечение старого значения со значением, полученным в ходе применения функции интерпретации. Если пересечения нет, значит, входные данные содержат противоречия и должны быть исключены из рассмотрения. Функции интерпретации определяются на основе связи между положением робота ($*X$, $*Y$, $*\theta$), положением наблюдаемых ориентиров (X_l , Y_l), и их измерениями, включающими в себя расстояние до объекта r_l и угол на него α_l :

$$*X = \pm \sqrt{(r + \Delta r_l)^2 - (Y - Y_l)^2} + X_l, \quad (1)$$

$$*Y = \pm \sqrt{(r + \Delta r_l)^2 - (X - X_l)^2} + Y_l, \quad (2)$$

$$*X = \frac{*Y - Y_l}{\tan(*\theta + *\Delta\alpha_l + \alpha_l)} + X_l, \quad (3)$$

$$*Y = (*X - X_l) \tan(*\theta + *\Delta\alpha_l + \alpha_l) + Y_l, \quad (4)$$

$$*\theta = \arctan\left(\frac{Y_l - *Y}{X_l - *X}\right) - (\alpha_l + *\Delta\alpha_l), \quad (5)$$

где $*\Delta r_l$ и $*\Delta\alpha_l$ – погрешности измерений r_l и α_l , выраженные интервалами, которые могут быть получены из среднеквадратичных ошибок по пра-

вилу N сигм: $*\Delta r_l = [-N\sigma r_l, N\sigma r_l]$. В формулах (1)–(5) часть переменных представляется в виде мультиинтервалов и для их вычисления требуется применение интервальной арифметики [34]. Для каждого наблюдаемого ориентира, полученного в ходе решения задачи распознавания сцен, формируется свой набор функций интерпретации (1)–(5). Процедура Н-вычислений сводится к итеративному применению функций к набору Н-переменных. После применения функции она удаляется из списка активных функций, однако если была обновлена переменная (изменилось ее значение), которая является входной для какой-либо функции присваивания вне активного списка, то эта функция возвращается в список активных функций. Вычислительная процедура продолжается до тех пор, пока список активных функций не пуст; также на практике можно ограничить число итераций или установить желаемую точность полученных значений переменных.

В то же время имеется ряд ситуаций, которые не могут быть разрешены методами распознавания сцен. Сюда относятся ситуации симметрии, поскольку алгоритмы на графах неспособны различать порядок появления объектов по мере увеличения угла на них, так как граф содержит информацию только о расстоянии между двумя объектами. Эта ситуация может быть решена на уровне Н-моделей введением дополнительных функций проверки корректности. На основе текущих ограничений ($*X$, $*Y$, $*\theta$) рассчитываются значения углов на самый левый и самый правый объект в сцене и выполняется проверка неотрицательности их разности:

$$-\left(\arctan\left(\frac{*Y - Y_{\text{left}}}{*X - X_{\text{left}}} \right) - \arctan\left(\frac{*Y - Y_{\text{right}}}{*X - X_{\text{right}}} \right) < 0 \right), (6)$$

где индекс left соответствует самому левому (по углу) ориентиру, а индекс right – самому правому. Применять дополнительные функции проверки корректности лучше не к полному набору ($*X$, $*Y$, $*\theta$), а к их декартовому произведению, где каждый элемент – это кортеж из трех интервалов для каждой переменной. Те кортежи, что не прошли проверку (6), исключаются из общего множества.

На выходе получаются значения Н-переменных, представленные набором кортежей интервалов. Такие кортежи описывают объединение пря-

моугольных областей (со сторонами, параллельными осям x и y , иногда именуемых брусами [34]) с указанием ограничения на размер угла. Однако полученный в такой форме ответ может быть несколько избыточным и содержать лишние значения (рис. 2, а).

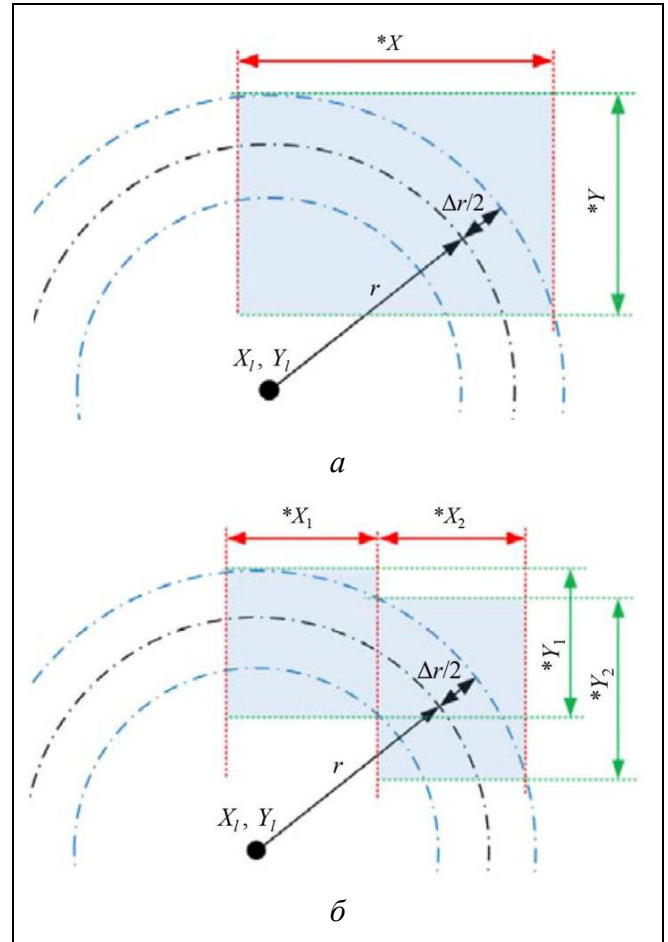


Рис. 2. Графическое представление избыточного оценивания: а – ситуация избыточного оценивания области положения робота, ограниченного синими штрихпунктирными линиями, где $*X$ – входная Н-переменная, а $*Y$ – рассчитанная; б – пример исчерпывающего оценивания области путем деления по входной переменной

Для таких ситуаций в интервальном анализе [34] применяются алгоритмы исчерпывающего оценивания, заключающегося в итеративном разбиении интервалов, превосходящих заданное ограничение и повторном применении вычислительной процедуры к получившемуся результату. На рис. 2, б показан пример такого разделения после первой итерации алгоритма по переменной $*X$ и вычисление новых значений $*Y_1$ и $*Y_2$. Таким образом, общий алгоритм глобальной локализации на основе Н-вычислений (алгоритм Н-локализации) представим в виде псевдокода.

