

ISSN 2712-8687

ПРОБЛЕМЫ УПРАВЛЕНИЯ

1/2026

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

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. Угольницкий,
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

I. A. Kalyaev, RAS¹ 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. Filimonyuk, 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. (Phych.),
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.),
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: ru@ipu.ru

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

Опубликовано: 25 февраля 2026 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

1.2026

СОДЕРЖАНИЕ

Анализ и синтез систем управления

Тремба А. А. Конструктивное D-разбиение для двух линейно
входящих в полином параметров. Ч. 2. Аппроксимация
областей устойчивости и анализ робастности 3

Управление в социально-экономических системах

Шумов В. В. Функция победы и ее применение
при моделировании конфликтов 22

Информационные технологии в управлении

Ефанов Д. В., Елина Е. И. Синтез самопроверяемых
дискретных устройств на основе логической коррекции
сигналов и композиции равновесных кодов «1 из 4»
и «3 из 4». Ч. 1. Метод синтеза с преобразованием всех
сигналов от объекта диагностирования 34

Кулинич А. А. Применение больших языковых моделей
в системах поддержки принятия решений. Ч. 1. Модели
объяснения и большие языковые модели 47

Управление подвижными объектами и навигация

Кулида Е. Л., Лебедев В. Г., Егоров Н. А. Методы решения
задачи оптимизации потока воздушных судов на посадку 66

Рудько И. М. Единое поле вероятности обнаружения
для группы стационарных наблюдателей 81

Хроника

**XXXIII Международная конференция «Проблемы управления
безопасностью сложных систем» 90**



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: February 25, 2026

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

1.2026

CONTENTS

Analysis and Design of Control Systems

Tremba, A. A. Constructive D-Partition for Two Parameters
Entering a Polynomial Linearly. Part II: Approximation
of Stability Regions and Robustness Analysis 3

Control in Social and Economic Systems

Shumov, V. V. The Victory Function and Its Application
in Conflict Modeling 22

Information Technology in Control

Efanov, D. V. and Yelina, Y. I. Design of Self-Checking
Discrete Devices Based on Boolean Signals Correction and
Composition of Constant-Weight Codes of the “1-Out-Of-4”
and “3-Out-Of-4” Types. Part I: The Design Method with
Conversion of All Signals from the Object under Diagnosis 34

Kulinich, A. A. Application of Large Language Models
in Decision Support Systems. Part I: Explanation Models
and Large Language Models 47

Control of Moving Objects and Navigation

Kulida, E. L., Lebedev, V. G., and Egorov, N. A. Methods
for Solving the Aircraft Landing Optimization Problem 66

Rudko, I. M. A Unified Detection Probability Field for a Group
of Stationary Observers 81

Chronicle

**33rd International Conference on Problems of Complex
Systems Security Control** 90

КОНСТРУКТИВНОЕ D-РАЗБИЕНИЕ ДЛЯ ДВУХ ЛИНЕЙНО ВХОДЯЩИХ В ПОЛИНОМ ПАРАМЕТРОВ. Ч. 2. Аппроксимация областей устойчивости и анализ робастности[#]

А. А. Тремба

Институт проблем управления им. В. А. Трапезникова РАН, г. Москва
Московский физико-технический институт, г. Долгопрудный

✉ atremba@ipu.ru

Аннотация. Для полинома, линейно зависящего от двух параметров, предлагаются несколько способов аппроксимации области устойчивости по отношению к заданной области локализации корней. Первый способ состоит в покрытии границы области устойчивости достаточно равномерной сеткой, гарантирующей такое покрытие с заданной точностью. Второй способ позволяет получить внутреннюю аппроксимацию области устойчивости с помощью отрезков или кривых, ограниченных областью устойчивости. Третий способ позволяет осуществить покрытие границы области устойчивости простыми множествами и тем самым получить кусочно-линейные внутренние и внешние аппроксимации области устойчивости. Все методы основаны на конструктивном D-разбиении, описывающем границу области устойчивости в виде набора участков дробно-рациональных кривых и отрезков. На плоскости параметров получена как простая оценка радиуса устойчивости, так и его точное значение. Реализация всех методов и алгоритмов сводится к нахождению вещественных корней полиномов.

Ключевые слова: конструктивное D-разбиение, дробно-рациональные кривые, аппроксимация области устойчивости, достаточно равномерная сетка, сеточные методы, полусеточные методы, опорная функция, радиус устойчивости.

ВВЕДЕНИЕ

Рассматривается анализ расположения корней полинома в зависимости от двух параметров, входящих в полином степени n линейно:

$$G(s, k_1, k_2) = k_1 P(s) + k_2 Q(s) + R(s). \quad (1)$$

Для заданного множества локализации корней $\mathbf{D} \subset \mathbb{C}$ требуется определить множество, называемое *областью устойчивости*, на плоскости параметров (k_1, k_2) , каждая точка которого соответствует устойчивому полиному

$$D_n = \{(k_1, k_2) : \text{все } n \text{ корней } G(s, k_1, k_2) \text{ лежат в } \mathbf{D}\}. \quad (2)$$

Напомним, что устойчивость (D-устойчивость¹) корней и полинома определяется относительно заданного множества локализации корней \mathbf{D} и обобщает понятие гурвицевых и шуровских полиномов, соответствующих устойчивым (в классическом смысле) системам непрерывного и дискретного времени. Задача поиска области устойчивости возникает при синтезе регуляторов для линейных систем управления в случае, когда два параметра регулятора входят в характеристический полином

[#] Результаты исследований, представленных в § 2 и 3, получены за счёт средств Российского научного фонда, грант № 21-71-30005-П, <https://rscf.ru/project/21-71-30005/>.

¹ Существует схожее по названию, но существенно отличающееся по смыслу понятие D-устойчивых (*D-stable*) матриц [1].

степени n линейно, или при анализе устойчивости в зависимости от двух (или одного, см. п. 2.1) параметров.

Выбором множества локализации корней \mathbf{D} можно гарантировать степень устойчивости, показатель колебательности замкнутой системы и пр. [2–4]. Предполагается, что \mathbf{D} – регулярное открытое множество. В этом случае область устойчивости также открытое множество.

Данная статья является второй частью исследования, в первой части которого [5] предложено конструктивное D-разбиение – описание границы области устойчивости с помощью D-разбиения плоскости параметров. Напомним идею этого метода (см. работы [6–10]).

Первое множество определяется отображением границы области локализации корней $\partial\mathbf{D}$ на плоскость параметров с помощью *основного уравнения*

$$K_{bnd} = \{(k_1, k_2) : G(s, k_1, k_2) = 0_{\mathbb{C}}, s \in \partial\mathbf{D}\}. \quad (3)$$

Второе множество определяется *условием понижения степени* на коэффициент при s^n :

$$K_{deg} = \{(k_1, k_2) : G_n(k_1, k_2) = 0_{\mathbb{C}}\}. \quad (4)$$

Множества (3) и (4) определяют D-разбиение – разделение плоскости параметров на односвязные области, некоторые из которых составляют область устойчивости. В каждой из областей D-разбиения число устойчивых корней не меняется при непрерывном изменении параметров внутри области при условии, что коэффициенты полинома зависят от параметров непрерывно. Уравнение в формуле (3) определяет граничный случай, при котором хотя бы один из корней полинома находится на границе $\partial\mathbf{D}$.

Приведём ряд свойств D-разбиения для полинома (1) вместе с результатами из первой части статьи [5].

Во-первых, множество K_{bnd} состоит из так называемой основной кривой и набора прямых, называемых *особыми прямыми*, определяемыми уравнениями

$$a_i k_1 + b_i k_2 + c_i = 0, \quad i = 1, \dots, M, \quad (5)$$

а K_{deg} является особой прямой (возможно, пустым множеством), определяемой уравнением

$$P_n k_1 + Q_n k_2 + R_n = 0.$$

Граница области устойчивости принадлежит объединению этих множеств:

$$\partial D_n \subset K_{bnd} \cup K_{deg}.$$

Во-вторых, если граница области локализации корней $\Gamma = \partial\mathbf{D}$ состоит из конечного набора участков дробно-рациональных кривых,

$$\Gamma = \bigcup_{\ell} \Gamma_{\ell}, \quad \Gamma_{\ell} = \{s_{\ell}(w) \in \mathbb{C} : w \in W_{\ell}\}, \\ \ell = 1, \dots, L,$$

то решение основного уравнения из формулы (3) эквивалентно решению L систем двух полиномиальных уравнений с некоторыми полиномами $P_{\ell,1}(w), P_{\ell,2}(w), Q_{\ell,1}(w), Q_{\ell,2}(w), R_{\ell,1}(w), R_{\ell,2}(w)$, зависящими от исходных полиномов P, Q, R и дробно-рациональных функций $s_{\ell}(w)$:

$$\begin{cases} k_1 P_{\ell,1}(w) + k_2 Q_{\ell,1}(w) + R_{\ell,1}(w) = 0, \\ k_1 P_{\ell,2}(w) + k_2 Q_{\ell,2}(w) + R_{\ell,2}(w) = 0, \end{cases} \quad (6) \\ w \in W_{\ell}, \quad \ell = 1, \dots, L.$$

Решением каждой системы являются дробно-рациональная кривая (*основная кривая*)

$$k_{\ell,1}(w) = \frac{1}{\det T_{\ell}(w)} (R_{\ell,2}(w) Q_{\ell,1}(w) - \\ - R_{\ell,1}(w) Q_{\ell,2}(w)), \\ k_{\ell,2}(w) = \frac{1}{\det T_{\ell}(w)} (R_{\ell,1}(w) P_{\ell,2}(w) - \\ - R_{\ell,2}(w) P_{\ell,1}(w)), \\ w \in W_{\ell}, \quad (7)$$

и набор особых прямых вида (5), соответствующих для каждого $\ell = 1, \dots, L$ таким вещественным корням (так называемым критическим частотам) уравнения

$$\det T_{\ell}(w) = P_{\ell,1}(w) Q_{\ell,2}(w) - \\ - Q_{\ell,1}(w) P_{\ell,2}(w) = 0, \quad w \in W_{\ell},$$

при которых система (6) имеет решение. Далее индекс ℓ может опускаться для краткости.

В-третьих, доказано, что при локализованном D-разбиении – ограниченном на компактном множестве \mathbf{K} – указанные дробно-рациональные кривые определены или доопределяются на замкнутых интервалах (отрезках), а особые прямые становятся отрезками. Более того, если граница \mathbf{K} состоит из конечного набора участков дробно-рациональных кривых, то D-разбиение также состоит из конечного числа участков дробно-рациональных кривых и отрезков. Предложенный метод конструктивного D-разбиения позволяет определить интервалы параметризации этих участков дробно-рациональных кривых и отрезков как



участков особых прямых с помощью нахождения вещественных корней полиномов.

В частности, граница области устойчивости (2) в пределах множества K состоит из конечного набора дробно-рациональных функций вида (7) и отрезков. Отрезки удобно представлять в параметрическом виде как

$$k(t) = t d + p, \quad t \in [t_1, t_2], \quad p, d \in \mathbb{R}^2, \quad d \neq 0,$$

где векторы $d = (-b, a)^T$, $p = -\left(\frac{ac}{a^2 + b^2}, \frac{bc}{a^2 + b^2}\right)$ получены из уравнений прямых (5), здесь и далее индексы опущены для краткости.

Параметр и точка пересечения прямой, исходяно заданной в виде $k(t) = t d + p, t \in (-\infty, +\infty)$, и прямой вида (5) задаются выражениями

$$\begin{aligned} t^* &= -\frac{c + ap_1 + bp_2}{ad_1 + bd_2} = -\frac{c + (a, b)p}{(a, b)d}, \\ k^* &= t^* d + p = -\frac{c + (a, b)p}{(a, b)d} d + p. \end{aligned} \quad (8)$$

Подчеркнем одну особенность построения локализованного D-разбиения. Дробно-рациональные участки границ области устойчивости вида (7) определены на замкнутых интервалах, но при этом может существовать бесконечный интервал, например, $(-\infty, +\infty)$ или $[w_1, +\infty)$. Для сведения области определения к конечному замкнутому интервалу используется замена параметра, см. п. 1.2.

Пересечение основных кривых с прямыми и локализованное D-разбиение

Напомним некоторые результаты, касающиеся пересечения дробно-рациональных кривых с прямыми. В частности, пересечение особых прямых (5) с основной кривой D-разбиения (7) задаётся уравнениями

$$\begin{aligned} &a_i (R_2(w)Q_1(w) - R_1(w)Q_2(w)) + \\ &+ b_i (R_1(w)P_2(w) - R_2(w)P_1(w)) + \\ &+ c_i (P_1(w)Q_2(w) - Q_1(w)P_2(w)) = 0, \\ &i = 0, \dots, K. \end{aligned} \quad (9)$$

Здесь индекс 0 соответствует особой прямой K_{deg} . Полученные уравнения относительно w полиномиальные, и их вещественные корни можно вычислить явно, обозначим их совокупность как $w_m, m = 1, \dots$. Эти корни вместе с критическими частотами w_i разбивают интервал W на отрезки и

интервалы, соответствующие простым непрерывным частям границ D-разбиения, участкам основной кривой. Собственно, эти участки и образуют криволинейную часть D-разбиения. Сами точки пересечения определяются из уравнения (7) как $k(w_m) = (k_1(w_m), k_2(w_m))$. Они, вместе с предельными точками $\lim_{w \rightarrow w_m} k(w)$, если таковые существуют, не только разбивают основную кривую на участки, но и разбивают особые прямые на отрезки или бесконечные интервалы (лучи). В случае локализованного D-разбиения бесконечные интервалы отсутствуют. Схожим образом можно найти точки самопересечений и пересечений нескольких основных кривых между собой (см. первую часть статьи [5]).

Локализованное D-разбиение удобно осуществлять в прямоугольнике $K = [\underline{k}_1, \bar{k}_1] \times [\underline{k}_2, \bar{k}_2]$, границы которого – вертикальные и горизонтальные отрезки. Для них уравнение (9) упрощается, так как достаточно рассмотреть каждую компоненту дробно рациональной кривой (7) отдельно:

$$\begin{aligned} &R_2(w)Q_1(w) - R_1(w)Q_2(w) = \\ &= x(P_1(w)Q_2(w) - Q_1(w)P_2(w)), \end{aligned} \quad (10)$$

$$x = \underline{k}_1, \bar{k}_1,$$

$$\begin{aligned} &R_1(w)P_2(w) - R_2(w)P_1(w) = \\ &= y(P_1(w)Q_2(w) - Q_1(w)P_2(w)), \end{aligned} \quad (11)$$

$$y = \underline{k}_2, \bar{k}_2.$$

После решения каждого из этих уравнений надо проверить, попадает ли вторая координата в нужный интервал, т. е. пересекается ли основная кривая со стороной прямоугольника. Например, если корни уравнения (10) равны w_m , то надо выбрать только те из них, для которых $k_2(w_m) \in [\underline{k}_2, \bar{k}_2]$, и наоборот.

Далее на основе конструктивного описания границы области устойчивости предлагаются несколько способов аппроксимации границы или самой области устойчивости. Кроме того, предложенный метод применяется для анализа положения области устойчивости (см. п. 4.3) и робастного анализа (см. п. 4.4). Полученная параметризация границы области устойчивости может быть использована для последующей оптимизации внутри области устойчивости и для иных задач, связанных с анализом и синтезом регуляторов с заданными характеристиками, определяемыми областью локализации корней.

1. АППРОКСИМАЦИЯ ГРАНИЦЫ D-РАЗБИЕНИЯ ТОЧКАМИ

Рассмотрим локализованное D-разбиение на ограниченном множестве K , например, на прямоугольнике $[\underline{k}_1, \bar{k}_1] \times [\underline{k}_2, \bar{k}_2]$. После применения алгоритма конструктивного D-разбиения по формулам (8) и (10), (11) можно определить пересечение особых прямых (5) и основных кривых (7) D-разбиения с границами множества K . В результате будет найдена область устойчивости D_n , ограниченная конечным числом отрезков $k_m(t)$ и участками дробно-рациональных основных кривых $k_\ell(w)$:

$$\partial D_n = \cup_m k_m(t) \cup_\ell k_\ell(w).$$

Согласно леммам 1 и 4 из первой части статьи [5], каждый из участков границы $k_m(t)$, $k_\ell(w)$ параметризован отрезком $W_\ell = [w_{\ell,1}, w_{\ell,2}]$ или замкнутым бесконечным интервалом. Более того, в теореме 1 [5] продемонстрирована возможность заменить бесконечный интервал параметризации отрезком (или отрезками). Далее эта замена в контексте аппроксимации кривой рассмотрена отдельно в п. 1.2.

Задача состоит в том, чтобы выбрать конечный набор точек – *узлов*, в совокупности образующих сетку $K_{grid} = \{k_r, r=1, \dots\}$ на этих кривых и отрезках, аппроксимирующий границу области устойчивости с заданной мелкостью $\rho > 0$. Мелкость сетки относительно множества определяется ограничением расстояния Хаусдорфа между набором точек и множеством, в данном случае – границей области устойчивости. Будем выбирать узлы, лежащие на границе, $k_r \in \partial D_n$. Они должны удовлетворять условию

$$\text{dist}_H(K_{grid}, \partial D_n) = \max_{k \in \partial D_n} \min_r \|k - k_r\| \leq \rho. \quad (12)$$

Таким образом, для построения сетки необходимо и достаточно, чтобы круги радиуса ρ , центры которых находятся в k_r , покрывали границу. Будем называть такую аппроксимацию и соответствующую сетку *достаточно равномерными*.

Аппроксимацию границы просто сделать для каждого участка границы по отдельности, явно добавив узлы в точках сопряжения участков границы. Для участков-отрезков $k_m(t) = t d_m + p_m$, $t \in [t_{m,1}, t_{m,2}]$, предлагается очевидная равномер-

ная аппроксимация с N интервалами, включающая оба конца отрезка:

$$k_{m,r} = (t_{m,1} + r\delta) d_m + p_m, \quad r = 0, 1, \dots, N, \\ \delta = \frac{t_{m,2} - t_{m,1}}{N}, \quad N = \left\lceil \frac{t_{m,2} - t_{m,1}}{2\rho} \right\rceil. \quad (13)$$

Расстояние между двумя узлами не превышает 2ρ . Эта сетка – оптимальная среди всех сеток, включающих крайние точки, с той точки зрения, что она равномерная и её число узлов минимально.

С криволинейными участками границы $k_\ell(w)$, определёнными на замкнутых интервалах, ситуация сложнее. Напомним, что на этих интервалах функция $k_\ell(w)$ непрерывна. Рассмотрим один участок, $k(w)$, $w \in W$, опустив индекс.

Формально можно получить равномерную аппроксимацию участка гладкой кривой с любой заданной точностью, если использовать так называемую натуральную (естественную) параметризацию участка кривой с помощью функции длины кривой от точки $k(w_1)$ в сторону увеличения значения параметра. Функцию длины можно определить по «скорости движения» точки вдоль кривой, $v(w) = \|k'(w)\| = \sqrt{k_1'(w)^2 + k_2'(w)^2}$, которую будем называть параметрической скоростью или просто скоростью. Функция длины равна

$$\lambda_{w_1}(w) = \int_{w_1}^w v(\tau) d\tau, \quad w \geq w_1 \\ \text{и} \\ \lambda_{w_1}(w) = - \int_w^{w_1} v(\tau) d\tau, \quad w < w_1. \quad (14)$$

Для монотонной функции $\lambda(w)$, $w \geq w_1$, строится обратная функция $w(\lambda)$. Затем отрезок $[0, \lambda(w_2)]$ разбивается равномерно числами λ_r на участки длины не больше 2ρ аналогично формуле (13). Итоговый набор $k_r = k(w(\lambda_r))$ разбивает кривую $k(w)$, $w \in W$, на участки равной длины, каждый из которых не больше 2ρ . Из непрерывности кривой и правила треугольника следует, что расстояние от каждой точки участка до одной из соседних узловых точек, включающих концы $k(w_1)$ и $k(w_2)$, не превышает ρ . К сожалению, аналитическое вычисление сначала интеграла, а потом и обратной функции обычно недоступно, так как компоненты вектора скорости



$k_1'(w), k_2'(w)$ задаются дробно-рациональными функциями.

1.1. Алгоритмы построения достаточно равномерной сетки

Предлагаются два алгоритма, использующие верхнюю оценку скорости на подынтервалах и последовательно добавляющие узлы сетки. Особенность алгоритмов состоит в том, что оценка скорости не равномерная, и это позволяет «расходовать» узлы сетки эффективнее. Полученная сетка, удовлетворяющая условию (12), будет достаточно равномерной, при этом в первом алгоритме узлы на кривой будут располагаться регулярным (по w) образом.

Вычислим набор точек, содержащих экстремальные значения скорости внутри отрезка W_ℓ с помощью необходимого условия экстремума:

$$\begin{aligned} & \text{Arg extr}_{w \in W_\ell} v(w) \in \\ & \in \text{Arg extr}_{w \in W_\ell} v(w)^2 \in \{w : (v(w)^2)' = 0\} = \quad (15) \\ & = \{w : k'_{\ell,1}(w)k''_{\ell,1}(w) + k'_{\ell,2}(w)k''_{\ell,2}(w) = 0\}. \end{aligned}$$

Из последнего выражения видно, что так как функции дробно-рациональные, поиск нахождения экстремальных точек сводится к вычислению корней соответствующего полинома. Обозначим множество экстремальных точек как $\{w_s\}$ и определим функцию $v_{\max}(w_a, w_b)$ для произвольного отрезка $[w_a, w_b] \in W_\ell, w_a < w_b$, возвращающую максимальное значение скорости на этом интервале:

$$\begin{aligned} v_{\max}(w_a, w_b) = \max \{ & v(w_a), v(w_b), \\ & \max \{v(w_s) : w_s \in (w_a, w_b)\} \}. \quad (16) \end{aligned}$$

Формула (16) использует теорему Ферма, т. е. тот факт, что на отрезке максимальное значение непрерывной дифференцируемой функции достигается либо на концах отрезка, либо внутри него. Для внутренних точек максимума выполняется необходимое условие экстремума, которому и удовлетворяют точки из набора $\{w_s\}$. Вычисление функции $v_{\max}(w_a, w_b)$ можно упростить, если среди экстремальных использовать только точки (локального) максимума, определяемые условием на вторую производную: $v''(w_s) < 0$. При равенстве второй производной нулю надо смотреть на производные высшего порядка или оставлять такие точки в качестве кандидатов.

Таким образом, согласно формуле (14), на отрезке $[w_a, w_b]$ имеет место оценка для длины кривой от конца $k_\ell(w_a)$ до промежуточной точки $k_\ell(w)$:

$$\lambda_{w_a}(w) \leq v_{\max}(w_a, w_b) \cdot (w - w_a), w \in [w_a, w_b]. \quad (17)$$

Аналогичная оценка верна и для длины кривой в обратную сторону, от w_b к w . Из этих оценок и правила треугольника для каждой половины кривой, до и после $w_c = (w_a + w_b) / 2$, имеют место верхние и нижние неравенства на длину кривой:

$$\begin{aligned} \|k_\ell(w_a) - k_\ell(w)\| & \leq \lambda_{w_a}(w) \leq \lambda_{w_a}(w_c) \leq \\ & \leq v_{\max}(w_a, w_b) \cdot \frac{w_b - w_a}{2}, w \in [w_a, w_c], \\ \|k_\ell(w_b) - k_\ell(w)\| & \leq \lambda_{w_b}(w) \leq \lambda_{w_b}(w_c) \leq \\ & \leq v_{\max}(w_a, w_b) \cdot \frac{w_b - w_a}{2}, w \in [w_c, w_b]. \end{aligned}$$

Из них следует достаточный критерий: если $v_{\max}(w_a, w_b) \cdot (w_b - w_a) \leq 2\rho$, то участок кривой $k_\ell(w), w \in [w_a, w_b]$, лежит в объединении двух кругов радиуса ρ с центрами в $k(w_a)$ и $k(w_b)$.

Предлагается следующий итеративный алгоритм нахождения достаточно равномерной сетки K_{set} , покрывающей участок кривой $k_\ell(w), w \in W_\ell$, использующий функцию $v_{\max}(w_a, w_b)$ на подынтервалах.

Алгоритм 1. Достаточно равномерная сетка для участка дробно-рациональной кривой.

Дано: дробно-рациональная кривая $k_\ell(w), w \in W_\ell = [w_1, w_2]$, знаменатель которой не обращается в нуль на отрезке W_ℓ , параметр мелкости $\rho > 0$.

1. Вычислить экстремальные точки $\{w_s\}$ для скорости по формуле (15) (или точки локальных максимумов), принадлежащие W_ℓ , и определить функцию верхней границы $v_{\max}(w_a, w_b)$.
2. Задать начальный список узлов из двух точек как $\{k(w_1), k(w_2)\}$. Задать начальный список интервалов, содержащий один элемент – интервал $[w_1, w_2]$.
3. Если список интервалов пуст, то завершить работу алгоритма и вернуть список узлов. В противном случае выбрать любой из интервалов (например, первый по счёту, или крайний левый);

обозначить его как $[w_a, w_b]$ и извлечь его из списка интервалов.

4. Если выполняется неравенство $v_{\max}(w_a, w_b) \times (w_b - w_a) \leq 2\rho$, то перейти к шагу 3.

5. Разбить интервал $[w_a, w_b]$ пополам точкой $w_c = (w_a + w_b) / 2 \in (w_a, w_b)$, добавить $k(w_c)$ в список узлов, добавить интервалы $[w_a, w_c]$ и $[w_c, w_b]$ в список интервалов и перейти к шагу 3.

Результат алгоритма: набор параметров из интервала $[w_1, w_2]$ и соответствующий ему набор узлов, являющийся достаточно равномерной сеткой.

Алгоритм 1 конечен, так как параметрическая скорость ограничена сверху на всём интервале W_ℓ , см. оценку (18) ниже. Из построения следует, что полученная алгоритмом 1 сетка не только осуществляет покрытие кривой окружностями радиуса ρ , но и разбивает саму кривую на участки длины не больше 2ρ . Алгоритм 1 тривиально обобщается на любые кривые, для которых есть оценка скорости $\bar{v}(w) \geq v(w)$. При этом он эффективно реализуется, если эта оценка скорости является полиномом или иной функцией с просто или заранее вычисляемыми экстремумами (максимумами). На рис. 1 приведена параметрическая скорость и достаточно равномерная сетка для участка границы области устойчивости в примере 1 (см. § 5).

В частном случае при использовании в качестве верхней оценки максимальной скорости $\bar{v}(w) = v_{\max}(w_1, w_2) = \text{const}$ результатом алгоритма 1 будет равномерная (бинарная) сетка:

$$k_r = k(w_r), \quad w_r = w_1 + r \delta_2, \quad \delta_2 = \frac{w_2 - w_1}{2^M},$$

$$M = \max \left\{ 0, \left\lceil \log_2 \frac{v_{\max}(w_1, w_2)(w_2 - w_1)}{\rho} \right\rceil - 1 \right\}, \quad (18)$$

$$r = 0, \dots, 2^M.$$

Равномерная бинарная сетка в общем случае хуже (по числу узлов, но не более чем в два раза) равномерной сетки, построенной по оценке максимальной скорости $v_{\max}(w_1, w_2)$ аналогично формуле (13) как

$$w_r = w_1 + r \delta, \quad r = 0, 1, \dots, N, \quad \delta = \frac{w_2 - w_1}{N},$$

$$N = \left\lceil \frac{v_{\max}(w_1, w_2)(w_2 - w_1)}{2\rho} \right\rceil. \quad (19)$$

При этом в алгоритме 1 расстояние между узлами сетки $k_r = k(w_r)$ адаптируется под максимальную скорость в каждом подынтервале, и на практике он эффективнее, так как требует существенно меньшего числа узлов, чем равномерная сетка (19). Это связано с тем, что функция скорости связана с дробно-рациональными функциями и потому существенно неравномерна, особенно на краях интервалов определения.

Отметим, что поскольку каждый раз при делении отрезка берётся его середина $w_c = (w_a + w_b) / 2$, то полученные алгоритмом 1 интервалы будут кратны самому маленькому интервалу $\delta_2 = 2^{-M}(w_2 - w_1)$ для того же M , поскольку хотя бы на одном из подынтервалов максимальное значение скорости достигается.

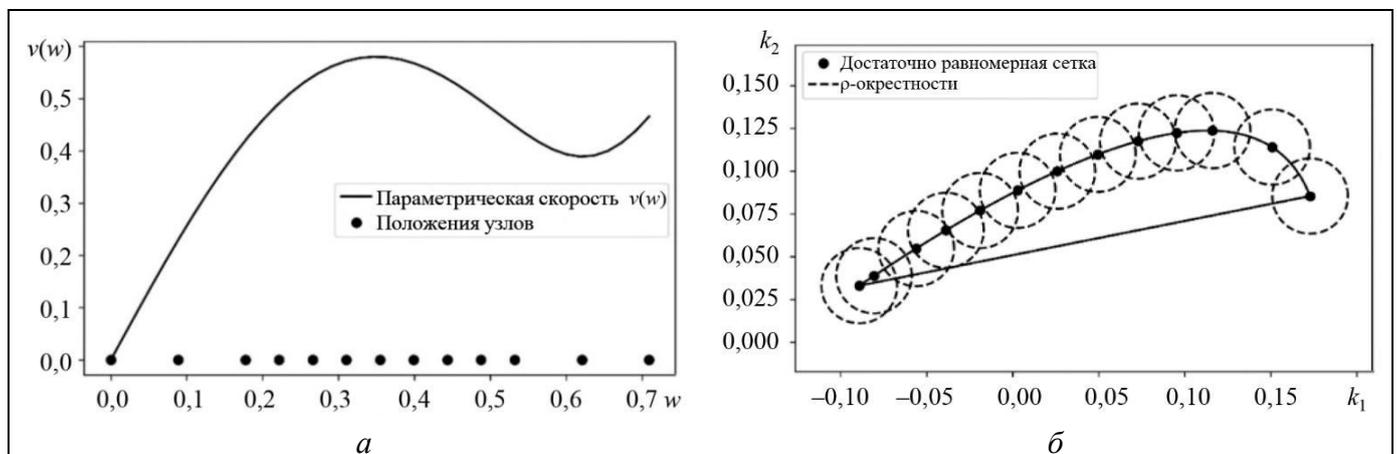


Рис. 1. *a* – Параметрическая скорость для криволинейной части границы области устойчивости примера 1, *b* – достаточно равномерная сетка, построенная по алгоритму 1, вместе с набором покрывающих кругов



Вычисление функции $v_{\max}(w_a, w_b)$ в формуле (17) становится тривиальным, если интервал не содержит экстремумов:

$$v_{\max}(w_a, w_b) = \max\{v(w_a), v(w_b)\},$$

если $\nexists w_s \in (w_a, w_b)$.

Используя этот факт, алгоритм 1 можно упростить, сразу разбив интервал $[w_1, w_2]$ точками, подозрительными на экстремум (или точками максимума). Тогда на каждом подынтервале максимальная скорость будет определяться исключительно скоростями на концах отрезка, то же относится и к последующему подразделению на более мелкие отрезки.

Алгоритм 2. Достаточно равномерная сетка для участка дробно-рациональной кривой (упрощённый вариант).

Дано: дробно-рациональная кривая $k_\ell(w)$, $w \in W_\ell = [w_1, w_2]$, знаменатель которой не обращается в нуль на отрезке W_ℓ , параметр мелкости $\rho > 0$.

1. Вычислить экстремальные точки $\{w_s\}$ для скорости по формуле (15) (или точки локальных максимумов), принадлежащие W_ℓ , и разбить ими интервал $[w_1, w_2]$.

2. Задать начальный список узлов как $\{k(w_1), k(w_2)\} \cup \{k(w_s)\}$. Задать начальный список интервалов, полученный разбиением интервала $[w_1, w_2]$ набором $\{w_s\}$.

3. Если список интервалов пуст, то завершить работу алгоритма и вернуть список узлов. В противном случае выбрать любой из интервалов (например, первый по счёту, или крайний левый, обозначим его как $[w_a, w_b]$), и извлечь его из списка интервалов.

4. Если выполняется неравенство $\max\{v(w_a), v(w_b)\} \cdot (w_b - w_a) \leq 2\rho$, то перейти к п. 3.

5. Разбить интервал $[w_a, w_b]$ пополам точкой $w_c = (w_a + w_b) / 2 \in (w_a, w_b)$, добавить $k(w_c)$ в список узлов, добавить интервалы $[w_a, w_c]$ и $[w_c, w_b]$ в список интервалов и перейти к п. 3.

Результат алгоритма 2 такой же, как у алгоритма 1: набор параметров из интервала $[w_1, w_2]$ и соответствующий ему набор узлов, являющийся достаточно равномерной сеткой.

По сравнению с алгоритмом 1, алгоритм 2 требует меньше вычислений, так как не требуется проверять принадлежность подозрительных на экстремум параметров $\{w_s\}$ текущему интервалу, но сетка при этом получается нерегулярной.

1.2. Бесконечные интервалы и перепараметризация

Дополняют алгоритмы 1 и 2 следующие рассуждения. Во-первых, для локализованного D-разбиения, согласно леммам 1–3 из первой части статьи [5], конечный интервал, на котором определён (или доопределён) участок дробно-рациональной кривой $k_\ell(w)$ вида (7) внутри множества K , всегда замкнутый. Чтобы применять аппроксимацию на бесконечных интервалах по w , делается замена переменных [11]. Пусть без ограничения общности дробно-рациональная кривая $k_\ell(w)$ задана и непрерывна на интервале $W_\ell = [w_1, \infty)$ и выбрана конечная точка $w_0 < w_1$ – некоторое конечное значение, $w_0 \neq -\infty$, например, $w_0 = w_1 - 1$. Поскольку D-разбиение локализовано, то существует предельная точка $k_\infty = \lim_{w \rightarrow \infty} k_\ell(w) \in K$ – в противном случае верно утверждение леммы 3 [5] и достаточно рассматривать конечный интервал, на котором кривая находится внутри множества K . Заменой переменной $w = w_0 + 1/u$, $u \in U = (0, 1/(w_1 - w_0)] \subset (0, 1]$, получается другая параметризация того же участка кривой $k_{\ell,u}(u) = k_\ell(w_0 + 1/u)$, $u \in U$, при этом вместо узла $k_{\ell,u}(0)$ берётся та же предельная точка $k_\infty = \lim_{u \rightarrow +0} k_{\ell,u}(u)$, а значение скорости в нем полагается равным нулю. Эта замена позволяет получить параметризацию конечных участков границы на конечном интервале. Случай $W_\ell = (-\infty, +\infty)$ приводит к разбиению на два конечных интервала после замены параметра.

Аналогичное доопределение скорости нулевым значением применяется при доопределении значения дробно-рациональной функции на конце интервала $[w_1, \dots)$, если w_1 является корнем знаменателя дробно-рациональной функции, но предельная точка $\lim_{w \rightarrow w_1} k_\ell(w)$ существует.

Отметим, что для не локализованного D-разбиения части границы $k_\ell(w)$ могут быть неограниченными из-за неограниченного либо открытого интервала W_ℓ , включая случаи, когда гра-

ница интервала W_ℓ является нулём знаменателя одной из компонент $k_\ell(w)$. В этих случаях предел $k_\ell(w)$ не существует, а неограниченную кривую нельзя приблизить конечным числом точек с заданной точностью. В таких случаях дробно-рациональная кривая (7) имеет асимптоты с соотношением

$$k_{\ell,1}(w) : k_{\ell,2}(w) = R_{\ell,2}(w) Q_{\ell,1}(w) - R_{\ell,1}(w) Q_{\ell,2}(w) : \\ : R_{\ell,1}(w) P_{\ell,2}(w) - R_{\ell,2}(w) P_{\ell,1}(w) \\ \text{при } w \rightarrow \infty,$$

и в качестве множества, аппроксимирующего D-разбиение, можно рассмотреть сочетание набора точек и лучей. Анализ качества и точности такой аппроксимации является отдельной задачей и выходит за рамки настоящей статьи.

2. ПОЛУСЕТОЧНЫЕ МЕТОДЫ

Наряду с изложенным в § 1 методом поточечной аппроксимации границ можно предложить ряд численных методов оценки как границы D-разбиения, так и областей D-разбиения. Их можно разделить на два способа.

Идея первого способа состоит в параметризации плоскости параметров k_1, k_2 двумя вспомогательными параметрами и формировании дискретной сетки по одному из вспомогательных параметров. Далее надо взять полученный набор линий (непрерывных по второму вспомогательному параметру, например, прямых или кривых) и найти пересечение границы D-разбиения с этой сеткой. Во-первых, с помощью пересечений будет получена некоторая поточечная аппроксимация границы. Во-вторых, если на этих линиях определить интервалы устойчивости – те части линий, которые попадут в область устойчивости – то будет получена внутренняя аппроксимация области устойчивости. Соответствующие методы, основывающиеся на расслоении пространства параметров, назовём полусеточными (*semi-grid*). В качестве простейшего примера можно взять набор горизонтальных (или вертикальных) линий.

По сути, используются одномерная параметризация (декомпозиция, сечение, расслоение) плоскости или пространства параметров и сетка для этой параметризации. При трёх и более параметрах аналогичная идея применяется для визуализации трёхмерного D-разбиения, см. п. 2.5.

Второй способ состоит в разбиении плоскости параметров на простые множества-ячейки K_i и выборе тех из них, которые пересекаются с грани-

цей области устойчивости. Эти множества образуют покрытие границы области устойчивости и, следовательно, дополнение к их объединению содержит внутреннюю аппроксимацию области устойчивости на плоскости параметров. При этом по одной части границы объединения покрывающего множества формируется внутренняя, а по другой – внешняя аппроксимация области устойчивости, см. рис. 3. Проверки пересечения границы D-разбиения и множеств K_i предлагается сделать эффективными, выбирая K_i как элементы регулярной сетки (*grid*).

Оба способа используют результат конструктивного D-разбиения в виде набора участка границ области устойчивости или одномерное D-разбиение. При этом для второго способа конструктивное D-разбиение необходимо, так как оно позволяет гарантировать отсутствие «пропущенных» компонент области устойчивости, целиком лежащих в одной из ячеек.

Приведём ряд полусеточных методов, а затем, в § 3, ряд сеточных методов, предварив их описанием одномерного D-разбиения.