Алгоритм глобальной локализации по визуальным ориентирам с использованием Н-вычислений

1.	Algorithm <code>sdm_global_localization</code> ($\{ \langle X_1, Y_1, r_1, \alpha_1, \sigma_r, \sigma_\alpha \rangle \}, N\sigma, xy_lim$):
2.	-- X_1, Y_1 – положения объектов на карте
3.	-- r_1, α_1 – измерения объектов на сцене
4.	-- σ_r, σ_α – погрешности измерений объектов
5.	-- $N\sigma$ – множитель для перевода погрешностей в интервал
6.	-- xy_lim – максимально допустимый размер интервалов по x и y
7.	$V = \emptyset$
8.	$*X = *Y = [[-\infty, \infty]]$
9.	$*\theta = [[-\pi, \pi]]$
10.	$R = \text{init_R}(\{ \langle X_1, Y_1, r_1, \alpha_1, \sigma_r, \sigma_\alpha \rangle \}, N\sigma)$ -- (1)–(5)
11.	$Status, *X, *Y, *\theta = \text{sdm_process}(*X, *Y, *\theta, R)$
12.	if ! $Status$
13.	return V
14.	$C = \text{init_C}(\{ \langle X_1, Y_1, r_1, \alpha_1, \sigma_r, \sigma_\alpha \rangle \}, N\sigma)$ -- (6)
15.	function <code>process_leaf</code> ($*x, *y, *y$)
16.	$Status, *x, *y, *y = \text{sdm_process}(*x, *y, *y, R)$
17.	if $Status$
18.	if $ *x > xy_lim$:
19.	$*x1, *x2 = \text{divide}(*x)$
20.	<code>process_leaf</code> ($*x1, *y, *y$)
21.	<code>process_leaf</code> ($*x2, *y, *y$)
22.	elif $ *y > xy_lim$:
23.	$*y1, *y2 = \text{divide}(*y)$
24.	<code>process_leaf</code> ($*x, *y1, *y$)
25.	<code>process_leaf</code> ($*x, *y2, *y$)
26.	else
27.	add ($*x, *y, *y$) to V
28.	for ($*x, *y, *y$) in $*X \times *Y \times *\theta$ do
29.	if $C(*x, *y, *y)$ then
30.	<code>process_leaf</code> ($*x, *y, *y$)
31.	return V

Алгоритм `sdm_global_localization` принимает на вход набор положений ориентиров и их измерений, множитель $N\sigma$ для перевода среднеквадратичных ошибок в интервальные, лимит на размер области xy_lim . Вспомогательная функция `init_R` инициализирует параметры функции присваивания (1)–(5) на основе данных об ориентирах. Функция `sdm_process` производит вычисление Н-переменных описанным выше образом, согласно изложенному в работе [28], и возвращает статус (успех либо неуспех) и обновленные значения Н-переменных. Функция `init_C` инициализирует параметры дополнительных функций проверки корректности((6). Функция `divide` разбивает интервал на два путем деления по его центру, и алгоритм возвращает набор кортежей интервалов, не превосходящих по осям x и y заданное ограничение xy_lim и прошедших все дополнительные проверки.

2.2. Локализация с помощью фильтра частиц в ограничениях

Полученные путем применения алгоритма `sdm_global_localization` ограничения описывают область, в которой может находиться робот. Имея такие ограничения, можно выполнить поиск с использованием ФЧ. Каждая частица в данном контексте – это гипотеза о положении робота. Требуется инициализировать начальное распределение в полученных ограничениях равномерным способом. Для работы ФЧ требуется оценивать качество каждой частицы. Предлагается использовать данные плотной карты в виде облака точек и облако точек, полученное с робота. Для сравнения двух облаков точек предлагается рассчитать кратчайшие расстояния от точек облака с робота до точек облака карты. Среднее арифметическое этих рас-



стояний покажет качество наложения облаков друг на друга. Однако для ускорения сходимости ФЧ рекомендуется использовать более «чувствительную» оценку:

$$p(d, \sigma) = \frac{e^{-(d/\sigma)^2}}{\sqrt{2\pi}\sigma}, \quad (7)$$

где p – желаемая оценка частицы; d – среднее кратчайших расстояний; σ – желаемый порог чувствительности. На каждой итерации работы ФЧ происходит расчет весов всех частиц путем переноса облака точек с робота в координаты частицы и расчета оценки (7). Далее происходит процедура перевыборки (*resampling*), цель которой – отсеять частицы с низким весом и размножить частицы с высоким; вероятность, с которой частица попадет в новую выборку, пропорциональна ее весу. Поскольку в процессе перевыборки некоторые части-

цы представлены в нескольких экземплярах, то в начале каждой итерации имеет смысл добавить к каждой частице некоторый случайный шум, чтобы разделить одинаковые гипотезы и плотнее покрыть область поиска. Процедуру, состоящую из указанных шагов, можно проводить либо фиксированное число итераций, либо до достижения целевых значений каких-либо критериев, например, среднего веса частиц, значений матрицы ковариации, построенной по распределению частиц, и т. п. Для выбора искомого положения существуют различные практики: взвешенное среднее всех частиц, максимальная по весу частица, первичная кластеризация частиц и применение указанных способов к локальным кластерам. После расчёта искомого положения робота его качество можно оценить с использованием той же метрики (7). Алгоритм глобальной локализации с помощью ФЧ выглядит следующим образом.