2.1. Одномерное D-разбиение и его связь с конструктивным D-разбиением

Одномерное D-разбиение относится к полиному $G(s, t)$, коэффициенты которого зависят от скалярного параметра $t \in \mathbb{R}$. В этом случае вещественная прямая параметра разбивается некоторыми значениями t_i на отрезки и лучи, соответствующие областям D-разбиения. Точки t_i , определяющие границы одномерного D-разбиения, будем называть критическими точками. Как и в случае двух параметров, они определяются основным уравнением

$$G(s, t) = G_n(t)s^n + G_{n-1}(t)s^{n-1} + \dots \\ + G_1(t)s + G_0(t) = 0_c, s \in \Gamma, \quad (20)$$

и уравнением понижения степени $G_n(t) = 0$. Если граница области локализации корней Γ описывается кусочно-дробно-рациональной кривой от параметра w , а полином G зависит от t полиномиально, то основное уравнение можно свести к системе двух полиномиальных уравнений от двух неизвестных.

Ситуация упрощается, если полином зависит от t линейно, тогда основное уравнение принимает вид:

$$G(s, t) = t P_t(s) + R_t(s) = 0_c. \quad (21)$$



Его можно решить явно, предполагая отсутствие общих корней² полиномов $P_i(s), R_i(s)$ на Γ . Критические точки t_i , разбивающие прямую параметров на отрезки (и лучи) с постоянным числом устойчивых корней, удовлетворяют уравнению

$$t = -\frac{R_i(s(w))}{P_i(s(w))}. \quad (22)$$

С учётом комплекснозначной правой части, можно исключить переменную t , решив уравнение относительно w :

$$\operatorname{Im} \frac{R_i(s(w))}{P_i(s(w))} = 0.$$

Это уравнение эквивалентно уравнению

$$\begin{aligned} & \operatorname{Im} R_i(s(w)) \operatorname{Re} P_i(s(w)) = \\ & = \operatorname{Re} R_i(s(w)) \operatorname{Im} P_i(s(w)), \quad s \in \Gamma, \end{aligned} \quad (23)$$

за исключением тех точек, где $P_i(s(w)) = 0$. Такие точки заведомо не являются решением исходного уравнения (21), так как в них $R_i(s(w)) \neq 0$ из-за отсутствия общих корней. Уравнение (23) сводится к полиномиальному относительно w для дробно-рациональных функций $s(w)$. Найденные корни w_i остаётся подставить в уравнение (22), получив t_i . К ним добавляется корень уравнения понижения степени $t_0 = R_{i,n} / P_{i,n}$, если он вещественный и $P_{i,n} \neq 0$. Наконец, для интервалов (t_i, t_{i+1}) , на которые разбивается прямая параметра, определяется число устойчивых корней, например, по корням полинома в серединах отрезков $G(s, (t_i + t_{i+1})/2)$ и в двух точках за пределами интервала $[\min_i t_i, \max_i t_i]$.

Аналогично двумерному D-разбиению, для дробно-рациональной функции границы $s(w)$ можно сначала привести уравнение (21) к полиномиальной зависимости от w , а только затем переходить к соотношениям (22) и (23). Если граница множества локализации корней Γ состоит из не-

скольких участков, то уравнение (21) надо решать для каждого участка.

В общем случае для полинома, линейно зависящего от m параметров, одномерное D-разбиение позволяет в явном виде найти пересечение области устойчивости с произвольной прямой (или отрезком)

$$k(t) = td + p, \quad t \in [t_1, t_2], \quad p, d \in \mathbb{R}^m,$$

где границы могут быть как ограниченными, так и нет. В двумерном случае полином (1) принимает вид (21):

$$\begin{aligned} G(s, t) = G(s, k_1(t), k_2(t)) = \\ = (d_1 P(s) + d_2 Q(s))t + p_1 P(s) + p_2 Q(s) + R(s). \end{aligned} \quad (24)$$

Критические точки t_i являются корнями некоторого полинома согласно формуле (23). Среди них выбираются только те, которые принадлежат интервалу $[t_1, t_2]$. Этот метод применяется для анализа произвольных прямых и отрезков в пространстве параметров любой размерности, в том числе как вспомогательный для анализа неявно заданных множеств – в качестве оракула, определяющего пересечение прямой с областью устойчивости [12].

Отметим, что если для полинома, зависящего от двух параметров, с помощью конструктивного D-разбиения получена характеристика границ области устойчивости в виде набора кривых и прямых, то пересечение области устойчивости с прямыми в пространстве параметров легко получить, применяя результаты о пересечении границ D-разбиения с прямой. В результате будут получены те же отрезки устойчивости на исследуемой прямой, что и с помощью одномерного D-разбиения.

Априори сравнить эффективность этих двух способов – пересечения с кривыми участками границы и одномерного D-разбиения – сложно, так как их трудоёмкость зависит от числа участков границы области устойчивости и степеней вспомогательных полиномов. При этом стоит иметь в виду, что координаты расположения участков границы областей устойчивости на плоскости легко оцениваются, см. п. 4.3. Оценка расположения участков границы позволяет эффективно исключать пересечение отрезков и прямых с областями D-разбиения, например, путём сравнения прямоугольников, содержащих участки границ областей D-разбиения. С другой стороны, одномерное D-разбиение для полинома вида (24), линейного относительно параметра, осуществляется с помощью решения уравнения (21) напрямую и не требует деления границы D-разбиения на основную кривую и особые прямые.

² В противном случае полином заведомо неустойчив, так как корень на границе не принадлежит открытому множеству локализации корней \mathbf{D} . Этот случай аналогичен случаю двух параметров с отсутствием общих корней у полиномов P, Q, R . Если рассматриваются не открытые множества \mathbf{D} и общий корень на границе считается устойчивым, его можно сократить.

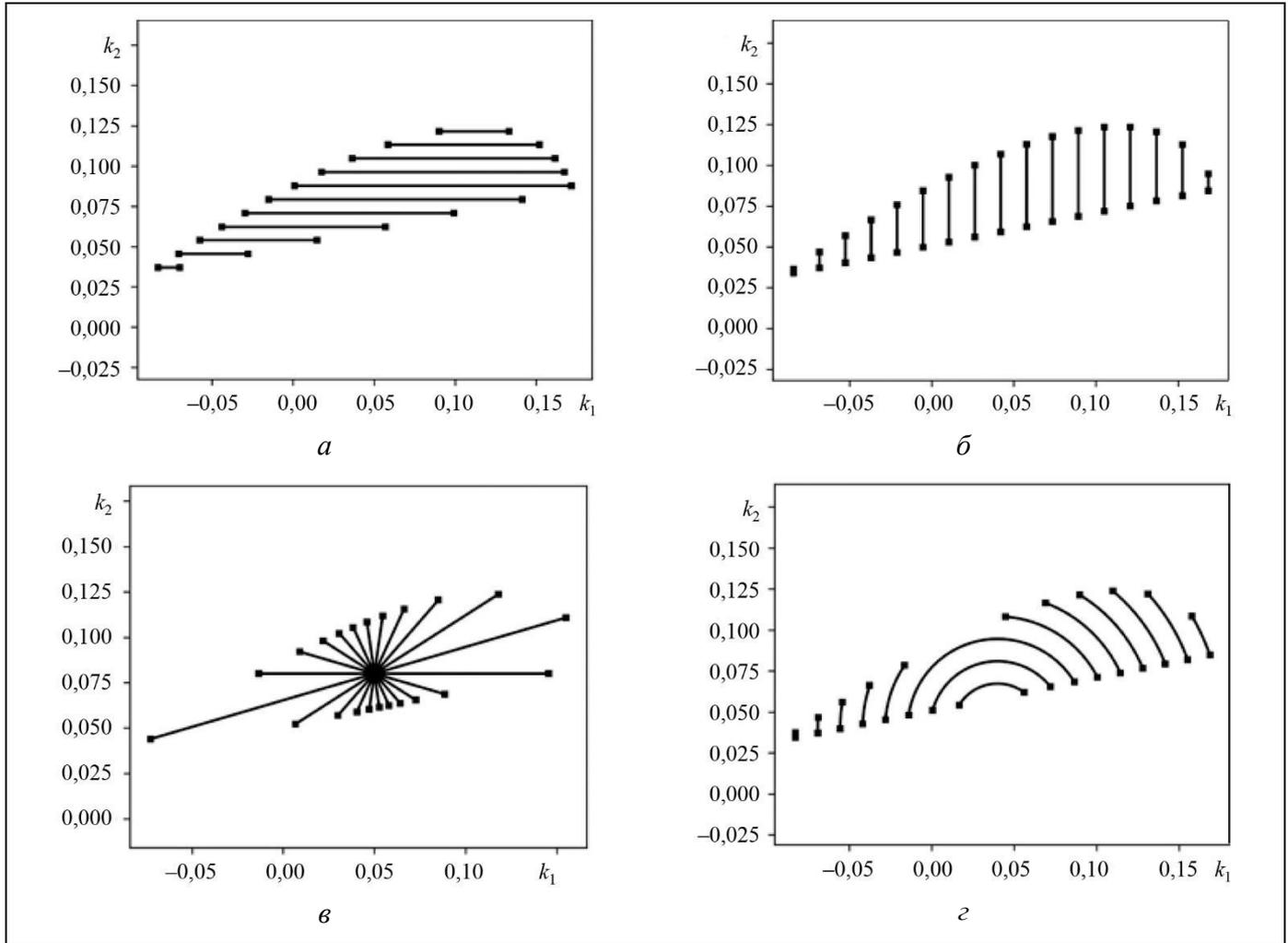


Рис. 2. Внутренние аппроксимации множества устойчивости на сетках: *а* – из горизонтальных прямых, *б* – вертикальных, *в* – концентрических окружностей, *г* – по угловой сетке

2.2. Полусеточный метод с параллельными прямыми

Как было упомянуто выше, проще всего применять метод, зафиксировав один из параметров. В случае двух параметров фиксированному первому параметру на плоскости k_1, k_2 соответствуют вертикальные прямые, а фиксированному второму – горизонтальные. Если задана область локализации $K = K_1 \times K_2 = [k_1, \bar{k}_1] \times [k_2, \bar{k}_2]$, проще всего брать равномерную сетку, отталкиваясь либо от мелкости, либо от числа прямых $N+1$, см. примеры на рис. 2, *а* и *б*. Например, вертикальные прямые K_i задаются в явном и параметрическом виде как

$$k_1 - k_{1,i} = 0, k_{1,i} = k_1 + i \frac{\bar{k}_1 - k_1}{N}, i = 0, \dots, N, \quad (25)$$

$$\left(\begin{array}{c} 0 \\ \bar{k}_2 - k_2 \end{array} \right) t + \left(\begin{array}{c} k_{1,i} \\ k_2 \end{array} \right), t \in [0, 1], i = 0, \dots, N.$$

Пересечение границ D-разбиения определяется уравнениями (10) и (11) для основной кривой и (8) для прямых либо можно использовать одномерное D-разбиение по t полинома (24) с использованием параметризации (25). На каждой прямой выделяются отрезки, для которых полином устойчив. Совокупность таких отрезков на всех прямых K_i , соответствующих областям устойчивости, образует внутреннюю аппроксимацию области устойчивости.

2.3. Полусеточный метод с лучами (угловая сетка)

Рассмотрим точку k_0 на плоскости параметров, например, соответствующую устойчивому полиному. Выберем несколько лучей, проходящих через эту точку, и рассмотрим пересечение этих лучей с границами D-разбиения. Для построения удобнее брать не лучи, а прямые, у которых угол



наклона относительно оси равномерно распределён на полуокружности. Например, для N прямых можно выбрать сетку в виде уравнения

$$(k_1 - k_{0,1}) \sin \frac{i\pi}{N} - (k_2 - k_{0,2}) \cos \frac{i\pi}{N} = 0, \quad (26)$$

$$i = 0, \dots, N-1,$$

или, в параметрической записи, $k_0 + t \left(\cos \frac{i\pi}{N}, \sin \frac{i\pi}{N} \right)^T$. В отличие от случая прямых, параллельных координатным осям, точки пересечения с участками границы области устойчивости надо искать с помощью формул (8) и (9). В качестве альтернативы можно использовать одномерное D-разбиение. Если точка k_0 , через которую проходят прямые, соответствует устойчивому полиному, то надо брать минимальные по модулю положительные и отрицательные параметры прямой: $\max_{m:t_m < 0} t_m \leq t \leq \min_{m:t_m > 0} t_m$, где t_m соответствуют точкам пересечения с основной кривой и особыми кривыми. В этом случае будет найдено пересечение прямой с компонентой области устойчивости, содержащей точку k_0 . Пример угловой сетки приведён на рис. 2, в.

Отметим, что при построении сетки по углу можно брать и набор случайно направленных прямых, а точка k_0 не обязательно должна соответствовать устойчивому полиному.

2.4. Полусеточный метод с концентрическими окружностями (радиальная сетка)

Наконец, рассмотрим параметризацию с помощью набора концентрических окружностей $K_i = \{k : \|k - k_0\| = r_i\}$. Точки пересечения границ D-разбиения с окружностями находятся явно: для особых прямых в параметрической форме $td + p$ из квадратичного уравнения

$$(td_1 + p_1 - k_{0,1})^2 + (td_2 + p_2 - k_{0,2})^2 = r_i^2, \quad (27)$$

а для участков основной кривой $k_\ell(w)$, $w \in W_\ell$, — из сводимых к полиномиальным по w уравнений:

$$(k_{\ell,1}(w) - k_{0,1})^2 + (k_{\ell,2}(w) - k_{0,2})^2 = r_i^2. \quad (28)$$

На рис. 2, г приведён пример радиальной сетки с центром вне исследуемого множества.

2.5. Полусеточный метод с параллельными (гипер-)плоскостями

Для полинома, зависящего от трёх параметров, возникающего, в частности, при анализе системы, замкнутой ПИД-регулятором, по одному из параметров берётся сетка. Для каждого фиксированного значения указанного параметра строится D-разбиение по оставшимся двум параметрам, и итоговые плоскости «стыкуются» между собой. Таким образом, построение и визуализация трёхмерной области сводится к построению ряда двумерных областей устойчивости. Этот классический приём используется не только для синтеза стабилизирующих регуляторов, но и для синтеза H_∞ -регуляторов и пр. [3].

В общем случае фиксация одного параметра уменьшает число свободных параметров, и можно построить или аппроксимировать D-разбиение по остальным параметрам.

3. СЕТОЧНЫЕ МЕТОДЫ (РАЗБИЕНИЕ ПЛОСКОСТИ ПАРАМЕТРОВ)

В полусеточных методах отсутствует оценка того, как граница области устойчивости ведёт себя вне множеств, образующих расслоение пространства параметров. Известны только точки её пересечения с прямыми (окружностями, гиперплоскостями), образующими границы расслоения. Конечно, если задана область локализации K , расслоения порождают ряд ограниченных множеств (отрезков или сегментов окружности), но их размер сравним с размером области K .

В сеточных методах предлагается более мелкое разбиение пространства параметров или области локализации на множества K_i , управляемое двумя сетками. Подчеркнём отличие предлагаемого метода от изложенного в работе [13]. Обычно предлагается искать области, для которых все полиномы устойчивы. Здесь же предлагается определять границу области устойчивости в явном виде с помощью конструктивного D-разбиения, а потом проверять её пересечение с конечным набором границ областей K_i , иначе говоря, явно искать пересечение границ ячеек с границами области устойчивости.

Также предлагается использовать ячеистую сетку, в частности, образованную конечным набором линий, и явно обнаруживать множества ячеек K_i , содержащие границу, по пересечениям

границ области устойчивости с линиями, образующими сетку. Эти линии одновременно являются границами ячеек. При этом ясно, что остальные ячейки либо целиком принадлежат какой-то области D-разбиения, в том числе области устойчивости, либо содержат какую-то область D-разбиения целиком³. В этом смысле подход близок к идеям адаптивного разбиения плоскости параметров: квадратичному дереву (*quadtree*) [14] или с учётом разбиения на пересекающиеся участки [11]. Предложенный метод отличается от методов слежения вдоль границы множества с помощью пересекающих её отрезков (например, таких, как ортогональный, симплексный, гусеничный и пр. методы [2]), так как гарантируется обнаружение всех множеств, покрывающих границу.

3.1. Ортогональная линейная сетка

Напрашивается очевидная комбинация одномерных сеток по каждому из параметров k_1, k_2 из п. 2.2, которая приводит к анализу множества прямоугольников. Удобство такого подхода состоит в том, что достаточно проанализировать пересечение границы D-разбиения с вертикальными и гори-

зонтальными прямыми (см. формулы (28), (29)), а точки пересечения линий сетки между собой очевидны. При этом автоматически будут получены прямоугольники-ячейки K_i , содержащие границу области устойчивости, см. пример на рис. 3, а.

3.2. Полярная сетка

Аналогичный подход работает при комбинации лучевой и радиальной сетки. Полученная полярная сетка состоит из концентрических окружностей и радиальных прямых, проходящих через центр окружностей. Радиусы окружностей не обязательно должны быть равномерными, как и сетка по углам прямых. Если выбранный центр лежит внутри области устойчивости, полярная сетка позволяет найти простую секторную внутреннюю (и внешнюю) аппроксимацию области устойчивости.

3.3. Треугольная сетка

Приведём ещё один вид сетки, образованной прямыми, – треугольную сетку, которая требует в полтора раз больше вычислений, но порождает более «гладкую» аппроксимацию границы. Для

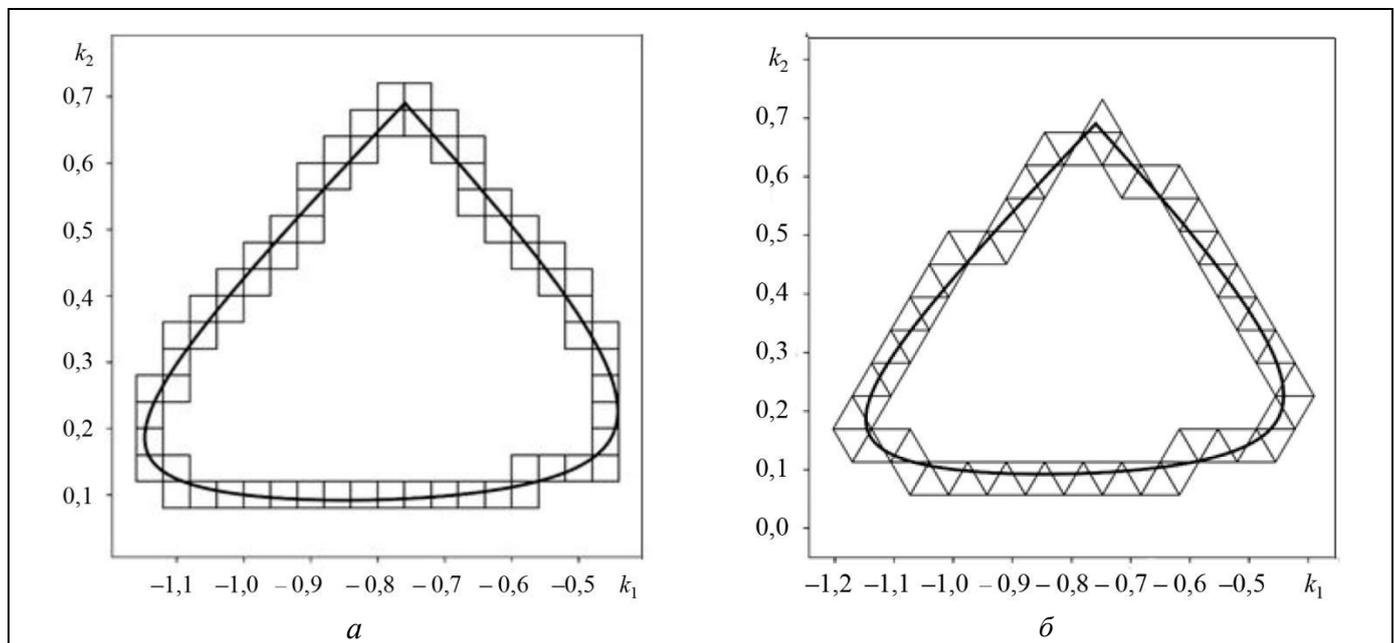


Рис. 3. Аппроксимация границы компоненты области устойчивости примера 2: а – ортогональной сеткой, б – треугольной сеткой

³ Собственно, это и есть основной недостаток традиционного «сеточного» подхода – он не обнаруживает явно множества, лежащие внутри ячейки K_i . Это принципиальная особенность, которую нельзя обойти напрямую, уменьшая размер ячеек. Однако, поскольку в указанных предположениях верна теорема 1 из первой части статьи [5], то граница области устойчивости состоит из конечного числа участков, и все они явно перечислены. Таким образом, можно проверить локализацию каждого из этих участков, см. п. 4.3, и проверить их положение относительно ячеек K_i .



этого берётся ряд горизонтальных прямых и ряд наклонных, см. пример для равномерных треугольников на рис. 3, б. Треугольная сетка удобна тем, что, граница накрывающего границу D-разбиения множества более «равномерная» и «гладкая». Также она состоит из отрезков кратной длины (с углами 60° , 120° , и 240°), если треугольники равносторонние.

4. ПРИМЕНЕНИЕ КОНСТРУКТИВНОГО D-РАЗБИЕНИЯ

Без ограничения общности результаты в этом параграфе также рассматриваются относительно области устойчивости D_n , а не произвольной области D-разбиения D_d .

4.1. Области устойчивости на кривой

Конструктивное D-разбиение позволяет найти участки устойчивости вдоль кривой в пространстве параметров. Пусть на плоскости параметров кривая K задана в явном виде как

$$F(k_1, k_2) = 0. \quad (29)$$

Её точки пересечения с границами областей D-разбиения разбивают её на участки, на которых число устойчивых корней постоянно. В частности, для нахождения «участков устойчивости», соответствующих устойчивым полиномам, достаточно рассмотреть пересечение кривой с границами области устойчивости $k_\ell(w)$. Эти точки удовлетворяют уравнениям

$$F(k_{\ell,1}(w), k_{\ell,2}(w)) = 0, \quad w \in W_\ell, \quad (30)$$

для пересечения с границами – участками основной кривой $k_\ell(w)$ и уравнениям

$$F(d_1t + p_1, d_2t + p_2) = 0 \quad (31)$$

для пересечения с границами – особыми кривыми, заданными в параметрическом виде $td + p$.

Если кривая K алгебраическая, т. е. функция $F(k_1, k_2)$ является полиномом, то решение обоих уравнений сводится к вычислению корней некоторого полинома. Такой приём использовался в п. 2.4 для анализа пересечений границы D-разбиения с окружностями с помощью уравнений (27) и (28) вместо использования записи окружности в форме (26). Также отметим, что уравнения (30), (31) можно рассматривать как граничные уравнения для кривой $k_\ell(w)$ [15].

Более того, для алгебраических кривых вида (29) можно получить параметризацию $K = \{k(v) : v \in \mathbb{R}\}$, а по точкам пересечения с границами D-разбиения можно восстановить параметры v_i [16]. Эти точки разбивают кривую на участки с постоянным числом корней, некоторые из которых соответствуют участкам устойчивости. Также при известной параметризации кривой $K = \{k(v) : v \in \mathbb{R}\}$ можно сразу подставить уравнение кривой в основное уравнение D-разбиения $G(s, k_1(v), k_2(v)) = 0$, $s \in \partial D$, и выполнить одномерное D-разбиение по параметру v . Если параметризация кривой дробно-рациональная, то его можно свести к одномерному D-разбиению для полинома вида (20).

Предложенный подход позволяет анализировать полиномы, коэффициенты которых зависят от одного параметра нелинейно, например $G(s, t) = f(t)P(s) + g(t)Q(s) + R(s)$. Сначала находится конструктивное D-разбиение полинома $G(s, f, g) = fP(s) + gQ(s) + R(s)$ относительно f и g как свободных параметров. Затем определяется параметризация кривой $(f(t), g(t))^T$ в виде (29) и решаются уравнения вида (30) и (31). Аналогичный приём может помочь анализировать полиномы с несколькими нелинейно входящими в коэффициенты параметрами.

4.2. Пересечение с множествами дополнительных ограничений

Описание области устойчивости в виде набора участков границы, полученное в результате конструктивного D-разбиения, позволяет легко учитывать дополнительные ограничения на коэффициенты k_1, k_2 . Пусть эти ограничения заданы множествами K_i . В этом случае сперва определяются те части границы ∂K_i , которые лежат внутри области устойчивости, например, как указано в п. 4.1. Затем границы области устойчивости обновляются с учётом этих, новых границ, и с исключением участков границы области устойчивости вне множеств K_i . Эта процедура повторяется для всех областей дополнительных ограничений K_i . По сути, каждый раз повторяется отсечение основных кривых и особых прямых аналогично тому, как это сделано в локализованном D-разбиении.

4.3. Локализация областей D-разбиения на плоскости параметров и опорная функция

Термин «локализация» может быть использован не только для априорного ограничения множества интересующих параметров, но и для обратной цели – идентификации положения области устойчивости или её компонент на плоскости.

Имея описание множества (в частности, области устойчивости) в виде набора участков границ $k_i(w)$, $w \in W_i = [w_{i,a}, w_{i,b}]$, включающих как участки основных кривых, так и отрезки на конечных интервалах W_i , нетрудно описать его положение. Так, можно найти крайние точки, используя тот факт, что участки границ заданы дробно-рациональными кривыми. Например, определение минимальных и максимальных значений компонент, т. е. прямоугольника, содержащего участок границы $k_i(w)$, осуществляется по экстремальным точкам аналогично тому, как это сделано для скорости в п. 1.1. Для этого ищутся корни $w_{i,1,m}$, $m=1, \dots$, $w_{i,2,\ell}$, $\ell=1, \dots$, двух уравнений

$$\begin{aligned} w_{i,1,m} : k'_{i,1}(w) &= 0, \quad w \in [w_{i,a}, w_{i,b}], \\ w_{i,2,\ell} : k'_{i,2}(w) &= 0, \quad w \in [w_{i,a}, w_{i,b}]. \end{aligned}$$

Как и для оценки скорости (16), имеют место неравенства, причём крайние значения достигаются в концах участков кривых или в точках $w_{i,1,m}$, $w_{i,1,\ell}$, подозрительных на экстремум:

$$\begin{aligned} \min \{ k_{i,1}(w_{i,a}), k_{i,1}(w_{i,b}), \min_m k_{i,1}(w_{i,1,m}) \} &\leq k_{i,1}(w) \leq \\ &\leq \max \{ k_{i,1}(w_{i,a}), k_{i,1}(w_{i,b}), \max_m k_{i,1}(w_{i,1,m}) \}, \\ \min \{ k_{i,2}(w_{i,a}), k_{i,2}(w_{i,b}), \min_\ell k_{i,2}(w_{i,2,\ell}) \} &\leq k_{i,2}(w) \leq \\ &\leq \max \{ k_{i,2}(w_{i,a}), k_{i,2}(w_{i,b}), \max_\ell k_{i,2}(w_{i,2,\ell}) \}. \end{aligned}$$

Вместо экстремальных точек можно использовать только минимумы или максимумы соответственно для нижних и верхних границ.

Для отрезков интервалы определяются по его концам: например, для k_1

$$\begin{aligned} \min \{ k_{i,1}(w_{i,a}), k_{i,1}(w_{i,b}) \} &\leq \\ &\leq k_{i,1}(w) \leq \max \{ k_{i,1}(w_{i,a}), k_{i,1}(w_{i,b}) \}, \end{aligned}$$

аналогично для второй координаты $k_{i,2}(w)$.

Пример локализации отдельных компонент области устойчивости приведён на рис. 4.

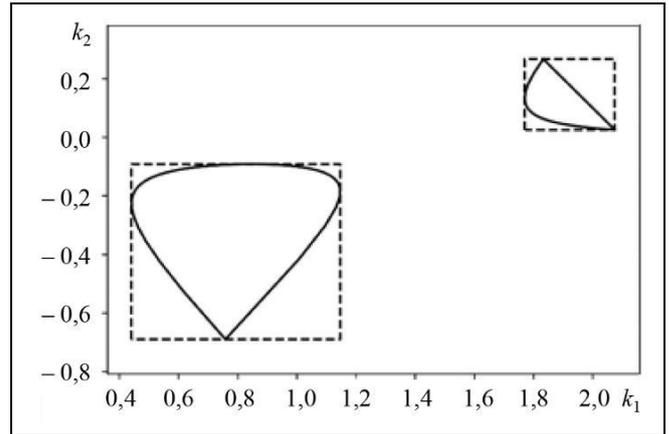


Рис. 4. Локализации двух компонент области устойчивости для примера 2, см. § 5

Схожим образом можно легко получить опорную функцию $\text{supp}_{D_n}(d) = \max_{k \in D_n} d^T k$ и её опорный элемент $\text{suppel}(d)$ для области устойчивости в направлении вектора $d \in \mathbb{R}^2$. Для этого используется опорная функция для каждого участка границы $k_i(w)$:

$$\begin{aligned} \text{supp}_{D_n}(d) &= \max_i \max_{w \in W_i} d_1 k_{i,1}(w) + d_2 k_{i,2}(w), \\ \text{suppel}(d) &= k_j(w_j), \\ w_j &= \arg \max_{w \in W_i} d_1 k_{i,1}(w) + d_2 k_{i,2}(w), \\ j &= \arg \max_i d_1 k_{i,1}(w_i) + d_2 k_{i,2}(w_i). \end{aligned}$$

Поиск максимума дробно-рациональной функции на интервале, на котором знаменатель не обращается в нуль, сводится к вычислению корней полинома в числителе производной. Более того, можно строить опорную функцию не только для всей области устойчивости целиком, а для её отдельной компоненты, взяв в рассмотрение только границы этой компоненты.

В общем случае для первичного анализа области устойчивости предлагается трёхэтапный подход.

1. Получить конструктивное описание границы области устойчивости.

2. Выбрать достаточно большое множество локализации \mathbf{K} , чтобы получить конечные и замкнутые интервалы параметризации W_i , применив также результаты из п. 3.2 по замене параметризации.

3. Локализовать область устойчивости по участкам границ внутри множества \mathbf{K} .

Для начальной локализации с большим множеством \mathbf{K} используются уравнения (8), (10), (11)



для прямоугольника, (27), (28) – для K в виде круга и т. д.

Кроме того, локализацию области устойчивости можно осуществить численно, с использованием достаточно равномерной сетки мелкости ρ . Пусть, согласно п. 1.1, такая сетка получена в результате применения алгоритма 1 или 2 ко всем участкам границы области устойчивости: $K_{grid} = \{k_r, r = 1, \dots\}$. Последовательным соединением узлов сетки будет получен многоугольник, приближающий область устойчивости с точностью не хуже ρ (параметр сетки). Для этого многоугольника также можно получить явное выражение для опорной функции и опорного элемента по его вершинам:

$$\text{supp}_{D_n}(d) = \max_r d^T k_r,$$

$$\text{suppel}_{D_n}(d) = k_r, \quad r = \arg \max_r d^T k_r.$$

Приведённые выражения являются приближением к опорной функции и элементу, а сама область устойчивости гарантированно лежит в прямоугольнике

$$\begin{aligned} & \left[-\rho + \min_r k_{r,1}, \max_r k_{r,1} + \rho \right] \times \\ & \times \left[-\rho + \min_r k_{r,2}, \max_r k_{r,2} + \rho \right], \end{aligned}$$

где ρ – мелкость достаточно равномерной сетки.

Аналогичным образом для набора точек можно вычислить минимальный покрывающий круг с центром k_0 и радиусом R . Круг с таким же центром и радиусом $R + \rho$ будет содержать область устойчивости.

4.4. Расстояние до области устойчивости (области неустойчивости)

Для точки k_0 , характеризующей устойчивый полином $G(s, k_{0,1}, k_{0,2})$, естественно ставится задача о расстоянии до ближайшей неустойчивой точки, т. е. задача определения круга максимального радиуса $\|k - k_0\| \leq R$, целиком лежащего в области устойчивости. Этот радиус называется радиусом устойчивости [3].

Имея конструктивное D-разбиение, радиус устойчивости можно вычислить точно или получить его оценку по достаточно равномерной сетке на границе. Точное решение определяется задачей минимизации, учитывающей все участки границы области устойчивости $k_i(w)$:

$$R = \min_i \min_{w \in W_i} \|k_i(w) - k_0\|.$$

Эта задача распадается на набор подзадач поиска минимумов дробно-рациональных функций на интервале $\min_{w \in W_i} \|k_i(w) - k_0\|^2$. В свою очередь, каждая из подзадач сводится к вычислению корней некоторого полинома аналогично тому, как это сделано для параметрической скорости (15), и к проверке концов участков.

Подобным образом, решая задачу максимизации расстояния от точки k_0 до точек границы, можно найти минимальный круг, содержащий область устойчивости, см. рис. 5.

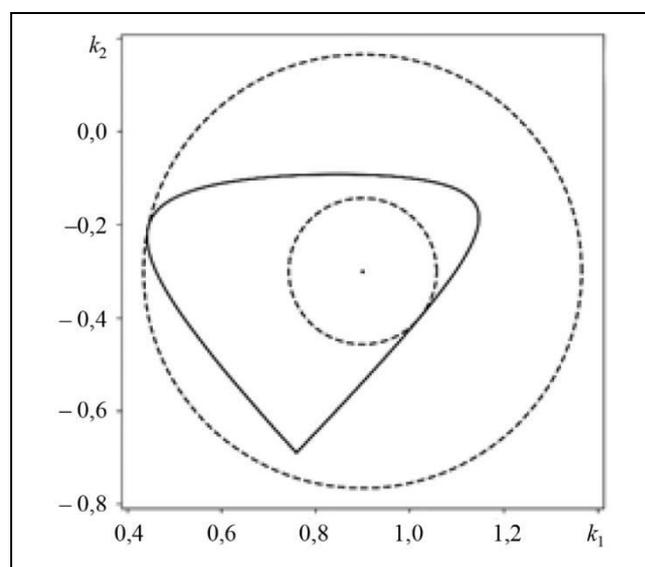


Рис. 5. Круг максимального радиуса с заданным центром, содержащийся в области устойчивости, и круг минимального радиуса, содержащий компоненту области устойчивости, для примера 2, см. § 5

Расстояние до границы области устойчивости можно оценить и по точкам достаточно равномерной сетки на границе как

$$\min_i \min_r \|k_r - k_0\| - \rho \leq R \leq \min_i \min_r \|k_r - k_0\| + \rho.$$

Здесь k_r – узлы сетки, покрывающей границу области устойчивости. Она состоит из объединения (без повторов) узлов сеток каждого из участков границы $k_i(w)$, $w \in W_i$.

4.5. Применение к робастному D-разбиению

В задачах робастного анализа рассматриваются полиномы, зависящие не только от «управляющих» параметров, но и от неопределённых параметров, обозначим их как q :

$$G(s, k_1, k_2, q) = k_1 P(s) + k_2 Q(s) + R_0(s) + \sum_i q_i R_i(s), \quad q \in Q. \quad (32)$$

Здесь неопределённость в задании полинома характеризуется ограниченным множеством Q . Таким образом, для каждого фиксированного k рассматривается семейство полиномов. Полином, для которого выбрано конкретное значение параметров $q = q_0$, называется номинальным. В этом случае полином (32) принимает форму (1) с $R(s) = R_0(s) + \sum_i q_{0,i} R_i(s)$.

Задача состоит в определении областей робастной устойчивости, т. е. таких k , что все корни полинома устойчивы для всех параметров $q \in Q$. Его построение, будучи идейно таким же, как и построение D-разбиения, существенно сложнее из-за того, что области D-разбиения с постоянным (для любого $q \in Q$) числом устойчивых корней разделяются не одномерными линиями, а двумя множествами. Эти множества определяются принципом исключения нуля, обобщающим уравнения (3) и (4):

$$K_{bnd} = \{k_1, k_2 : G(s, k_1, k_2, q) = 0_{\mathbb{C}}, s \in \partial D, q \in Q\}$$

и

$$K_{deg} = \{G_n(k_1, k_2, q) = 0_{\mathbb{C}}, q \in Q\}.$$

Соответственно, границами областей робастного D-разбиения являются границы этих двух множеств. Границы множества робастной устойчивости всегда лежат внутри множества устойчивости номинального полинома. Это следует из того, что каждый полином семейства (32) должен быть устойчивым, включая номинальный полином $G(s, k_1, k_2, q_0)$. Таким образом, при описании границ множества робастной устойчивости можно ограничиться только теми границами множеств K_{bnd} и K_{deg} , которые лежат внутри области устойчивости номинального полинома.

Кроме того, можно многократно повторять конструктивное D-разбиение для полиномов $G(s, k_1, k_2, q_r)$, $r = 1, 2, \dots$, где каждый полином определяется случайно выбранными параметрами $q_r \in Q$. Множество робастной устойчивости лежит в пересечении областей устойчивости всех выбранных полиномов.

Пусть граница множества локализации корней ∂D имеет параметризацию $s(w)$, и номинальный полином имеет границу, параметризованную набором кривых и прямых $k_i(w)$, $w \in W_i$. Оказывается,

что граница областей робастного D-разбиения характеризуется подобным образом: K_{bnd} является огибающей семейства множеств $K_{bnd}(w) = \{k_1, k_2 : G(s(w), k_1, k_2, q) = 0_{\mathbb{C}}, q \in Q, s(w) \in \Gamma\}$ [3, 17].

Можно было бы ожидать, что интервалы по w множеств $K_{bnd}(w)$, порождающих границу области робастной устойчивости, входят в интервалы W_i границ области устойчивости без неопределённых параметров. К сожалению, в общем случае гипотеза о вложенности интервалов неверна, так как граница множества $K_{bnd}(w)$ при $w \notin W_i$ может влиять на итоговое множество робастной устойчивости. Это зависит от значений функции основной кривой $k(w)$ для номинального D-разбиения и размера множества Q . Исследование и построение конструктивного робастного D-разбиения является открытой проблемой даже для случая параметров, линейно входящих в характеристический полином.

5. ПРИМЕРЫ

На рисунках были использованы множества устойчивости или их компоненты из двух примеров, подробно разобранных в первой части статьи [5].

Пример 1 [18, с. 77]. Рассматривается замкнутая система непрерывного времени с характеристическим полиномом

$$G(s, k_1, k_2) = k_1 s(s-1)(s-2) + k_2 (s-1)(s-2) + s(s+1)(s^2 + s + 1). \quad (33)$$

Требуется исследовать его устойчивость относительно области локализации корней $D = \{s : \operatorname{Re} s < 0, 2\}$ с параметризацией границы $s(w) = -0,2 + jw$. Поскольку полином (33) имеет вещественные коэффициенты, достаточно взять верхнюю часть границы, $W = [0, \infty)$.