Алгоритм глобальной локализации по облакам точек с помощью фильтра частиц в ограничениях

```

1.  Algorithm pf_global_localization( $V, \sigma, N, iter, \Delta, PCD_{map}, PCD_{robot}$ ):
2.      --  $V$  – ограничения на положение робота
3.      --  $\sigma$  – чувствительность
4.      --  $N$  – число частиц
5.      --  $iter$  – число итераций
6.      --  $\Delta$  – параметры шума
7.      --  $PCD_{map}$  – облако точек карты
8.      --  $PCD_{robot}$  – облако точек с робота
9.       $P = \emptyset$ 
10.     for  $n$  in  $N$  do
11.          $v = \text{sample\_element}(V)$ 
12.          $p = \text{sample\_particle}(v)$ 
13.         add  $p$  to  $P$ 
14.     for  $i$  in  $iter$  do
15.          $W = \emptyset$ 
16.          $P' = \emptyset$ 
17.         for  $p$  in  $P$  do
18.              $p = \text{random\_shift}(p, \Delta)$ 
19.             add  $p$  to  $P'$ 
20.              $PCD = \text{transform\_cloud}(PCD_{robot}, p)$ 
21.              $d = \text{get\_mean\_p2p}(PCD, PCD_{map})$ 
22.              $w = \text{get\_w}(d, \sigma) \text{ -- (7)}$ 
23.             add  $w$  to  $W$ 
24.          $P = \text{resampling}(P', W)$ 
25.          $pose_{final} = \text{get\_pose}(P, W)$ 
26.          $PCD_{final} = \text{transform\_cloud}(PCD_{robot}, pose_{final})$ 
27.          $d_{final} = \text{get\_mean\_p2p}(PCD_{final}, PCD_{map})$ 
28.          $w_{final} = \text{get\_w}(d_{final}, \sigma) \text{ -- (7)}$ 
29.     return  $pose_{final}, w_{final}$ 

```


Алгоритм **pf_global_localization** принимает на вход набор ограничений, полученных от алгоритма **sdm_global_localization**, ряд значений параметров, облака точек с робота и карты. Функция **sample_element** выбирает один из кортежей интервалов случайным образом, но пропорционально общему размеру области, полученному путем произведения размеров всех интервалов. Функция **sample_particle** равновероятностно генерирует частицу, включающую в себя положение по осям x и y и углу так, что все значения принадлежат выбранным интервалам. Функция **random_shift** добавляет к частицам случайный белый шум с указанными параметрами. Функция **transform_cloud** производит геометрический перенос облака точек в координаты частицы. Функция **get_mean_p2p** производит расчет кратчайших расстояний между точками первого облака и точками второго облака. Функция **get_w** рассчитывает вес частицы согласно формуле (7). Функция **resampling** производит пере выборку частиц, функция **get_pose** рассчитывает положение робота одним из предложенных ранее способов, например, взятием взвешенного среднего, путем умножения матрицы частиц на вектор весов. Алгоритм возвращает положение робота и вес этого положения как критерий качества.

Таким образом, к каждой выбранной гипотезе о сопоставлении объектов сцены объектам карты последовательно применяются алгоритмы **sdm_global_localization** и **pf_global_localization**, в результате чего определяется набор решений с рассчитанными критериями качества, из которых выбирается лучшее по введенной метрике.

3. ЭКСПЕРИМЕНТАЛЬНОЕ ИССЛЕДОВАНИЕ

Экспериментальное исследование предложенной технологии было проведено на адаптированном открытом наборе данных KITTI-360 [35] для задачи распознавания сцен, предложенной в работе [32]. Подготовленные данные¹ содержат описания семантических карт и сцен в виде наборов объектов с указанием соответствий между этими множествами. В качестве входных данных для алгоритма **sdm_global_localization** были выбраны результаты работы метода распознавания сцен на основе поиска изоморфного подграфа *sig_lite_clip* [32] как показавшего лучшие результаты по метрике нахождения полностью верного ответа. Ответ для задачи распознавания сцен представляет собой соответствие между наблюдаемыми объектами и

объектами на карте, это соответствие используется для формирования задачи локализации по данным об ориентирах. Тем не менее, значительная часть ответов не содержала правильного ответа (когда все объекты сопоставлены корректно) на первом месте в упорядоченном по коэффициенту уверенности (численная оценка результатов распознавания сцен) списке гипотез. Чтобы проверить эффективность указанной технологии в работе с такими входными данными, были выбраны максимальные по коэффициенту уверенности десять ответов для каждой входной сцены.

Выбранные сцены были разбиты на несколько групп: *super_true*, *has_true* и *wrong*. Группа *super_true* соответствует тем результатам распознавания сцен, в которых правильный ответ сопоставления объектов сцены объектам карты присутствует и имеет самый высокий коэффициент уверенности среди выбранных десяти ответов. Группа *has_true* содержит ответы, среди которых есть правильный, но он имеет не самый высокий коэффициент уверенности среди остальных гипотез. Группа *wrong* содержит результаты, в которых нет полностью верных ответов. Размер групп соотносится как 68 % – 27% – 4 % к общему числу сцен, оставшееся число не было обработано методами распознавания. Все эксперименты были выполнены на компьютере с процессором AMD Ryzen7 2700X Eight-Core 3.70GHz и ОЗУ 32ГБ, графические ускорители не использовались.

Для полученных десяти гипотез для каждой сцены была проведена процедура Н-локализации (SDM, *Sub Definite Models*) с $N\sigma = 4$ в варианте с дополнительным разделением интервалов с $xy_lim = 10$ м, и без него (см. рис. 2). В варианте без дополнительного деления переменных оценивался результат, полученный на десятой строчке алгоритма **sdm_global_localization**. Полученные интервальные ограничения на положение робота для всех гипотез одной сцены оценивались на предмет точности интервалов и охвата интервалов (табл. 1). Точность интервалов рассчитывалась путем определения, принадлежит ли истинное положение робота, известное в используемом наборе данных, рассчитанным интервальным ограничениям (чем больше значение, тем лучше). Охват интервалов отражает сокращение области поиска. Он оценивался путем равномерного распределения миллиона частиц (положений робота) по площади всей карты, определения числа частиц, принадлежащих любому из полученных наборов ограничений, и взятию отношения полученного числа ко всему набору (чем меньше его значение, тем лучше).

¹ URL: https://github.com/MoscowskyAnton/scene_recognition_kitti_360



Таблица 1

Результаты тестирования подхода Н-локализации

Метод	Точность интервалов, %, ↑			Охват интервалов, %, ↓			t , с
	Группа			Группа			Все группы
	super_true	has_true	wrong	super_true	has_true	wrong	
SDM с делением	97,1	98	17,6	0,23	0,41	0,65	0,69
SDM без деления	99,8	100	37	3,22	4,44	0,37	0,03

Из табл. 1 видно, что группа *wrong*, в которой нет верного ответа, сильно отстает по точности интервалов. Также видна разница на порядок в охвате интервалов для варианта алгоритма с делением при небольшой потере точности интервалов относительно варианта без деления. Однако уменьшение охвата интервалов на порядок приводит к увеличению времени выполнения алгоритма на тот же порядок.

Для оценки точности определения положения робота предлагаемая технология сравнивалась с другими методами глобальной локализации. Все исследуемые подходы были реализованы автором лично либо взяты из открытых библиотек. Методы, работающие с результатом распознавания сцен, получали для каждой сцены на вход все десять вариантов и по численной оценке выбирали наилучший, для этого лучшего решения рассчитывались ошибки по положению и по углу и время исполнения. Ошибки оценивались при помощи математического ожидания ($|.$) медианы (M). Время работы метода распознавания сцен не учитывалось, однако для выбранного метода и выбранной последовательности составило $41,55 \pm 46,5$ с [32] на том же вычислительном оборудовании. Стоит отметить, что в наборе KITTI-360 облака точек для карты представлены не в монолитном варианте, а в виде нескольких пересекающихся областей (подкарт); там, где это возможно, выбиралась одна из подкарт, наиболее подходящая к области поиска. В исследовании были рассмотрены следующие подходы:

– **RANSAC+ICP** – классический подход к локализации исключительно по облакам точек с расчетом FPFH-дескрипторов [3], поиском их соответствий методом RANSAC [8] и дальнейшим уточнением с помощью ICP [12]. Данный подход был выбран для рассмотрения ввиду доступных открытых реализаций. Поиск соответствий методом RANSAC проводился для каждой подкарты, а уточнение ICP – для той позиции, которой соответствовало наибольшее значение меры *fitness*.

– **SVD** – метод локализации по данным распознавания сцен [19]. Качество решения для разных результатов распознавания сцен определялось средней ошибкой положения каждого объекта между сценой и картой.

– **SVD+ICP** – локализация по данным распознавания сцен методом SVD и дальнейшее уточнение полученного положения с помощью ICP по облакам точек. В качестве численной оценки использовался параметр *fitness* метода ICP.

– **SDM+PF** – предлагаемая в работе технология Н-локализации по ориентирам алгоритмом **sdm_global_localization** ($N\sigma = 4$, $x_{y_lim} = 10$ м) и использование ФЧ в ограничениях алгоритмом **pf_global_localization** ($N = 150$; $\sigma = 0,1$; $iter = 5$; $\Delta = (0,5; 0,1)$). В качестве решения выбиралась частица с наибольшим весом, который также служил критерием отбора решений для разных результатов распознавания сцен.

– **SDM+RANSAC+ICP** – подход **RANSAC+ICP**, который применялся к облаку точек, полученному на основе ограничений, рассчитанных алгоритмом **sdm_global_localization**. Рассчитанные ограничения были расширены на дальность работы дальности. В качестве меры оценки решений также бралось значение *fitness* метода ICP. Использовался вариант метода SDM без деления, остальные параметры соответствовали параметрам в **SDM+PF**.

Дополнительно для рассматриваемых результатов была рассчитана часто встречающаяся в задачах глобальной локализации метрика *recall@k* ($R@k$), показывающая процент попадания ответа в некоторую доверительную область для k ответов. В литературе [36, 37] для задачи глобальной локализации доверительная область выбирается, как правило, равной 20 м, исходя из соображения, что для решения данной задачи с практической точки зрения уже неважно, насколько сильно ошибка превышает это значение.

Численные результаты применения исследуемых методов для группы *super_true* приведены в табл. 2.