Основная кривая одна, $w \in [0, \infty)$, с компонентами

$$k_1(w) = \frac{4,6 w^4 - 7,112 w^2 + 0,62016}{-w^4 - 6,28 w^2 - 6,9696},$$

$$k_2(w) = \frac{-w^6 + 8,68 w^4 - 5,4208 w^2 - 0,230784}{-w^4 - 6,28 w^2 - 6,9696}.$$

Критической частоте $w_0 = 0$ соответствует единственная особая прямая, $-0,528k_1 + 2,64k_2 - 0,1344 = 0$, с параметризацией $p + td$, где $p = (-0,00979021; 0,04895105)^T$, $d = (-2,64; -0,528)^T$.



Область устойчивости ограничена одним участком основной кривой $k(w)$, $w \in [0; 0,70951628]$ и отрезком $p+td$, $t \in [-0,06916025; 0,02999640]$ (рис. 6).

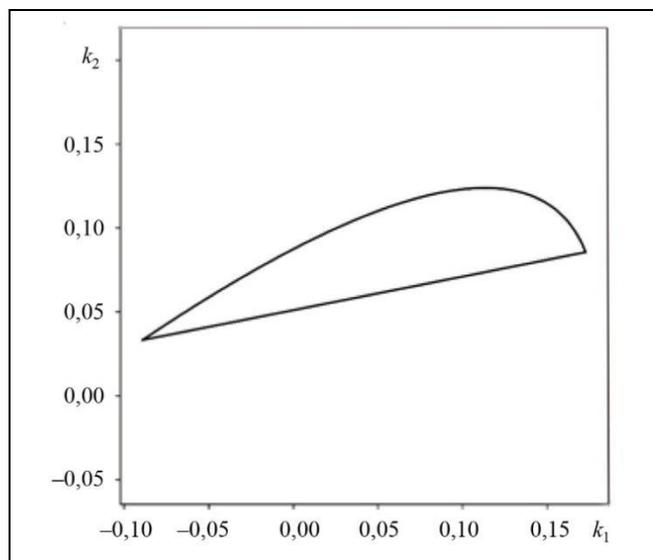


Рис. 6. Область устойчивости для примера 1

На рис. 2 изображён ряд внутренних аппроксимаций области устойчивости с помощью одномерных линий. На рис. 3, а изображены параметрическая скорость $v(w)$ указанного участка кривой и достаточно равномерные значения параметров, полученные с помощью алгоритма 1 для $\rho = 0,02$. На рис. 3, б приведены узловые точки и их окрестности радиуса ρ , покрывающие этот участок кривой.

Пример 2 [19]. Рассмотрен характеристический полином дискретной системы $G_0(z, k_1, k_2) = z^n + k_1 z^{n-1} + (1+\varepsilon)z^{n-2} + k_2$ при $n = 5$ и $\varepsilon = 0,1$. Устойчивость дискретного полинома эквивалентна гурвицевости полинома

$$G(s, k_1, k_2) = (s+1)^5 + (1+\varepsilon)(s-1)^2 (s+1)^3 + k_1 (s-1)(s+1)^4 + k_2 (s-1)^5.$$

Граница областей D-разбиения состоит из одной основной кривой

$$k_1(w) = \frac{-16,6 w^8 + 128,8 w^6 - 221,2 w^4 + 128,8 w^2 - 16,6}{8(w^8 - 6 w^6 + 6 w^2 - 1)},$$

$$k_2(w) = \frac{-0,2 w^8 - 0,8 w^6 - 1,2 w^4 - 0,8 w^2 - 0,2}{8(w^8 - 6 w^6 + 6 w^2 - 1)}$$

и двух особых прямых. Одна из них соответствует $w=0$ и имеет параметризацию $td+p$, где $p = (1,05; 1,05)$, $d = (1,-1)$. Вторая особая прямая опре-

деляется условием понижения степени (4) с параметризацией $td+p$, где $p = (-1,05; -1,05)$, $d = (-1; 1)$.

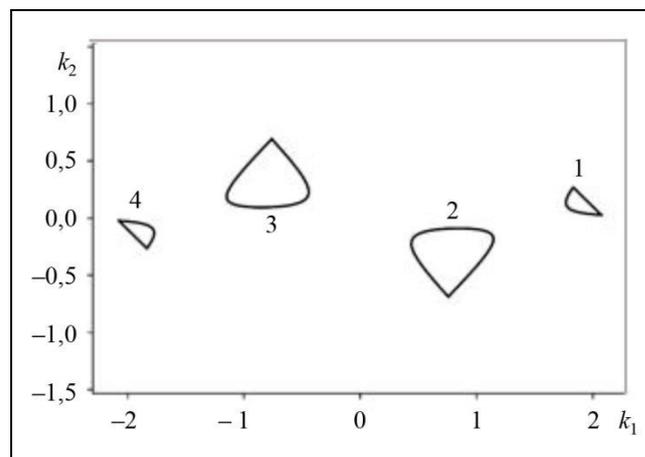


Рис. 7. Компоненты области устойчивости для примера 2

Область устойчивости состоит из четырёх компонент, см. рис. 7:

1) из участка первой особой прямой при $t \in [1,025; 1,45]$ и участка основной кривой при $w \in [0; 0,37796447]$,

2) участка основной кривой при $w \in [0,42972375; 0,96431209]$,

3) участка основной кривой при $w \in [1,03700867; 2,32707640]$,

4) участка второй особой прямой при $t \in [1,025; 1,45]$ и участка основной кривой при $w \in [2,64575131; \infty)$.

Последний участок кривой можно записать как $k_{u,2}(u) = k_\ell(1/u)$, $u \in [0; 0,37796447]$, согласно п. 3.2, при этом значение в $u=0$ определено и совпадает с $k_{(\infty)}$:

$$k_{u,1}(u) = \frac{16,6 u^8 - 128,8 u^6 + 221,2 u^4 - 128,8 u^2 + 16,6}{8(u^8 - 6 u^6 + 6 u^2 - 1)},$$

$$k_{u,2}(u) = \frac{0,2 u^8 + 0,8 u^6 + 1,2 u^4 + 0,8 u^2 + 0,2}{8(u^8 - 6 u^6 + 6 u^2 - 1)}.$$

Отметим, что полученная параметризация совпадает с исходной параметризацией по w с точностью до знака, а интервал совпадает с интервалом первого участка. Это связано с симметрией исходного множества локализации корней дискретной системы в виде единичного круга и его параметризации.

На рис. 3 изображены аппроксимации второй компоненты области устойчивости с помощью ортогональной и треугольной сеток. На рис. 4 представлены огра-

ничающие области $[1,7706; 2,075] \times [0,025; 0,2667]$ для первой компоненты и $[0,4414; 1,1472] \times [-0,6901; -0,09185]$ для второй компоненты области устойчивости, третья и четвёртая компоненты симметричны второй и первой. На рис. 5 изображены круги с центрами $k_1 = 0,9$, $k_2 = -0,3$, меньший из которых полностью лежит в области устойчивости и определяет радиус устойчивости 0,4417 для указанного регулятора. Большой круг радиуса 0,9538 целиком содержит вторую компоненту области устойчивости.

ЗАКЛЮЧЕНИЕ

На основе метода конструктивного D-разбиения, развитого в первой части статьи [5], предложены два алгоритма построения достаточно равномерной сетки на границе области устойчивости с заданной мелкостью. Алгоритмы используют оценку параметрической скорости дробно-рациональных кривых. Предложен ряд полусеточных внутренних аппроксимаций области устойчивости, а также способ нахождения регулярного покрытия границы области устойчивости прямоугольниками и треугольниками. Указано применение предложенных методов для ряда задач исследования области устойчивости, включающего вопросы локализации компонент области устойчивости на плоскости, построение их опорных функций и их приближений, а также вычисление радиуса устойчивости.

Благодарности. Автор выражает благодарность анонимным рецензентам и корректору, чьи замечания позволили улучшить статью.

ЛИТЕРАТУРА

1. *Kushel, O.Y.* Unifying Matrix Stability Concepts with a View to Applications // *SIAM Review*. – 2019. – Vol. 61, no. 4. – P. 643–729.
2. *Дидук Г.А.* Машинные методы исследования автоматических систем. – Л.: Энергоатомиздат. Ленингр. отд-ние, 1983. – 176 с. [*Diduk, G.A.* Machine methods for studying automatic systems. – Leningrad: Energoatomizdat, 1983. – 176 s. (In Russian)]
3. *Ackermann, J.* Robust control: The parameter space approach. (2nd ed.). – London: Springer, 2002. – 483 p.
4. *Violet, G.* Continuity argument revisited: geometry of root clustering via symmetric products / arXiv:1512.08645. – 2016. – P. 1–45. – DOI: <https://doi.org/10.48550/arXiv.1512.08645>
5. *Тремба А.А.* Конструктивное D-разбиение для двух линейно входящих в полином параметров. Ч. 1. Описание границ областей D-разбиения. – 2025. – № 6. – С. 39–57. [*Tremba, A.A.* Constructive D-Partition for Two Parameters Entering Polynomial Linearly: Part I. Description of the Boundaries of the D-Partition Regions // *Control Sciences*. – 2025. – No. 6. – P. 33–49.]
6. *Грязина Е.Н., Поляк Б.Т., Тремба А.А.* Современное состояние метода D-разбиения // *Автоматика и телемеханика*. – 2008. – № 12. – С. 3–40. [*Gryazina, E.N., Polyak, B.T., Tremba, A.A.* D-decomposition technique state-of-the-art // *Automation and Remote Control*. – 2008. – Vol. 69, no. 12. – P. 1991–2026.]
7. *Неймарк Ю.И.* Об определении значений параметров, при которых система автоматического регулирования устойчива // *Автоматика и телемеханика*. – 1948. – Т. 9, № 3. – С. 190–203. [*Neimark, Yu.I.* On the determination of the values of the parameters at which the automatic control system is stable // *Automation and Remote Control*. – 1948. – Vol. 9, no. 3. – P. 190–203. (In Russian)]
8. *Неймарк Ю.И.* Динамические системы и управляемые процессы. – М.: Наука, 1978. – 336 с. [*Neimark, Yu.I.* Dinamicheskie sistemy i upravlyаемые processy. – Moscow: Nauka, 1978. – 336 s. (In Russian)]
9. *Gryazina, E.N., Polyak, B.T.* Stability regions in the parameter space: D-decomposition revisited // *Automatica*. – 2006. – Vol. 42, no. 1. – P. 13–26.
10. *Siljak, D.D.* Analysis and synthesis of feedback control systems in the parameter plane, I-linear continuous systems // *IEEE Transactions on Applications and Industry*. – 1964. – Vol. 83, no. 75. – P. 449–458.
11. *Tan, L., Li, B., Zhang, B., Cheng, J.S.* An algorithm for the intersection problem of planar parametric curves // *Proceedings of International Workshop on Computer Algebra in Scientific Computing*. – Cham: Springer Nature Switzerland, 2023. – P. 312–329.
12. *Polyak, B.T., Gryazina, E.N.* Randomized methods based on new Monte Carlo schemes for control and optimization // *Annals of Operations Research*. – 2011. – Vol. 189. – P. 343–356.
13. *Прышников П.Ф.* D-разбиение при полиномиальной зависимости коэффициентов многочлена от двух параметров // *Автоматика и телемеханика*. – 2021. – № 3. – С. 32–46. [*Pryashnikova, P.F.* D-decomposition in the case of polynomial dependence of the coefficients of a polynomial on two parameters // *Automation and Remote Control*. – 2021. – Vol. 82, no. 3. – P. 398–409.]
14. *Farouki, R.T.* Reduced difference polynomials and self-intersection computations // *Applied Mathematics and Computation*. – 2018. – Vol. 324. – P. 174–190.
15. *Saydy, L., Tits, A.L., Abed, E.H.* Guardian maps and the generalized stability of parametrized families of matrices and polynomials // *Mathematics of Control, Signals and Systems*. – 1990. – Vol. 3, no. 4. – P. 345–371.
16. *Goldman, R.N., Sederberg, T.W., Anderson, D.C.* Vector elimination: A technique for the implicitization, inversion, and intersection of planar parametric rational polynomial curves // *Computer Aided Geometric Design*. – 1984. – Vol. 1, no. 4. – P. 327–356.
17. *Тремба А.А.* Робастное D-разбиение при l_p -ограниченных параметрических неопределенностях // *Автоматика и телемеханика*. – 2006. – № 12. – С. 21–35. [*Tremba, A.A.* Robust D-decomposition under l_p -bounded parametric uncertainties // *Automation and Remote Control*. – 2006. – Vol. 67, no. 12. – P. 1878–1892.]



18. *Francis, B.* A Course in H_∞ Control Theory. – Berlin: Springer-Verlag, 1987. – Lecture Notes in Control and Information Sciences. – Vol. 88. – 150 p.
19. *Грязина Е.Н.* К теории D-разбиения // Автоматика и телемеханика. – 2004. – № 12. – С. 15–28. [*Gryazina, E.N.* The D-Decomposition Theory // Automation and Remote Control. – 2004. – Vol. 65, no. 12. – P. 1872–1884.]

Статья представлена к публикации членом редколлегии
С. А. Красновой.

Поступила в редакцию 06.08.2025,
после доработки 08.10.2025.
Принята к публикации 23.10.2025.

Тремба Андрей Александрович – канд. физ.-мат. наук,
Институт проблем управления РАН, г. Москва; Московский
физико-технический институт, г. Долгопрудный.
✉ atremba@ipu.ru
ORCID iD: <https://orcid.org/0000-0001-5783-7600>

© 2026 г. Тремба А. А.



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

CONSTRUCTIVE D-PARTITION FOR TWO PARAMETERS ENTERING A POLYNOMIAL LINEARLY. PART II: Approximation of Stability Regions and Robustness Analysis

A. A. Tremba

Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia
Moscow Institute of Physics and Technology, Dolgoprudny, Russia

✉ atremba@ipu.ru

Abstract. For a polynomial linearly dependent on two parameters, several methods are proposed to approximate its stability region with respect to a given root localization region (also called a root clustering set in the literature). The first method is to apply a sufficiently uniform grid to the stability region boundary that ensures its complete coverage with a given accuracy. The second (semi-grid) method yields an internal approximation of the stability region using line segments or curve arcs bounded by the stability region. The third method is to cover the stability region boundary with simple sets (cells) in order to obtain piecewise linear internal and external approximations of the stability region. All methods are based on the constructive D-partition (constructive D-decomposition) method, which describes the stability region boundary as a set of line segments and rational curve arcs. The exact stability radius and its simple estimate are derived in the parameter plane. Implementation of all methods and algorithms is reduced to finding the real roots of polynomials.

Keywords: constructive D-partition, rational curves, approximation of the stability region, sufficiently uniform grid, grid methods, semi-grid methods, support function, stability radius.

Acknowledgments. The research presented in Sections 2 and 3 was supported by the Russian Science Foundation, project no. 21-71-30005-II, <https://rscf.ru/project/21-71-30005/>. The author is grateful to the anonymous reviewers and the editorial proof-reader for careful reading of the manuscript and helpful remarks.

ФУНКЦИЯ ПОБЕДЫ И ЕЕ ПРИМЕНЕНИЕ ПРИ МОДЕЛИРОВАНИИ КОНФЛИКТОВ

В. В. Шумов

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

✉ v.v.shumov@yandex.ru

Аннотация. Функция победы в боевых и специальных действиях является отдельным классом функций технологий конфликта, описывающих вероятность победы участника в конкурсе или конфликте. Дана краткая характеристика агрегированных функций, к которым относят производственные функции, функции полезности, функции технологии конфликта, индексы мощи и могущества государств и т. д. Сформулированы содержательные (постулаты) и формальные (аксиомы, свойства) требования к функции победы. Представлено теоретико-вероятностное (на основе закона поражения целей А. Н. Колмогорова и распределения Вейбулла) и содержательное (положения военной науки и практики) обоснование функции победы. Приведен обзор экономических моделей конфликта и присвоения. С использованием модели безопасности и функции победы выполнена постановка трех задач распределения спорного ресурса (территории и населения) между странами. Перспективным направлением исследований следует считать развитие теории конфликта на нескольких уровнях – на театре военных действий, межстрановом и геополитическом.

Ключевые слова: математическая модель, безопасность, боевые и специальные действия, функция победы, постулаты технологии конфликта, модели конфликта и присвоения.

ВВЕДЕНИЕ

Во многих сферах деятельности применяются *агрегированные функции* – математические зависимости ожидаемого результата (возможностей) от ряда факторов (усилий, ресурсов, технологий и т. д.). Агрегированная функция ставит в соответствие набору количественных показателей (видов ресурса) действительное число, характеризующее не наблюдаемый результат, а ожидаемые возможности.

В информатике и статистике под агрегированием понимается переход к сумме (среднему, минимуму, максимуму, медиане и т. д.). В системном анализе агрегирование (объединение нескольких элементов в единое целое) тесно связано с эмерджентностью (наличием у системы свойств, не присущих ее компонентам по отдельности) [1] и неопределенностью (категория неопределенности может быть достаточной для учета факторов сложности и эмерджентности [2]).

К наиболее исследованным агрегированным функциям можно отнести: производственные функции [3], функции полезности [4–6], функции технологии конфликта (см. обзор [7]), индексы (модели) геополитической мощи и могущества государств (см. статью [8] и обзор в монографии [9]) и др. Несмотря на критику агрегированных производственных функций («кембриджские споры» [10, 11]), они остаются основой прикладных статистических исследований и теорий экономического роста.

Со второй половины XX в. предметная область традиционной экономики расширилась: помимо производства и торговли стали исследоваться задачи присвоения (захват чужой продукции или защита своей) [12]. В теории конфликта функции технологии конфликта (*contest success functions*), описывающие вероятность победы участника в конкурсе (конflikте), являются аналогом производственных функций в теории производства и функций полезности в теории потребления. Обычно конфликт считался результатом неполной и



асимметричной информации или даже результатом иррациональности. Показано, что боевые действия изменяют стратегические позиции противников в долгосрочной перспективе. В этом случае сторона, учитывающая краткосрочные и долгосрочные эффекты, может выбрать войну и этот выбор будет соответствовать рациональному и дальновидному поведению, основанному на полной информации. Если одна или несколько сторон считают будущее достаточно важным, следует ожидать начала войны [7].

Цель настоящей работы – исследование функции победы в боевых и специальных действиях как отдельного класса функций технологии конфликта, постановка задач по ее применению в моделях конфликтов и присвоения.

1. ТРЕБОВАНИЯ К ФУНКЦИЯМ ТЕХНОЛОГИИ КОНФЛИКТА

Требования к агрегированным функциям технологии конфликта (победы в боевых и специальных действиях) разделим на *содержательные* (вытекающие из положений военной науки и системного анализа) и *формальные* (требования аксиоматического характера).

1.1. Содержательные требования

На рис. 1 показаны работы, оказавшие существенное влияние на развитие военной науки и практики.

Трактат Я. Бернулли «Искусство предположений» [13] явился первым систематическим изложением теории вероятностей. В нем представлены комбинаторика, биномиальное распределение, первый вариант закона больших чисел.

«Классический этап» развития теории вероятностей был завершен работой П. С. Лапласа «Аналитическая теория вероятностей» [14]. В ней исследованы дискретные и непрерывные случайные величины¹, дано понятие плотности распределения вероятности и характеристической функции, при-

ведена формула полной вероятности, доказана сходимость биномиального распределения к нормальному (теорема Муавра – Лапласа) при увеличении числа испытаний. Как часто случалось и в более поздние времена (XX в., успехи кибернетики и разочарование в ней, см. книгу [15]), в XIX в. число работ по теории вероятностей продолжало расти и предпринимались компрометирующие науку попытки распространить ее методы далеко за разумные пределы: на мораль, психологию, правоприменение, богословие и т. д.