Таблица 2

Численные результаты экспериментов для группы решений *super_true*

Метод	$ r $, м	$M(r)$, м	$ \alpha $, рад	$M(\alpha)$, рад	$ t $, с	$M(t)$, с	$R@1$, %	$R@5$, %
RANSAC+ICP	233,3	190,1	0,71	0,12	4418,1	4171,5	37	—
SVD	4,7	2,3	0,26	0,06	0,005	0,004	94,4	95,6
SVD+ICP	39,5	1,6	0,29	0,06	18,6	17,7	94,9	96,1
SDM+PF	83,7	2,1	0,39	0,06	95,1	97,1	98,4	99,3
SDM+RANSAC+ICP	100,6	1,4	0,28	0,05	1009	976	69,4	83,7

По результатам, представленным в табл. 2, можно сделать вывод, что распространенный подход SVD без привязки к облакам точек значительно превосходит остальные методы практически по всем параметрам. Связано это с тем, что коэффициент уверенности, использующийся для нормирования ответов в методах распознавания сцен, идеологически совпадает с метрикой оценки метода SVD, поэтому для группы *super_true* этот метод практически не ошибается с выбором гипотезы, выбирая самую высокую по коэффициенту уверенности, которая и является правильной. Это же видно в сравнении с расширением этого метода SVD + ICP, где значительно выросла средняя ошибка, так как появились сцены, где была предпочтена другая гипотеза, однако медианная ошибка в сравнении с SVD была уменьшена, что говорит о том, что более чем в половине случаев ICP смог улучшить положение по данным с облаков точек. Другие методы, использующие результаты распознавания сцен, показывают сравнимую медиану, но сильно увеличенное математическое ожидание из-за неверно выбранных гипотез, которые могут находиться на других концах карты, тем самым значительно повлияв на среднее значение. Влияние «удаленных» от верного ответа ошибок также косвенно подтверждается и превосходящими значениями $R@k$ для метода SDM + PF, поскольку метрика $R@k$ описывает пороговую оценку.

Результаты для группы *has_true* представляют больший интерес в рамках исследования, поскольку отражают возможность методов находить верный ответ в наборе близких гипотез. Численные результаты для данной группы приведены в табл. 3.

В табл. 3 в сравнении с табл. 2 метод SVD, который показывал лучшие результаты, значительно потерял в точности и демонстрирует одни из худших значений показателей. Это происходит ввиду того, что он не использует дополнительные данные в виде облаков точек, как это делают осталь-

ные методы, что подтверждает важность предлагаемого комбинированного подхода. Однако добавление коррекции с помощью ICP позволило значительно повысить точность и получить практически лучшие результаты по всем параметрам и хорошие временные показатели работы алгоритма. Показатели точности методов, основанных на результатах распознавания сцен и уточнения положения по облакам точек, в целом не сильно отличаются от группы *super_true* (см. табл. 2), что говорит о возможности нахождения решения среди множества близких гипотез. Сильное снижение значения метрики $R@1$ обусловлено неправильностью лучшего ответа в группе *has_true*, снижение значения метрики $R@5$ также происходит потому, что правильный ответ мог не попасть в пять лучших решений.

Группа решений, отнесенных к категории *wrong*, которая не содержит полностью верного ответа, тоже представляет интерес в задаче локализации по визуальным ориентирам, поскольку постановка содержит заведомо ложные данные. Численные результаты представлены в табл. 4.

По результатам экспериментов, представленным в табл. 4, можно сделать вывод о том, что отсутствие верного ответа от системы распознавания сцен отрицательно влияет на результаты (все показатели точности значительно снижаются в сравнении с табл. 2 и табл. 3), что свидетельствует о важности разработки робастных методов распознавания сцен. Предложенный подход N-локализации на основе функций проверки корректности позволяет отбрасывать некоторые из гипотез, в среднем было отброшено 17,5 % всех обработанных гипотез и эта величина не сильно варьируется в пределах исследуемых групп (16,6; 20,1 и 18,5 % соответственно).

На рис. 3 представлены значения рассчитанных показателей $recall@1$ и $recall@5$ без деления на группы.

Предложенные методы на основе сокращения области поиска (SDM+PF и SDM+RANSAC+ICP) показывают сравнимые по точности результаты



Таблица 3

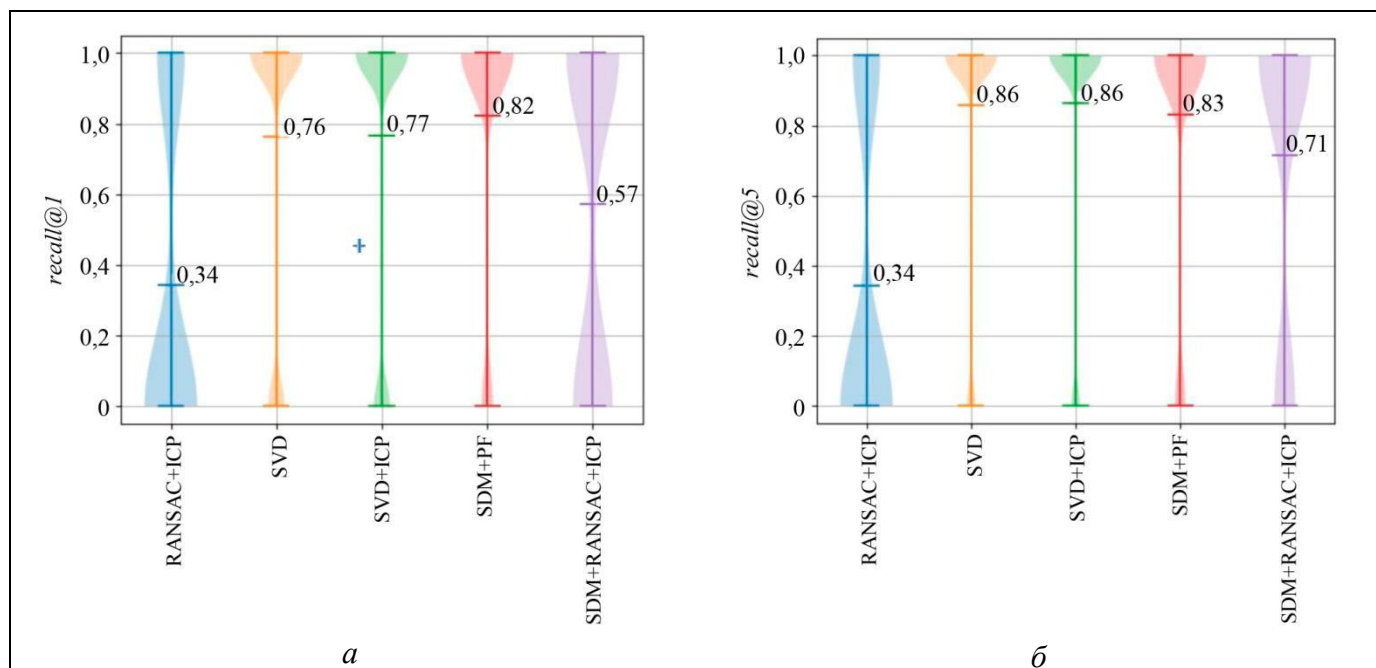
Численные результаты экспериментов для группы решений *has_true*