К числу классиков европейской и мировой науки Нового времени, бесспорно, принадлежит Карл Клаузевиц. Его фундаментальный труд «О войне» [16] оказал глубокое влияние на развитие военной науки и не утратил своей актуальности в настоящее время. Автором не преследуется цель изложить творческое наследие К. Клаузевица, это дело военных исследователей и практиков. Следуя традиции (см. работы [17, 18]), сравним взгляды К. Клаузевица и Л. Н. Толстого [19] на роль неопределенности в войне и бою.

Следуя работам [20, 2], *неопределенностью* военной (боевой) деятельности назовем возможность наступления в ходе этой деятельности некоторых событий, влияющих на реализацию деятельности и ее результат, но которые могут наступить, а могут и не наступить. Следствием неопределенности военной деятельности является невозможность априори предсказать характеристики результата военной (боевой) деятельности, момент его получения и усилия (ресурсы), которые будут для этого затрачены.

Измеримая неопределенность военной деятельности определяется как возможность наступления описываемых некоторыми закономерностями событий (которые могут наступить, а могут и не наступить). Для анализа таких событий могут применяться количественные методы (например, вероятностные (статистические)), основанные на предыдущих измерениях или фундаментальных законах (вместе с предположением о неизменности условий и закономерностей).



Рис. 1. Ключевые работы, оказавшие влияние на развитие военной науки и практики

¹ Понятия «случайная величина», «функция распределения» появились в XX в. в работах русской вероятностной школы.

Истинной неопределенностью военной деятельности назовем возможность наступления уникальных (или редко повторяющихся) событий, которые не объясняются известными закономерностями.

Принципиальным отличием истинной неопределенности от измеримой является то, что события первой из них возникают вследствие непознанных факторов (частый и важный, но частный случай – активный выбор индивида), в то время как события второй, хотя и непредсказуемы, но описываются известными закономерностями [2, с. 66].

К. Клаузевиц отмечает [16], что война – область недостоверного; три четверти того, на чем строится действие на войне, лежит в тумане неизвестности. Его способ устранения неопределенности – «тонкий, гибкий, пронизательный ум». Недостоверность известий и предположений приводит к тому, что воюющий в действительности сталкивается с совершенно иным положением дел, чем ожидал. Многие донесения, получаемые на войне, противоречат одно другому; ложных донесений еще больше, а основная масса их недостоверна. Другой способ устранения неопределенности на стратегическом уровне – решительное численное превосходство (полутора-двукратное) над противником в силах и средствах, учет морального фактора. Неопределенность уменьшается при переходе с тактического уровня на стратегический. Третий способ противодействия неопределенности – наличие резерва. Резерв есть средство не только противостоять непредвиденным начинаниям противника, но и исправлять никогда не поддающийся предвидению исход боя в случае неблагоприятного его оборота.

В бою велика роль морального фактора. На тактическом уровне моральные силы постепенно восстанавливаются. Моральное влияние победы растет в увеличивающейся прогрессии в зависимости от масштаба боевых действий. И это может также свидетельствовать о неопределенности меньшего порядка на стратегическом уровне. Другое условие, определяющее размер морального веса победы, – это численное соотношение сил, сражавшихся друг против друга. Разбить противника малыми силами – это свидетельство общего превосходства над ним. Однако в действительности реальное соотношение сил на тактическом уровне обычно никогда не известно (за редким исключением, когда поле боя выделено), а на стратегическом уровне оно становится известным спустя многие годы, поэтому такое знание о моральном соотношении сторон никак не влияет на текущие события.

В заключение К. Клаузевиц пишет, что храбрость и дух войска во все времена повышали фи-

зические силы, так будет и впредь. Вместе с тем, в истории встречаются периоды, когда резкое превосходство в устройстве и вооружении войск давало значительный моральный перевес; в другие периоды превосходство давала подвижность войск²; далее оказывали влияние вновь вводимые системы тактики, затем военное искусство увлекалось стремлением к искусному использованию местности и т. д. Армии в наши дни настолько стали схожи между собой и вооружением, и снаряжением, и обучением, что между лучшими из них и худшими заметного отличия в этом отношении не существует. Степень подготовки научных сил еще представляет существенные различия, но приводит лишь к тому, что одни являются инициаторами и изобретателями тех или иных усовершенствований, а другие – их быстрыми подражателями. Чем больше будет равновесие во всех перечисленных выше факторах, тем более решительное влияние окажет численное соотношение сил.

Л. Н. Толстой рассматривает войны с более широких позиций (философских, цивилизационных, политических, культурных, этических и др.) и выделяет истинную неопределенность военной деятельности: «Наполеон, представляющийся нам руководителем всего этого движения (как диким представлялась фигура, вырезанная на носу корабля, силою, руководящею корабль), Наполеон во все это время своей деятельности был подобен ребенку, который, держась за тесемочки, привязанные внутри кареты, воображает, что он правит» [19]. Бесчисленное количество свободных сил (ибо нигде человек не бывает свободнее, как во время сражения, где дело идет о жизни и смерти) влияет на направление сражения и это направление никогда не может быть известно вперед и никогда не совпадает с направлением какой-нибудь силы. Историки постфактум пишут: *случай* сделал положение, *гений* воспользовался им. Но слова «случай» и «гений» не обозначают ничего действительно существующего и потому не могут быть определены. Офицеры – герои романа Л. Н. Толстого, – бродят по полю боя, ослепленные орудийным оружием дымом; не имея ни малейшего представления о происходящем. Истинная неопределенность войны объясняется автором романа тем, что человеческий разум не может постичь причины событий во всей их полноте, однако желание найти их заложено в человеческой душе.

² Превосходство в подвижности наблюдалось в 1930–1950-е гг. (переход армий к сплошной моторизации войск, созданию мотострелковых и механизированных частей и соединений), когда создавались моторизованные группы (дивизии), способные к действиям в глубине и глубокому охвату пехотных частей и соединений противника.



Сегодня для военных специалистов и исследователей идеи Л. Н. Толстого особенно актуальны и значимы, в том числе и в связи с внедрением систем искусственного интеллекта. Безусловно, искусственный интеллект будет применяться для решения простых и типовых задач (машинное зрение, обработка данных с датчиков в режиме времени, близком к реальному и др.). Т. Липский (военная академия США, Вест-Пойнт) отмечает: ввиду заинтересованности во внедрении искусственного интеллекта армия рискует забыть важность сохранения границы между человеком и машиной не только из этических соображений, но и потому что, как показал Л. Толстой³, в бою командир не может подчиняться алгоритмам и шаблонам. Командование остается «художественно-психологической» работой [18]. Искусственный интеллект плохо подходит для быстро меняющихся ситуаций, действия в которых в значительной степени зависят от контекста и нуждаются в человеческом суждении – так характеризуется любая война. Поэтому важно заложить в инструкции армейского руководства здоровый скептицизм по отношению к искусственному интеллекту. Военные должны научиться составлять планы и руководить их исполнением без посторонней помощи, прежде чем включать искусственный интеллект в этот процесс, и они должны регулярно проходить повторную сертификацию в этой области на протяжении всей военной карьеры [18]. В работе [21] показано, что машинное обучение не ведет к сильному искусственному интеллекту, а миф об искусственном интеллекте ослабляет веру в человеческий потенциал. О.П. Кузнецов отмечает, что понимание – это интерпретация в терминах картины мира человека; картину мира строит наш мозг, и она структурируется через категоризацию опыта человека; значения (смыслы) формируются раньше, чем формируются понятийные структуры; в основе значений лежат биологические и социальные цели; в когнитивных процессах участвует не только мозг, но и тело, а понимание связано с действиями в среде, знания о которой содержатся в картине мира.

Со второй половины XIX в. в России и европейских странах достижения естествознания и теории вероятностей активно внедряются в теорию

³ В романе «Война и мир» так описана роль М.И. Кутузова: «Долголетним военным опытом он знал и старческим умом понимал, что руководить сотнями тысяч человек, борющихся с смертью, нельзя одному человеку, и знал, что решают участь сраженья не распоряжения главнокомандующего, не место, на котором стоят войска, не количество пушек и убитых людей, а та неуловимая сила, называемая духом войска, и он следил за этой силой и руководил ею, насколько это было в его власти».

и практику стрельбы и управления огнем артиллерии [22]. В годы Первой мировой войны публикуются первые работы по моделированию военных и боевых действий (модели Осипова – Ланчестера [23, 24]).

На основе данных военной статистики о крупнейших сражениях регулярных армий М. П. Осипов разработал так называемую квадратичную модель боя, заложил основы теории боевых потенциалов, количественно оценил роль морального фактора и искусства полководца («победа зависит не от продолжительности боя, а главным образом от понесенных сторонами потерь; поэтому вернее будет считать, что бой длится до тех пор, пока потери одной из сторон не достигнут некоторого определенного процента. Таким процентом в среднем можно считать 20 %...» [23]).

Труд В. К. Триандафиллова «Характер операций современных армий» примечателен в двух отношениях: во-первых, в нем разработан новый раздел военного искусства – оперативное искусство (в дополнение к тактике и стратегии) и, во-вторых, названная работа дает все основания считать его автора основоположником системного анализа. В предисловии к первому изданию книги автор так излагает методологию своего исследования: сначала рассматривается материальная основа армейских операций – вооружение армий, их численность, организация и другие важнейшие данные обстановки, влияющие на характер боевых действий, а вслед за этим на основании всех этих данных исследуются вопросы современной тактики, отдельной операции и ряда последовательных операций. Все положения выражены на языке конкретных цифр, тактических и оперативных норм. Только таким путем можно показать различие между настоящим и прошлым, показать, куда растет военное дело. Причем весь цифровой материал имеет *ориентирующее значение*, для каждой конкретной обстановки нормы могут измениться. Установить характер и степень таких изменений – дело полководца, дело практического деятеля [25, с. 13, 14]. Вклад В.К. Триандафиллова в военную науку и системный анализ (раздел кибернетики) анализируется в монографии [9].

В 1930–1940-е гг. наблюдался редкий в истории случай, когда и теория игр (раздел исследования операций), и военная практика решали параллельно одну и ту же задачу – повышение эффективности деятельности путем исключения шаблона (применения так называемых «смешанных стратегий»). В 1944 г. опубликована хрестоматийная работа «Теория игр и экономическое поведение» по теории игр [5]. Краеугольным камнем современной теории некооперативных игр является концеп-

ция равновесия Нэша, существующая для всех конечных игр [26]. Обзор постановок теоретико-игровых задач представлен в работе [9].

Из всего наследия выдающегося советского полководца Г. К. Жукова отметим только два его результата (см. публикацию [27]). Во-первых, Г. К. Жуков выделил одни и те же факторы, влияющие на успех любого боя (сражения, операции), т. е. справедливые для всех уровней: тактического, оперативного и стратегического (оперативно-стратегического). Во-вторых, он начиная с 1943 г. руководствовался идеями «смешанных стратегий» при планировании стратегических наступательных операций (прорыв подготовленной обороны противника).

Исследование операций как прикладная математическая дисциплина возникло в годы Второй мировой войны. Его цель определялась как предоставление командующему количественных оснований для принятия решений. В последующем эта дисциплина стала одним из разделов кибернетики [28], которая ныне понимается как «наука об организации систем и управлении ими» [15, с. 7]. Сегодня чрезвычайно моден и продуктивен подход, называемый «сетевцентризмом» и включающий принципы организации и исследования любых сетей вообще и, в частности, «собирающихся» на время выполнения боевой задачи, в нужное время и в нужном месте [15, с. 75–78].

Из рассмотренного выше можно сделать следующие выводы (содержательные требования к функции победы).

- *Постулат Г. К. Жукова:* функция победы должна иметь одну и ту же форму для описания боевых действий на всех уровнях (тактическом, оперативном, оперативно-стратегическом), включая и специальные действия (партизанские, антипартизанские, подрывные, разведывательно-диверсионные, контртеррористические и др.).

- *Постулат К. Клаузевица и Л. Н. Толстого:* функция победы должна отражать как истинную, так и измеримую неопределенность.

- *Постулат М. П. Осипова и В. К. Триандафиллова:* функция победы должна позволять учитывать численности войск сторон, их моральные и технологические характеристики, особенности обстановки, существующие и перспективные системы вооружения.

1.2. Формальные требования

Исторически первая функция победы вытекает из модели боя М. П. Осипова. При отсутствии операционных потерь и резервов обычное сражение описывается системой дифференциальных уравнений

$$\frac{dx(t)}{dt} = -a_y y(t), \quad \frac{dy(t)}{dt} = -a_x x(t),$$

где $x(t)$ и $y(t)$ – численности войск первой и второй стороны в момент времени t ; a_x и a_y – их поражающие эффективности. Из условия равенства сил $y_0 = x_0 \sqrt{a_x / a_y}$ получим индикаторную функцию победы М. П. Осипова (вероятность победы первой стороны):

$$p_x(x, y) = \begin{cases} 1, & x\sqrt{a_x} > y\sqrt{a_y}, \\ 0,5, & x\sqrt{a_x} = y\sqrt{a_y}, \\ 0, & x\sqrt{a_x} < y\sqrt{a_y}. \end{cases} \quad (1)$$

Одна из моделей партизанской войны такова [29]:

$$\frac{dx(t)}{dt} = -a_y y(t), \quad \frac{dy(t)}{dt} = -a_x x(t) \frac{y(t)}{y_0},$$

где $x(t)$ – численность регулярных войск, а $y(t)$ – численность партизан. Из условия равенства сил $y_0 \sqrt{2a_y} = x_0 \sqrt{a_x}$ имеем функцию победы С. Дейчмана (следует из модели партизанской войны [29]):

$$p_x(x, y) = \begin{cases} 1, & x\sqrt{a_x} > y\sqrt{2a_y}, \\ 0,5, & x\sqrt{a_x} = y\sqrt{2a_y}, \\ 0, & x\sqrt{a_x} < y\sqrt{2a_y}. \end{cases}$$

То есть при прочих равных условиях для достижения паритета численность регулярных войск должна быть в $\sqrt{2} \approx 1,4$ раза выше численности партизан (ср. с выражением (1)).

Достаточно хорошо исследованным является следующий класс функций победы (успеха в конкурсе, аукционе):

$$p_x(x, y) = \frac{f_x(x)}{f_x(x) + f_y(y)}, \quad (2)$$

где $f_x(\cdot)$ и $f_y(\cdot)$ – неотрицательные строго возрастающие функции. Приведем некоторые наиболее часто встречающиеся функциональные формы модели (2) [7, 9].

Модель Г. Таллока

$$p_x(x, y) = \frac{x^\mu}{x^\mu + y^\mu} = \frac{(x/y)^\mu}{1 + (x/y)^\mu}, \quad (3)$$

где $0 < \mu \leq 1$ – параметр решительности сторон, относится к классу моделей на основе отношения сил сторон.



Модель Д. Макфаддена и Д. Хиршляйфера

$$p_x(x, y) = \frac{e^{\mu x}}{e^{\mu x} + e^{\mu y}} = \frac{1}{1 + e^{\mu(x-y)}} \quad (4)$$

относится к классу моделей на основе разности сил сторон. К этому же классу относится пробит-модель $p_x(x, y) = \Phi(x - y)$, где Φ – функция Лапласа.

Теоретико-вероятностное обоснование функций конфликта основано на анализе влияния неучитываемых факторов (случайных ошибок) на результат. Функции регрессии в общем случае имеют вид $Y_x = h(x, \varepsilon_x)$, $Y_y = h(y, \varepsilon_y)$, где функции ошибок ε_x и ε_y имеют равные нулю математические ожидания. Тогда вероятность победы первой стороны в конфликте равна

$$p_x(x, y) = P(Y_x > Y_y) = P(h(x, \varepsilon_x) > h(y, \varepsilon_y)).$$

Функции конфликта аксиоматизированы, в частности, Р. Люсом [30] и С. Скапердасом [31]. В основу аксиоматики положено *свойство независимости от посторонних альтернатив (Independence of Irrelevant Alternatives property)*: в контексте конфликта это свойство требует, чтобы исход конфликта между любыми двумя сторонами зависел только от количества оружия, которым владеют эти две стороны, а не от количества оружия, которым владеют третьи стороны. Следующее важное требование к функциям конфликта – их *однородность нулевой степени*, т. е. $p_x(tx, ty) = p_x(x, y)$ для всех $t > 0$. Модели (3) и (4) обладают *свойством симметрии или анонимности* в том смысле, что если усилия сторон поменять местами, то и вероятности их победы также поменяются местами.

С. Скапердас и др. отмечают, что несмотря на наличие значительного числа публикаций по моделированию конфликтов, конкурсов и аукционов, лишь в небольшом количестве публикаций затрагиваются вопросы верификации функций конфликта на реальных данных [32].

2. ТЕОРЕТИКО-ВЕРОЯТНОСТНОЕ И СОДЕРЖАТЕЛЬНОЕ ОБОСНОВАНИЕ ФУНКЦИИ ПОБЕДЫ

2.1. Теоретико-вероятностное обоснование функции победы

В 1945 г. А. Н. Колмогоров предложил критерий эффективности стрельбы, основанный на законе поражения целей – зависимости вероятности поражения одиночной или групповой цели (математического ожидания числа пораженных целей)

от количества выстрелов по ней. Вероятность поражения цели (событие A) при x попаданиях по ней вычисляется по формуле [33]

$$P(A|x) = 1 - e^{-\alpha x}, \quad (5)$$

где $\alpha > 0$ – параметр. Выражение (5) – экспоненциальное распределение (частный случай распределения Вейбулла) – имеет широкую область применения (надежность технических систем, системы массового обслуживания и др.) и тесно связано с понятием пуассоновского процесса, для которого промежутки времени между последовательными событиями суть независимые случайные величины, имеющие экспоненциальное распределение; при этом значение α равно среднему числу событий в единицу времени.

Обозначим X и Y случайные величины – количества попаданий, необходимых для поражения целей противника первой и второй сторонами. Используя распределение Вейбулла, определим вероятности решения сторонами поставленных задач (поражения целей противника):

$$F_x(x) = 1 - e^{-(\alpha_x x)^m}, \quad \alpha_x = \beta_x r_x,$$

$$F_y(y) = 1 - e^{-(\alpha_y y)^m}, \quad \alpha_y = \beta_y r_y,$$

где $m > 0$ – параметр масштаба боевых действий; r_x и r_y – количество боевых единиц в распоряжении первой и второй стороны соответственно; $\beta_x > 0$ и $\beta_y > 0$ – боевые эффективности единиц сторон.

Плотности распределения случайных величин X и Y равны

$$f_x(x) = \alpha_x m (\alpha_x x)^{m-1} e^{-(\alpha_x x)^m},$$

$$f_y(y) = \alpha_y m (\alpha_y y)^{m-1} e^{-(\alpha_y y)^m}.$$

Вероятность победы в боевых и специальных действиях первой стороны равна

$$\begin{aligned} P_x &= P(x < y) = 1 - P(x > y) = \\ &= 1 - \int_0^{\infty} f_x(x) \left[\int_0^x f_y(y) dy \right] dx \end{aligned} \quad (6)$$

(чем меньше попаданий требуется для поражения противника, тем эффективнее действуют боевые единицы).

Опуская промежуточные выкладки, из формулы (6) получим:

$$P_x = \frac{(\beta_x r_x)^m}{(\beta_x r_x)^m + (\beta_y r_y)^m}.$$

В статье [34] применен подобный подход, но в иной содержательной интерпретации параметров модели и без учета особенностей боевых действий.