Метод	$ r $, м	$M(r)$, м	$ \alpha $, рад	$M(\alpha)$, рад	$ t $, с	$M(t)$, с	$R@1$, %	$R@5$, %
RANSAC+ICP	207,5	158,9	1	0,46	4038,3	4002	27,7	—
SVD	191,1	28,9	1,15	0,83	0,005	0,004	44,3	85,8
SVD+ICP	38,4	2	0,35	0,04	19,6	19,6	44,5	85,9
SDM+PF	80,4	1,5	0,29	0,05	94,8	95,6	62,2	63,9
SDM+RANSAC+ICP	71,5	6,7	0,22	0,03	911,2	878,3	61,5	84,6

Таблица 4

Численные результаты экспериментов для группы решений *wrong*

Метод	$ r $, м	$M(r)$, м	$ \alpha $, рад	$M(\alpha)$, рад	$ t $, с	$M(t)$, с	$R@1$, %	$R@5$, %
RANSAC+ICP	289	262,3	0,87	0,06	4501	4255	33,3	—
SVD	351	298,3	1,22	0,92	0,004	0,004	24,3	30,7
SVD+ICP	206,7	95,2	1,13	0,86	17,8	17,2	24,5	30,7
SDM+PF	203,4	150,8	1,34	1	79,1	83,4	25,7	25,7
SDM+RANSAC+ICP	210,5	285,5	0,59	0,07	1076,6	1030,9	13,3	20

Рис. 3. Распределения и средние значения $recall@k$ для исследованных методов: *a* – $recall@1$; *б* – $recall@5$

с подходом на основе SVD, в некоторых случаях превосходя их (см. рис. 3), что говорит о перспективности данного направления для решения задач глобальной локализации. Значения временных параметров работы алгоритмов, приведенные в исследовании, носят скорее ознакомительный характер, поскольку при реализации методов не возникали вопросы, связанные с программной оптимизацией. Подход на основе ФЧ с использованием

поточечного расстояния для сравнения облаков обладает ограничением: он показывает хорошие результаты только вблизи верного ответа и с небольшими отклонениями по углу. Исследование и применение более чувствительных метрик сравнения облаков точек, в том числе с учетом их цветовых характеристик, также потенциально улучшит результат. Отдельно стоит отметить, что метод SDM+RANSAC+ICP показал как значительное

увеличение точности, так и уменьшение времени вычисления в сравнении с классическим подходом RANSAC+ICP, что говорит о возможности использования технологии сокращения пространства поиска на основе Н-локализации и для других методов глобальной локализации по облакам точек.

ЗАКЛЮЧЕНИЕ

В работе описан подход недоопределенной локализации, предложенный в рамках комплексной технологии глобальной локализации робота на трехмерных картах, снабженных семантическим слоем. Распознавая и локализуя объекты, находящиеся в поле зрения робота, можно решить задачу распознавания сцен, сопоставив их с объектами карты. По полученным данным с помощью подхода недоопределенной локализации определяются ограничения на положение робота, которые, в свою очередь, могут быть использованы для сокращения области поиска методом локализации на основе данных дальномеров.

Проведенные эксперименты показывают, что методы локализации по облакам точек, примененные в ограничениях, полученных предлагаемой комплексной технологией, значительно повышают точность и быстродействие алгоритма в сравнении с базовым подходом глобальной локализации без учета семантики. Сокращение зоны поиска до 0,2 % от размера карты позволяет говорить о применимости авторской технологии для любых методов локализации по данным дальномеров. Также эксперименты показали, что комбинирование методов локализации по ориентирам с методами локализации по облакам точек позволяет работать в ситуации, когда верный ответ распознавания сцен «затерян» среди нескольких близких друг к другу гипотез. Таким образом, результаты, полученные по итогам проведенных экспериментов, позволяют считать, что цель исследования, определенная как улучшение качества глобальной локализации, достигнута. Повышение качества работы алгоритма происходит благодаря применению недоопределенных вычислений, которые используются в рамках комплексной технологии локализации с использованием семантических карт.

Дальнейшим направлением работ предполагается исследование и применение более чувствительных методов регистрации облаков точек для увеличения точности определения положения робота и оценки гипотез. В дополнение целесообразно рассмотреть предлагаемый подход во взаимодействии с методами Place Recognition для сокращения пространства поиска среди имеющихся

ключевых кадров, которые зачастую снабжаются координатной меткой, по которой можно определить соответствие полученным в ходе Н-локализации ограничениям. Стоит отметить, что перспективной задачей является многокритериальная оптимизация параметров используемых методов, требующая значительных вычислительных ресурсов, однако она выходит за рамки данной работы, в которой параметры были выбраны эмпирическим путем.