Обозначив $x = r_x$, $y = r_y$, $\beta = \beta_x / \beta_y$, имеем функцию победы в боевых и специальных действиях:

$$p_x(x, y) = \frac{(\beta x)^m}{(\beta x)^m + (y)^m}, \beta = \varphi\rho, \quad (7)$$

где β – параметр боевого превосходства первой стороны над второй; φ – параметр морального превосходства; ρ – параметр технологического превосходства (превосходства в согласованности действий, разведке, огневом поражении и маневренности, см. работу [35]).

Пусть $q = \beta x/y$ – отношение сил сторон. Тогда

$$p_x = \frac{q^m}{q^m + 1} = 1 - s^{-m}, s^m = q^m + 1, s > 1, \quad (8)$$

т. е. получили распределение Парето. Случайная величина с функцией распределения (8) имеет плотность распределения и математическое ожидание

$$f(s) = ms^{-m-1}, s > 1, \\ M[S] = \int_1^{\infty} ms^{-m} ds = \left. \frac{ms^{1-m}}{1-m} \right|_1^{\infty}.$$

При $m \leq 1$ распределение (8) не имеет математического ожидания (оно равно бесконечности). Следовательно, функция победы (8) при $m \leq 1$ отражает истинную неопределенность, а при $m > 1$ – измеримую неопределенность.

Из формулы (8) найдем требуемое отношение сил для достижения победы с назначенной вероятностью:

$$q = \sqrt[m]{\frac{p_x}{1-p_x}}. \quad (9)$$

В статье [36] выполнена статистическая оценка параметра масштаба и по критерию хи-квадрат проверена гипотеза о соответствии модели (7) статистическим данным. На достаточно большом объеме статистических данных за XIX–XX вв. показано, что параметр m имеет следующие значения в следующих случаях:

- специальные действия – $m \approx 0,5$;
- бой (тактический уровень) – $m \approx 1$;
- боевые действия (оперативный уровень) – $m \approx 2$;
- военные действия (стратегический уровень) – $m \approx 3$.

В науках прикладного характера важнейшим является следующий вопрос: при какой вероятно-

сти событие можно считать практически достоверным?

Выделяются четыре уровня боеспособности соединений, частей и подразделений⁴: боеспособные (имеют не менее 75 % боеспособных организационных структур); ограниченно боеспособные (50–75 %); частично боеспособные (30–50 %); небоеспособные (менее 30 % боеспособных организационных структур).

Применительно к задачам подготовки боевых и специальных действий можно назначить следующие степени достоверности победы первой стороны:

- практически достоверная победа – вероятность победы равна 0,85–0,95;
- достаточная степень победы – вероятность равна 0,80–0,85;
- приемлемая степень победы – вероятность равна 0,7–0,8.

Степени достоверности назначаются с учетом обстановки. Назначение высоких степеней достоверности победы не всегда целесообразно, так как при этом требуется достаточно высокая концентрация войск и, следовательно, увеличивается риск группового поражения средствами противника.

Результаты расчетов по формулам (8) и (9) показывают, что с увеличением масштаба боевых действий неопределенность все больше смещается от истинной к измеримой. Так, на стратегическом уровне практически достоверная победа достигается при двукратном превосходстве над противником. Наибольшая неопределенность характерна для специальных действий, где даже для приемлемой степени победы над противником (партизанами, диверсионно-разведывательными и террористическими группами) требуется девятикратное превосходство.

2.2. Содержательное обоснование функции победы

По опыту Великой Отечественной войны боевой порядок советской стрелковой дивизии строился, как правило, в два эшелона, в полосе по фронту 8–12 км и в глубину 8–10 км. Полосы обороны армии и фронта имели относительно небольшую глубину и большой пространственный размах. То есть на оперативном и стратегическом уровнях в обороне для удержания полосы и нанесения контрудара по вторгшемуся противнику своевременно могла быть задействована относи-

⁴ Боевая способность. – URL: <https://encyclopedia.mil.ru/encyclopedia/dictionary/details.htm?id=3465@morfDictionary> (дата обращения: 10.08.2023).



тельно небольшая доля войск фронта (армии), см. работу [9, п. 3.2.3], что находит отражение в увеличении значения параметра масштаба m модели (7).

За прошедшие после Второй мировой войны годы сменилось несколько поколений вооружения, моторизация войск практически завершена, появились беспилотные (безэкипажные) системы, высокоточное оружие, автоматизированные системы управления войсками и оружием. Существенно выросли возможности дивизий и бригад по поражению противника в тактической и оперативной глубине.

В работе [9] оценено влияние современных систем вооружения (и, следовательно, новых тактических приемов и способов действий войск) на значение параметра масштаба функции победы на оперативном и стратегическом уровнях. В связи с увеличением дальности эффективной разведки противника и поражения его боевых единиц значения параметра масштаба на оперативном и стратегическом уровнях снизились и составляют $m \approx 1,5-2$ для оперативного уровня и $m \approx 2-3$ для стратегического уровня. Численная оценка параметра m для различных театров военных действий и условий обстановки представляется актуальной научной задачей.

3. ОБЗОР МОДЕЛЕЙ КОНФЛИКТА И ПРИСВОЕНИЯ

В традиционной экономической науке под присвоением понимается ненасильственный процесс, который гарантируется совершенными правами собственности и их беспрепятственным соблюдением. Экономика конфликта основана на модели соперничества между двумя и более игроками (агентами), каждый из которых выбирает между производством ресурсов (потребительских благ) и производством оружия (инструмента, предназначенного для присвоения ресурсов, произведенных другими игроками или совместно). С учетом известной проблемы безбилетника (англ. *free rider problem* – ситуации, когда отдельные потребители общественных благ не вносят свой вклад в их обеспечение, надеясь, что это за них сделают другие участники) установлено, что групповые структуры тем менее стабильны, чем больше в них участников (см. обзоры в работах [7, 37]).

В публикации [7] рассматривается среда, где действуют два одинаковых и нейтральных к риску агента (государства), оспаривающих R единиц ресурса, которые могут быть *потреблены напрямую*. Из-за несовершенства институтов управления и принуждения спор может быть разрешен конфликтом (угрозой конфликта). Целевые функции сто-

рон определяются следующим образом:

$$V_i(G_1, G_2) = p_i(G_1, G_2)R - G_i,$$

$$p_i(G_1, G_2) = \frac{G_i^\mu}{G_1^\mu + G_2^\mu}, \quad i = 1, 2,$$

где $0 < \mu \leq 1$ – параметр решительности; G_i – объем ресурса, потраченного i -й стороной на производство оружия.

Оптимальные значения целевых функций равны

$$V_i(G^*) = V^* = \frac{1-\mu/2}{2}R, \quad i = 1, 2.$$

Поскольку $\mu \leq 1$, то игрокам выгодно ассигновать средства на вооружение. Далее авторами работы [7] учитываются издержки конфликта и двухэтапные игры, в которых на первом этапе стороны одновременно и независимо выделяют ресурс для производства оружия. На втором этапе стороны вступают в переговоры о разделе спорного блага. В случае согласия на раздел это благо распределяется между сторонами. Иначе переговоры заканчиваются конфликтом, в котором победитель забирает все спорное благо.

А. Алесина и Е. Сполаоре исследовали проблему более общего характера – взаимосвязь между конфликтом и распределением размеров стран в модели, где возможны как мирные переговоры, так и военные конфликты [8]. Известно, что с ростом размера страны и численности населения в ней снижаются издержки граждан на общественное благо (оборона, безопасность, образование и т. д.). Но вместе с тем растут издержки на согласование интересов и предпочтений различных этнических и социальных групп. В случае высокой разнородности этих групп и попытки навязать всем группам некоторые действия издержки могут проявляться в форме межнациональных конфликтов и гражданских войн. Авторами показано, что существуют оптимальные размеры стран, которые могут быть установлены в форме переговоров или конфликтов.

4. ПРИМЕНЕНИЕ ФУНКЦИИ ПОБЕДЫ ПРИ МОДЕЛИРОВАНИИ КОНФЛИКТА

Рассмотрим две страны $i = 1, 2$ (без ограничения общности можно считать, что это два блока стран, управляемых двумя центрами). Для определенности положим, что имеется спорная территория ($i = 3$), например, ранее находившаяся в совместном управлении первой и второй стран (рис. 2).

Страна 1	Спорная территория	Страна 2
----------	--------------------	----------

Рис. 2. Две страны и спорный ресурс

Введем следующие обозначения: s_i – площадь территории i -й страны (спорной территории); z_i – численность населения i -й страны; $\mu_{ij} \geq 1$ – параметр этнической разнородности между населением стран i и j ; δ_i – параметр притяжения i -й страны; A_i – индекс ее социально-технологического развития. Положим, что $A_1 = A_2 = A_3 = A$, $z_1 > z_2 > z_3$.

Следуя сложившейся традиции (см. статью [8]), каждой стране поставим в соответствие аналог производственной функции – функцию безопасности [38]:

$$u_i = w_i q_i, \quad w_i = A \left(\frac{z_i}{z_{\max}} \right)^\omega \left(\frac{s_i}{s_{\max}} \right)^{1-\omega},$$

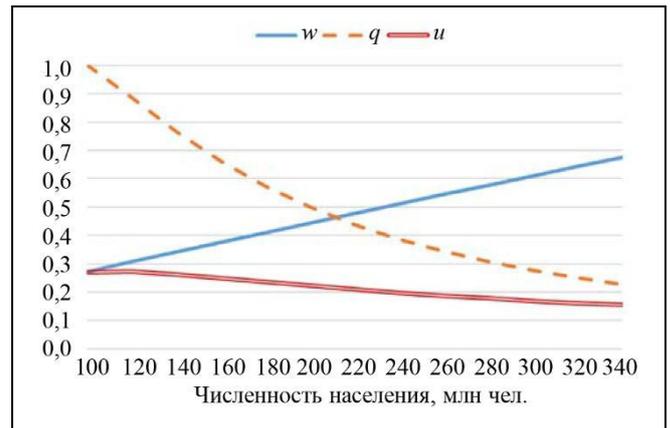
$$q_i = \left(\frac{\xi_i}{z_i} \right)^{\frac{\xi_i + \mu_i (z_i - \xi_i)}{\delta_i z_i}},$$

где z_{\max} – численность населения крупнейшей по этому показателю страны (в настоящий момент – Индии); s_{\max} – площадь территории крупнейшей по этому показателю страны (России); $\omega \approx 0,5$ – параметр эластичности по населению; ξ_i – численность основного этноса i -й страны; μ_i – параметр этнической разнородности в i -й стране (между основным этносом и другими); δ_i – параметр притяжения основного этноса i -й страны; w_i – функция суверенитета i -й страны; q_i – функция ее сохранения.

При $\mu_i = 1$ в стране отсутствуют этнические различия (с точки зрения участия этносов в значимых для общества действиях). С увеличением значения μ_i эти различия возрастают. При $\delta_i > 1$ этнос способен эффективно интегрировать в общество другие национальности. Малые значения параметра ($\delta_i < 0,5-0,6$) имеют народы, не создавшие своей государственности (см. работу [38]).

Вхождение в страну (союз, блок) новых стран (регионов, территорий) увеличивает значение функции суверенитета (а, следовательно, сокращает издержки индивидов на производство общественного блага), но вместе с тем уменьшает значение функции сохранения (рост издержек, связанных с межнациональными конфликтами)⁵. Пример динамики функций суверенитета, сохранения и безопасности представлен на рис. 3.

⁵ Этнический состав страны, а, следовательно, и значение ее функции безопасности может меняться и в результате неконтролируемой миграции.


 Рис. 3. Функции суверенитета (w), сохранения (q) и безопасности (u)

На *первом этапе* (переговоры) вычислим функции безопасности новых объединений стран (территорий). Функция безопасности страны 1 и спорной территории:

$$U_{13} = A \left(\frac{z_1 + z_3}{z_{\max}} \right)^\omega \left(\frac{s_1 + s_3}{s_{\max}} \right)^{1-\omega} \left(\frac{\xi_1}{z_1 + z_3} \right)^{\frac{\mu_1 z_1 + \mu_{13} z_3}{\delta_1 (z_1 + z_3)}}.$$

Функция безопасности страны 2 и спорной территории:

$$U_{23} = A \left(\frac{z_2 + z_3}{z_{\max}} \right)^\omega \left(\frac{s_2 + s_3}{s_{\max}} \right)^{1-\omega} \left(\frac{\xi_2}{z_2 + z_3} \right)^{\frac{\mu_2 z_2 + \mu_{23} z_3}{\delta_2 (z_2 + z_3)}}.$$

Функция безопасности стран 1 и 2 и спорной территории:

$$U_{123} = A \left(\frac{z_1 + z_2 + z_3}{z_{\max}} \right)^\omega \left(\frac{s_1 + s_2 + s_3}{s_{\max}} \right)^{1-\omega} \times$$

$$\times \left(\frac{\xi_1}{z_1 + z_2 + z_3} \right)^{\frac{\mu_1 z_1 + \mu_{12} z_2 + \mu_{13} z_3}{\delta_1 (z_1 + z_2 + z_3)}}.$$

Возможны следующие ситуации равновесия (в условиях отсутствия конфликтов) – выигрыши от переговоров в зависимости от значений функций безопасности отдельных стран и их объединений:

- спорная территория объявляет суверенитет;
- страна 1 и спорная территория объединяются;
- страна 2 и спорная территория объединяются;
- объединяются страны 1, 2 и спорная территория.

В случае несогласия одного или нескольких агентов (стран, правительств) может начаться *второй этап* (конфликт, боевые действия). Предметом конфликта может быть борьба стран 1 и 2 за спорную территорию. Определим целевые функции этих стран:



$$H_i = V_i \pi_i - C_i x_i, \quad \pi_1 = \frac{(\beta x_1)^m}{(\beta x_1)^m + (x_2)^m},$$

$$\pi_2 = 1 - \pi_1, \quad V_i = \frac{(z_i + z_3) \delta_i}{z_i \mu_{i3}}, \quad i = 1, 2,$$

где V_i – ценность спорной территории для i -й страны; π_i – вероятность победы i -й страны; x_1 (x_2) – ресурс, выделенный первой (второй) страной для ведения боевых действий; β – параметр боевого превосходства войск первой страны над войсками второй; C_i – издержки на приобретение и поддержание ресурса i -й страны. Параметр этнической разнородности μ_{i3} между субъектами i -й страны и спорной территории отражает издержки i -й страны при ведении боевых действий на спорной территории, а параметр притяжения δ_i – способность их снижать. Таким образом, ценность объекта пропорциональна степени увеличения населения страны и параметру притяжения, обратно пропорциональна параметру разнородности.

Положим $C_1 = C_2 = C$. Для нахождения равновесия Нэша вычислим частные производные и приравняем их к нулю. Опуская промежуточные вычисления, находим:

$$x_1^* = \frac{m(V_1 V_2)^m}{C[(V_1)^m + (V_2)^m]^2} \frac{V_1}{\beta}, \quad x_2^* = \frac{m(V_1 V_2)^m}{C[(V_1)^m + (V_2)^m]^2} V_2.$$

Вероятность победы в конфликте первой страны над второй будет равна

$$\pi_1 = \frac{(V_1)^m}{(V_1)^m + (V_2)^m} = \frac{1}{1 + \left(\frac{\delta_2(1 + z_3/z_2)\mu_{13}}{\delta_1(1 + z_3/z_1)\mu_{23}} \right)^m}.$$

Оценив ожидаемые результаты конфликта, стороны переходят к *третьему этапу* – оценке издержек на интеграцию спорной территории. Определим целевые функции индивидов первой и второй стран:

$$G_i = Y_i - T_i - (\mu_{i3} - 1) S_i \frac{\pi_i z_3}{z_i}, \quad i = 1, 2,$$

где Y_i – доход индивида i -й страны; T_i – его налоги; S_i – издержки индивида, связанные с проведением контртеррористических и специальных операций.

Из последнего выражения видно, что издержки граждан существенно зависят от значения параметра разнородности между страной и населением спорной территории. Эти издержки также зависят от соотношения численности населения спорной территории и численности населения страны.

Рассчитав показатели безопасности, издержек в ходе конфликта и влияния на благосостояние

граждан интеграции нового населения в состав страны, правительству можно принять более обоснованные решения и заручиться поддержкой общественного мнения.

ЗАКЛЮЧЕНИЕ

В настоящей работе сформулированы содержательные и формальные требования к функции победы в боевых и специальных действиях как к одной из разновидностей функций технологии конфликта (конкурса).

С использованием закона поражения целей А. Н. Колмогорова и распределения Вейбулла выполнено теоретико-вероятностное обоснование конкретного вида функции победы. Содержательное обоснование функции и ее параметров основано на использовании положений военной науки и практики. Выполненное обоснование функции победы в боевых и специальных действиях дает основания для выделения ее в отдельный класс функций технологии конфликта (конкурса).

В последние десятилетия предметная область политической экономики и экономической науки расширилась. Помимо задач производства и распределения благ стали исследоваться вопросы присвоения как результат конфликта между агентами (странами) и установления оптимальных границ между странами. С использованием модели безопасности и функции победы выполнена постановка трех задач распределения спорного ресурса (территории и населения) между странами. На первом этапе для обоснования переговорных позиций вычисляются функции безопасности стран в случае вхождения в их состав спорной территории. Если вопрос не решается мирным путем, правительства стран угрожают конфликтом (боевыми действиями), выделяя соответствующий ресурс на вооруженные силы. Оценка исхода конфликта – основной результат второго этапа. На третьем этапе оцениваются издержки в случае присоединения спорной территории к одной из стран, что позволяет ее правительству заручиться поддержкой общества при ведении переговоров или в ходе конфликта.

Перспективным направлением исследований следует считать развитие теории конфликта на нескольких уровнях – на театре военных действий, межстрановом и геополитическом.

Благодарности. Автор выражает признательность А. А. Галяеву и Д. А. Новикову за содержательные замечания, все недостатки работы относит на свой счет.