ЛИТЕРАТУРА

1. Chen, P., Zhao, X., Zeng, L., et al. A Review of Research on SLAM Technology Based on the Fusion of LiDAR and Vision // *Sensors*. – 2025. – Vol. 25, no. 5. – Art. no. 1447.
2. Yin, H., Xu, X., Lu, S., et al. A Survey on Global LiDAR Localization: Challenges, Advances and Open Problems // *Int. J. Comput. Vis.* – 2024. – Vol. 132, no. 8. – P. 3139–3171.
3. Rusu, R.B., Blodow, N., Beetz, M. Fast Point Feature Histograms (FPFH) for 3D Registration // *Proceedings of 2009 IEEE International Conference on Robotics and Automation*. – Kobe, Japan, 2009. – P. 3212–3217.
4. Saarinen, J., Andreasson, H., Stoyanov, T., et al. Normal Distributions Transform Monte-Carlo Localization (NDT-MCL) // *proceedings of 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*. – Tokyo, Japan, 2013. – P. 382–389.
5. Mildenhall, B., Srinivasan, P.P., Tancik, M., et al. NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis // *Commun. ACM*. – 2022. – Vol. 65, no 1. – P. 99–106.
6. Komorowski, J. MinkLoc3D: Point Cloud Based Large-Scale Place Recognition // *Proceedings of 2021 IEEE Winter Conference on Applications of Computer Vision (WACV)*. – Waikoloa, HI, USA, 2021. – P. 1789–1798.
7. Rosen, D.M., Doherty, K.J., Terán Espinoza, A., et al. Advances in Inference and Representation for Simultaneous Localization and Mapping // *Annu. Rev. Control. Robot. Auton. Syst.* – 2021. – Vol. 4, no. 1. – P. 215–242.
8. Fishler, M., Bolles, R. Random Sample Consensus: A Paradigm for Model Fitting Applications to Image Analysis and Automated Aartography // *Communication of the ACM*. – 1981. – Vol. 24. – P. 381–395.
9. Yang, H., Shi, J., Carlone, L. TEASER: Fast and Certifiable Point Cloud Registration // *IEEE Trans. Robot.* – 2021. – Vol. 37, no. 2. – P. 314–333.
10. Lusk, P.C., How, J.P. CLIPPER: Robust Data Association Without an Initial Guess // *IEEE Robot. Autom. Lett.* – 2024. – Vol. 9, no. 4. – P. 3092–3099.
11. Choy, C., Dong, W., Koltun, V. Deep Global Registration // *Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. – Seattle, WA, USA, 2020. – P. 2511–2520.
12. Rusinkiewicz, S., Levoy, M. Efficient Variants of the ICP Algorithm // *Proceedings of 3rd International Conference on 3-D Digital Imaging and Modeling*. – Quebec City, Canada, 2001. – P. 145–152.
13. Vizzo, I., Guadagnino, T., Mersch, B., et al. KISS-ICP: In Defense of Point-to-Point ICP – Simple, Accurate, and Robust



- Registration If Done the Right Way // IEEE Robot. Autom. Lett. – 2023. – Vol. 8, no. 2. – P. 1029–1036.
14. Barfoot, T.D. State Estimation for Robotics. – Cambridge: Cambridge University Press, 2017.
 15. Chen, R., Yin, H., Jiao, Y., et al. Deep Samplable Observation Model for Global Localization and Kidnapping // IEEE Robot. Autom. Lett. – 2021. – Vol. 6, no. 2. – P. 2296–2303.
 16. Zhang, X., Wang, L., Su, Y. Visual Place Recognition: A Survey from Deep Learning Perspective // Pattern Recognit. – 2021. – Vol. 113. – Art. no. 107760.
 17. Xia, L., Cui, J., Shen, R., et al. A Survey of Image Semantics-Based Visual Simultaneous Localization and Mapping: Application-Oriented Solutions to Autonomous Navigation of Mobile Robots // Int. J. Adv. Robot. Syst. – 2020. – Vol. 17, no. 3. – Art. no. 172988142091918.
 18. Jin, X., Li, X., Zhu, Y., et al. A Survey of Visual Semantic Mapping // Proceedings of 2022 3rd Asia-Pacific Conference on Image Processing, Electronics and Computers. – New York, NY, USA, 2022. – P. 295–301.
 19. Arun, K.S., Huang, T.S., Blostein, S.D. Least-Squares Fitting of Two 3-D Point Sets // IEEE Trans. Pattern Anal. Mach. Intell. – 1987. – Vol. PAMI-9, no. 5. – P. 698–700.
 20. Pramatarov, G., De Martini, D., Gadd, M., et al. BoxGraph: Semantic Place Recognition and Pose Estimation from 3D LiDAR // arXiv:2206.15154v1. – 2022. – DOI: 10.48550/arXiv.2206.15154
 21. Wu, Y., Zhang, Y., Zhu, D., et al. An Object SLAM Framework for Association, Mapping, and High-Level Tasks // arXiv:2305.07299. – 2023. – DOI: <https://doi.org/10.48550/arXiv.2305.07299>
 22. Ankenbauer, J., Lusk, P.C., Thomas, A., et al. Global Localization in Unstructured Environments Using Semantic Object Maps Built from Various Viewpoints // Proceedings of 2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). – Detroit, MI, USA, 2023. – P. 1358–1365.
 23. Guo, X., Hu, J., Chen, J., et al. Semantic Histogram Based Graph Matching for Real-Time Multi-Robot Global Localization in Large Scale Environment // IEEE Robot. Autom. Lett. – 2021. – Vol. 6, no. 4. – P. 8349–8356.
 24. Wang, Y., Jiang, C., Chen, X. GOREloc: Graph-Based Object-Level Relocalization for Visual SLAM // IEEE Robot. Autom. Lett. – 2024. – Vol. 9, no. 10. – P. 8234–8241.
 25. Matsuzaki, S., Sugino, T., Tanaka, K., et al. CLIP-Loc: Multi-modal Landmark Association for Global Localization in Object-based Maps // Proceedings of 2024 IEEE International Conference on Robotics and Automation (ICRA). – Yokohama, Japan, 2024. – P. 13673–13679.
 26. Liu, Y., Petillot, Y., Lane, D., et al. Global Localization with Object-Level Semantics and Topology // Proceedings of 2019 International Conference on Robotics and Automation (ICRA). – Montreal, Canada, 2019. – P. 4909–4915.
 27. Handbook of Constraint Programming / Ed. by Rossi F., van Beek P., Walsh T. – Amsterdam: Elsevier, 2006. – 978 p.
 28. Нариньяни А.С. Введение в недоопределенность // Информационные технологии. – 2007. – № 4. – С. 1–32. [Narinian, A.S. Introduction to Subdefinitions // Informational Technologies. – 2007. – No. 4. – P. 1–32. (In Russian)]
 29. Moscovsky, A. Subdefinite Computations for Reducing the Search Space in Mobile Robot Localization Task / In: Artificial Intelligence. RCAI 2021. Ed. by S.M. Kovalev, S.O. Kuznetsov, and A.I. Panov. – Lecture Notes in Computer Science. – Cham: Springer, 2021. – Vol. 12948.
 30. Moscovsky, A.D. Extended Object Detection: Flexible Object Description System for Detection in Robotic Tasks / In: Smart Electromechanical Systems. Ed. by A.E. Gorodetskiy, and I.L. Tarasova. – Studies in Systems, Decision and Control. – Cham: Springer, 2022. – Vol. 419. – P. 27–43.
 31. Contreras, M., Jain, A., Bhatt, N.P., et al. A survey on 3D Object Detection in Real Time for Autonomous Driving // Front. Robot. AI. – 2024. – Vol. 6, no. 11. – Art. no. 1212070
 32. Московский А.Д. Распознавание сцен в задаче глобальной локализации мобильного робота с использованием моделей векторных представлений изображений и графовых подходов // Управление большими системами. – 2025. – № 114. – С. 307–344. [Moscowsky, A.D. Scene Recognition for the Mobile Robot Global Localization Problem Based on Image Vectorization and Graphs Approaches // Large-Scale Systems Control. – 2025. – No. 114. – P. 307–344. (In Russian)]
 33. Radford, A., Kim, J.W., Hallacy, C., et al. Learning Transferable Visual Models from Natural Language Supervision // Proceedings of International Conference on Machine Learning. – Vienna, Austria, 2021. – P. 8748–8763.
 34. Баженов А.Н., Жилин С.И., Кумков С.И., и др. Обработка и анализ интервальных данных. – Москва, Ижевск: Институт компьютерных исследований, 2024. – 356 с. [Bazhenov, A.N., Zhilin, S.I., Kumkov, S.I., et al. Processing and Analysis of Interval Data. – Moscow, Izhevsk: Institute of computer researches, 2024. – 356 p. (In Russian)]
 35. Liao, Y., Xie, J., Geiger, A. KITTI-360: A Novel Dataset and Benchmarks for Urban Scene Understanding in 2D and 3D // IEEE Trans. Pattern Anal. Mach. Intell. – 2023. – Vol. 45, no. 3. – P. 3292–3310.
 36. Keetha, N., Mishra, A., Karhade, J., et al. Anyloc: Towards Universal Visual Place Recognition // arXiv:2308.00688. – 2023. – DOI: <https://doi.org/10.48550/arXiv.2308.00688>
 37. Puligilla, S., Omata, M., Zaidi, H., et al. LIP-Loc: LiDAR Image Pretraining for Cross-Modal Localization // Proceedings of 2024 IEEE/CVF Winter Conference on Applications of Computer Vision Workshops (WACVW). – Waikoloa, HI, USA, 2024. – P. 939–948.

Статья представлена к публикации членом редколлегии
Р. В. Мецераковым.

Поступила в редакцию 05.05.2025,
после доработки 12.05.2025.
Принята к публикации 16.07.2025.

Московский Антон Дмитриевич – начальник группы, Национальный исследовательский центр «Курчатовский институт», г. Москва

✉ moscovskyad@yandex.ru

ORCID iD: <https://orcid.org/0000-0002-6546-8697>

© 2025 г. Московский А. Д.



Эта статья доступна по лицензии Creative Commons «Attribution» («Атрибуция») 4.0 Всемирная.

APPLICATION OF SUBDEFINITE MODELS IN THE GLOBAL LOCALIZATION OF A MOBILE ROBOT

A. D. Moscovsky

National Research Center "Kurchatov Institute," Moscow, Russia

✉ moscovskyad@yandex.ru

Abstract. This paper considers the application of subdefinite (SD) models, a variation of constraint programming, to the localization problem of a mobile robot. A complex technology with semantic maps and point cloud maps is proposed. The technology is intended to accelerate and increase the accuracy of global localization in large, symmetric, and periodic environments. The conventional localization approach is based on data from rangefinders generating point clouds; the idea proposed instead is, first, to match the objects observed by the robot to those on the semantic map (recognize the scene), and then apply SD computations to perform localization via visual landmarks. SD computations are used to determine interval constraints on the robot's positions, represented by several sets for each hypothesis obtained during the scene recognition. Within the interval constraints, the robot is localized using rangefinder data based on a particle filter initialized within these constraints. According to the experiments conducted on the open KITTI-360 dataset, localization based on SD computations can reduce the search space to 0.2% of the original map size. The complex technology shows a significant advantage compared to approaches involving point clouds or visual landmarks only, especially in scenarios with multiple hypotheses about the matches of observed objects and those on the semantic map.

Keywords: global localization, subdefinite models, constraint programming, scene recognition, semantic maps, mobile robot.

Acknowledgments. This work was carried out under the state assignment of the National Research Center "Kurchatov Institute."