ЛИТЕРАТУРА

1. *Перегудов Ф.И., Тарасенко Ф.П.* Основы системного анализа: учеб., 2-е изд., доп. – Томск: Изд-во НТЛ, 1997. – 396 с. [*Peregudov, F.I., Tarasenko F.P.* Fundamentals of Systems Analysis: ucheb., 2-e izd., dop. – Tomsk: Izd-vo NTL, 1997. – 396 s. (In Russian)]
2. *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.]
3. *Cobb, C.W., Douglas, P.H.* A Theory of Production // *The American Economic Review*. – 1928. – Vol. 18, no. 1. – P. 139–165.
4. *Bernoulli, D.* Specimen theoriae novae de mensura sortis // *Commentarii Academiae Scientiarum Imperialis Petropolitanae*. Petropoli. – 1738. – Vol. V. – P. 175–192.
5. *von Neumann, J., Morgenstern, O.* Theory of Games and Economic Behavior. – Princeton: Princeton University Press, 1944. – 668 p.
6. *Schoemaker, P.J.H.* The Expected Utility Model: Its Variants, Purposes, Evidence and Limitations // *Journal of Economic Literature*. – 1982. – Vol. XX, no.2. – P. 529–563.
7. *Garfinkel, M.R., Skaperdas, S.* Economics of Conflict: An Overview // *Handbook of Defense Economics*. – 2007. – No. 2. – P. 649–709.
8. *Alesina, A., Spolaore, E.* Conflict, Defense Spending, and the Number of Nations // *European Economic Review*. – 2006. – No. 50. – P. 91–120.
9. *Модели военных, боевых и специальных действий / под ред. Д.А. Новикова.* – М.: ЛЕНАНД, 2025. – 528 с. [*Models of Military, Combat and Special Operations / Ed. by D.A. Novikov.* – Moscow: LENAND, 2025. – 528 p. (In Russian)]
10. *Robinson, J.* The Production Function and the Theory of Capital // *The Review of Economic Studies*. – 1953. – Vol. 21, no. 2. – P. 81–106.
11. *Solow, R.M.* The Production Function and the Theory of Capital // *The Review of Economic Studies*. – 1955. – Vol. 23, no. 2. – P. 101–108.
12. *Haavelmo, T.* A Study in the Theory of Economic Evolution. – Amsterdam: North-Holland, 1954. – 114 p.
13. *Bernoulli, J.* Ars conjectandi, opus posthumum. Accedit Tractatus de seriebus infinitis, et epistola Gallice scripta De ludo pilae reticularis. – Basileae: impensis Thurnisiorum, fratrum, 1713. – 358 s.
14. *Laplace, P.S.* Théorie analytique des probabilités. – Paris: Ve Courcier, 1812. – 612 p.
15. *Новиков Д.А.* Кибернетика: Навигатор. История кибернетики, современное состояние, перспективы развития. – М.: ЛЕНАНД, 2016. – 160 с. [*Novikov, D.A.* Cybernetics: From Past to Future. – Cham: Springer, 2016.]
16. *Clausewitz, K.* Vom Krieg. – 1832. – 462 s.
17. *Гродецкая А.Г.* О войне и мире на берегах Гудзона: конференция «“War and Peace” at West Point» // *Русская литература*. – 2010. – № 4. – С. 115–119. [*Grodetskaya, A.G.* “War and Peace” at West Point // *Russian Literature*. – 2010. – No. 4. – S. 115–119. (In Russian)]
18. *Lipsky, T.* Tolstoy's Complaint: Mission Command in the Age of AI // *Modern War Institute*. – 07.01.2025. – URL: <https://mwi.westpoint.edu/tolstoys-complaint-mission-command-in-the-age-of-artificial-intelligence/> (дата обращения: 08.09.2025). [Accessed September 8, 2025.]
19. *Толстой Л.Н.* Война и мир. В 4 т. – М.: Юрайт. – Т. 1 и 2, 2019. – 583 с. Т. 3 и 4, 2025. – 593 с. [*Tolstoy, L.N.* War and Peace. Vol. 4 t. – М.: Yurajt. – Т. 1 i 2, 2019. – 583 s. Т. 3 i 4, 2025. – 593 s. (In Russian)]
20. *Knight, F.* Risk, Uncertainty and Profit / Hart Schaffner and Marx Prize Essays. No. 31. – Boston and New York: Houghton Mifflin, 1921. – 381 p.
21. *Кузнецов О.П.* О машинном обучении, мифах о сильном искусственном интеллекте и о том, что такое понимание // *Онтология проектирования*. – 2024. – Т. 14, № 4 (54). – С. 466–482. [*Kuznetsov, O.P.* On Machine Learning, Myths About General AI, and What Understanding Is // *Ontology of Designing*. – 2024. – Vol. 14, no. 4 (54). – P. 466–482. (In Russian)]
22. *Забудский Н.А.* Теория вероятностей и применение ее к стрельбе и пристрелке. – СПб.: Типография Императорской Академии Наук, 1898. – 426 с. [*Zabudskij, N.A.* Probability Theory and its Application to Shooting and Sighting. – SPb.: Tipografiya Imperatorskoj Akademii Nauk, 1898. – 426 s. (In Russian)]
23. *Осипов М.П.* Влияние численности сражающихся сторон на их потери // *Военный сборник*. – 1915. – № 6. – С. 59–74; № 7. – С. 25–36; № 8. – С. 31–40; № 9. – С. 25–37. [*Osirov, M.P.* Influence of the Number of Fighting Parties on Their Losses // *Voennyi Sbornik*. – 1915. – No. 6. – S. 59–74; no. 7. – S. 25–36; no. 8. – S. 31–40; no. 9. – S. 25–37. (In Russian)]
24. *Lanchester, F.* Aircraft in Warfare: the Dawn of the Fourth Arm. – London: Constable and Co., 1916. – 243 p.
25. *Триандафиллов В.К.* Характер операций современных армий. 3-е изд. – М.: Воениздат, 1936. – 260 с. [*Triandafillov, V.K.* The Nature of Operations of Modern Armies. 3-e izd. – М.: Voenizdat, 1936. – 260 s. (In Russian)]
26. *Nash, J.* Non-cooperative Games // *Annals of Mathematics*. – 1951. – Vol. 54, no. 2. – P. 286–295.
27. *Речь Г.К. Жукова на военно-научной конференции, декабрь 1945 г.* // *Военная мысль*. – 1985. Специальный выпуск (февраль). – С. 3, 17–33. [*Speech by G. K. Zhukov at a Military-Scientific Conference, December 1945.* // *Voennaya mysl'*. – 1985. Special'nyj vypusk (fevral'). – S. 3, 17–33. (In Russian)]
28. *Wiener, N.* Cybernetics or Control and Communication in the Animal and the Machine. – New York: John Wiley & Sons, Inc. Paris: Hermann et cie, 1948. – 194 p.
29. *Deitchman, S.* A Lanchester Model of Guerrilla War // *Operational Research*. – 1962. – No. 10. – P. 818–827.
30. *Luce, R. D.* Individual Choice Behavior: A Theoretical Analysis. – New York: Wiley, 1959. – 176 p.
31. *Skaperdas, S.* Contest Success Functions // *Economic Theory*. – 1996. – No. 7. – P. 283–290.
32. *Jia, H., Skaperdas, S., Vaidya, S.* Contest Functions: Theoretical Foundations and Issues in Estimation // *International Journal of Industrial Organization*. – 2013. – No. 31. – P. 211–222.
33. *Колмогоров А.Н.* Число попаданий при нескольких выстрелах и общие принципы оценки эффективности системы стрельбы // *Тр. Матем. ин-та им. В.А. Стеклова*. – 1945. – Т. 12. – С. 7–25. [*Kolmogorov, A.N.* Number of Hits with Multiple Shots and General Principles for Evaluating the Effectiveness of a Firing System // *Tr. Matem. in-ta im. V.A. Steklova*. – 1945. – Vol. 12. – S. 7–25. (In Russian)]



34. *Jia, H.A.* Stochastic Derivation of the Ratio Form of Contest Success Functions // *Public Choice*. – 2008. – No. 135 (3). – P. 125–130.
35. *Корепанов В.О., Шумов В.В.* Моделирование военных, боевых и специальных действий // *Военная мысль*. – 2023. – № 1. – С. 28–41. [*Korepanov, V.O., Shumov, V.V.* Simulation of Military, Combat and Special Operations // *Voennaya mysl'*. – 2023. – No 1. – S. 28–41. (In Russian)]
36. *Шумов В.В.* Исследование функции победы в бою (сражении, операции) // *Проблемы управления*. – 2020. – № 6. – С. 19–30. [*Shumov, V.V.* A Study of Contest Success Function for Battles (Combats, Operations) // *Control Sciences*. – 2020. – No. 6. – P. 19–30. (In Russian)]
37. *Елагин Д.П.* Политическая экономия внутренних вооруженных конфликтов: обзор теоретических объяснений их возникновения (часть I) // *Мировое и национальное хозяйство*. – 2021. – № 1 (54). – Ст. № 6. [*Elagin, D.P.* The Political Economy of Internal Armed Conflicts: A Review of Theoretical Explanations of Their Occurrence (Part I) // *Mirovoe i nacional'noe hozyajstvo*. – 2021. – No. 1 (54). – St. no. 6. (In Russian)]
38. *Шумов В.В.* Национальная безопасность: моделирование и прогнозирование. – М.: ЛЕНАНД, 2023. – 138 с. [*Shu-*

mov, V.V. National Security: Modeling and Forecasting. – Moscow: LENAND, 2023. – 138 s. (In Russian)]

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

Поступила в редакцию 27.09.2025.

Принята к публикации 23.10.2025.

Шумов Владислав Вячеславович – д-р техн. наук, профессор, Институт проблем управления им. В. А. Трапезникова РАН, г. Москва,

✉ v.v.shumov@yandex.ru

ORCID iD: <https://orcid.org/0000-0002-5722-7770>

© 2026 г. Шумов В. В.



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

THE VICTORY FUNCTION AND ITS APPLICATION IN CONFLICT MODELING

V. V. Shumov

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

✉ v.v.shumov@yandex.ru

Abstract. The victory function (VF) in combat and special operations is a separate class of contest success functions (CSFs) describing the probability of a participant's success in a competition or conflict. This paper briefly characterizes aggregate functions, which include production functions, utility functions, CSFs, power (might) indices of countries, etc. Substantive (postulates) and formal (axioms, properties) requirements for the VF are given. Probabilistic (based on A.N. Kolmogorov's law of target destruction and the Weibull distribution) and substantive (based on the framework of military science and practice) justifications of the VF are presented. Economic models of conflict and appropriation are overviewed. Using the security model and the VF, three problems of distributing a disputed resource (territory and population) between countries are formulated. A promising line of further research is to develop conflict theory at several levels, namely, in the theater of military operations, the intercountry, and geopolitical levels.

Keywords: mathematical model, security, combat and special operations, victory function, postulates of conflict technology, models of conflict and appropriation, military cybernetics.

Acknowledgments. The author is grateful to A.A. Galyaev and D.A. Novikov for careful reading of the manuscript and helpful remarks.

СИНТЕЗ САМОПРОВЕРЯЕМЫХ ДИСКРЕТНЫХ УСТРОЙСТВ НА ОСНОВЕ ЛОГИЧЕСКОЙ КОРРЕКЦИИ СИГНАЛОВ И КОМПОЗИЦИИ РАВНОВЕСНЫХ КОДОВ «1 ИЗ 4» И «3 ИЗ 4».

Ч. 1. Метод синтеза с преобразованием всех сигналов от объекта диагностирования

Д. В. Ефанов*, Е. И. Елина**

***Санкт-Петербургский политехнический университет Петра Великого, г. Санкт-Петербург;

*Институт проблем транспорта им. Н. С. Соломенко РАН, г. Санкт-Петербург

*✉ TrES-4b@yandex.ru, **✉ eseniya-elina@mail.ru

Аннотация. Предложено при синтезе самопроверяемых дискретных устройств на основе логической коррекции сигналов использовать композицию равновесных кодов «1 из 4» и «3 из 4». Для данной композиции равновесных кодов можно реализовать простую и компактную самопроверяемую схему тестера, требующую для полной проверки всего четырех тестовых комбинаций. Описана структура самопроверяемого дискретного устройства. Подразумевается преобразование всех четырех сигналов от объекта диагностирования при синтезе для него схемы встроенного контроля. Разработан простейший алгоритм синтеза самопроверяемой схемы встроенного контроля на основе логической коррекции сигналов и композиции равновесных кодов «1 из 4» и «3 из 4». Приведены результаты экспериментов с тестовыми комбинационными схемами, демонстрирующие преимущество предлагаемого метода перед известным методом дублирования и перед методом, основанным на применении равновесного кода «1 из 4» совместно с логической коррекцией сигналов, по показателям структурной избыточности конечного самопроверяемого дискретного устройства. Использование разработанного авторами метода перспективно при синтезе самопроверяемых дискретных устройств на разной элементной базе в различных сферах применения.

Ключевые слова: самопроверяемое дискретное устройство, схема встроенного контроля, логическая коррекция сигналов, композиция равновесных кодов «1 из 4» и «3 из 4», структурная избыточность, контролепригодная структура.

ВВЕДЕНИЕ

Часто при синтезе самопроверяемых дискретных устройств применяют методы, основанные на внесении избыточности не в сам исходный объект (объект диагностирования), а в специализированную схему встроенного контроля (СВК) [1–3]. Такая схема строится с учетом особенностей объекта диагностирования и выбираемого диагностического признака для фиксации неисправностей и ошибок в вычислениях. По сути, реализуется система

рабочего диагностирования для заданного объекта, которая косвенно, по результатам вычисления им своих функций, фиксирует факт наличия или отсутствия неисправностей [4]. Схемы встроенного контроля реализуются самопроверяемыми, что позволяет исключить проблему «сторожа над сторожем», когда разработчику требуется задумываться над тем, как обеспечить контроль исправности самой СВК. Полностью самопроверяемые СВК являются самотестируемыми и защищенными от неисправностей [5]. Первое свойство подразумева-



ет существование для всех неисправностей СВК из заданного класса хотя бы одного набора значений аргументов, при подаче которого на ее входы устанавливается контрольный сигнал ошибки. Второе свойство говорит о том, что при возникновении любой неисправности в СВК из заданного класса на ее выходах либо вычисляются верные значения, либо фиксируется сигнал ошибки.

Существуют два основных подхода к синтезу СВК, каждый из которых порождает множество методов. Первый из них является классическим и известен с середины прошлого столетия [6]. Он подразумевает дополнение информационных сигналов, формируемых на параллельных выходах объекта диагностирования при подаче на входы каждого из наборов значений аргументов, в СВК контрольными сигналами таким образом, чтобы в исправном состоянии объекта диагностирования было установлено некое соответствие между информационными и контрольными сигналами. Зачастую это делается с использованием диагностического признака принадлежности кодовых слов, образуемых информационными и контрольными сигналами, некоторому наперед заданному равномерному блоковому коду [7–9]. Неисправности объекта диагностирования в процессе его эксплуатации вызывают ошибки в вычислениях, что приводит к нарушению соответствия между значениями информационных и контрольных сигналов, что, в свою очередь, и фиксируется в СВК. Второй подход известен с конца прошлого столетия – начала нынешнего [10–13]. Он подразумевает не дополнение информационных сигналов контрольными, а преобразование всех или части сигналов таким образом, чтобы на каждом наборе значений аргументов было зафиксировано требуемое кодовое слово с заданными диагностическими свойствами.

Оба подхода имеют свои преимущества и недостатки. Например, подход, основанный на дополнении сигналов, позволяет учитывать напрямую свойства обнаружения ошибок равномерными блоковыми кодами и использовать особые схемотехнические способы, позволяющие всегда обнаруживать ошибки определенных видов в СВК [14, 15]. При этом, однако, для заданного элементного базиса и способа построения компонентов СВК ее структура оказывается единственной, и разработчик не может повлиять на характеристики эффективности реализации конечного самопроверяемого дискретного устройства (например, на показатели структурной избыточности, контролепригодности, энергопотребления и пр.). Второй же подход, основанный на логической коррекции сигналов (ЛКС), позволяет на каждом наборе значений ар-

гументов выбрать кодовое слово, в которое преобразуется двоичный вектор с выходов объекта диагностирования, из некоторого множества, что определяет широкие возможности для построения различных СВК даже для заданных элементной базы и методов их реализации [16]. При этом можно решать оптимизационные задачи для того или иного показателя эффективности СВК. Но из-за использования ЛКС для вектора, формируемого на выходах объекта диагностирования, оказывается невозможным применение напрямую известных схемотехнических приемов для покрытия отдельных видов ошибок в СВК. Требуются иные подходы, например, выделение подмножеств выходов с особыми используемыми при организации контроля свойствами (например, на которых невозможно возникновение ошибок с заданной их кратностью d (одно-, двух-, трехкратных и т. д.)) [17]. Во всех этих случаях существуют два варианта построения СВК, первый из которых подразумевает полное покрытие ошибок на выходах объекта диагностирования, вызванных неисправностями в его внутренней структуре, а второй – проявление неисправностей хотя бы на одном наборе значений аргументов [18, 19].

Наиболее просто при использовании второго подхода реализуются СВК, основанные на применении неразделимых блоковых кодов. К ним относятся равновесные коды [20], коды Плоткина (Адамара) [21], коды Бордена [22] и другие разнообразные композиции равновесных кодов [23]. Особенностью методов, использующих такие коды, является то, что при организации СВК можно легко схемотехническими приемами добиться исключения ошибок определенного вида в кодовых словах, тогда как для разделимых кодов приходится еще учитывать потенциальные одновременные искажения и информационных, и проверочных символов [24].

Исследования авторов настоящей статьи показали, что помимо традиционных равновесных кодов существует особая их композиция, образованная путем объединения кодовых слов, принадлежащих кодам «1 из 4» и «3 из 4» [23]. Для нее может быть построен тестер с простой и компактной структурой, для полной проверки которого требуются всего четыре тестовые комбинации. Это минимальное значение мощности множества кодовых слов, образующих проверяющий тест, для самопроверяемых тестеров [5].

Данная работа посвящена описанию методов синтеза СВК на основе ЛКС с применением композиции равновесных кодов «1 из 4» и «3 из 4». Она состоит из двух частей: в первой авторы представляют структуру организации СВК на основе

ЛКС с преобразованием сигналов со всех выходов объекта диагностирования и простейший алгоритм синтеза самопроверяемой СВК, а во второй приводятся методы, позволяющие снижать структурную избыточность путем использования свойств композиции равновесных кодов «1 из 4» и «3 из 4» и уменьшения числа элементов преобразования в СВК.

1. ОРГАНИЗАЦИЯ СХЕМЫ ВСТРОЕННОГО КОНТРОЛЯ НА ОСНОВЕ КОМПОЗИЦИИ РАВНОВЕСНЫХ КОДОВ «1 ИЗ 4» И «3 ИЗ 4»

Структура организации СВК на основе композиции равновесных кодов «1 из 4» и «3 из 4» изображена на рис. 1. Она строится для множества из четырех выходов объекта диагностирования. Выбор множества выходов мощностью 4 связан с длиной кодового вектора композиции равновесных кодов «1 из 4» и «3 из 4».

Объектом диагностирования является комбинационная часть дискретного устройства, или просто комбинационное дискретное устройство $F(X)$, которое вычисляет значения булевых функций $f_1(X), f_2(X), f_3(X), f_4(X)$ при подаче на входы заданного множества наборов значений аргументов $\langle X \rangle = \langle x_i x_{i-1} \dots x_2 x_1 \rangle$ (далее полагается, что это множество полное). Таким образом, на каждом наборе значений аргументов формируется булев вектор $\langle f_4(X) f_3(X) f_2(X) f_1(X) \rangle$. Для контроля корректности вычислений обустраивается специализированная СВК. Она состоит из трех функциональных блоков: блока $G(X)$ вычисления значений функций коррекции, непосредственно блока коррекции сигналов (БКС) и самопроверяемого тесте-

ра композиции равновесных кодов «1 из 4» и «3 из 4» TSC (*totally self-checking checker*).

Блок $G(X)$ предназначен для вычисления значений функций $g_1(X), g_2(X), g_3(X), g_4(X)$ для коррекции сигналов от объекта диагностирования $F(X)$ при подаче на входы обоих блоков одинаковых наборов значений аргументов. Каждая из функций коррекции сигналов предназначена для реализации в БКС преобразования вида

$$h_i(X) = f_i(X) \oplus g_i(X), \quad i = \overline{1, 4}, \quad (1)$$

где $h_i(X)$ – скорректированное значение на выходе БКС, получаемое при подаче на входы конкретного набора значений аргументов.

Для коррекции сигналов используются двухвходовые элементы сложения по модулю два (элементы XOR). Их использование позволяет преобразовать любое значение 0 (1) в любое из значений 0 (1)¹. Тем самым, на каждом наборе значений аргументов можно преобразовать кодовый вектор $\langle f_4(X) f_3(X) f_2(X) f_1(X) \rangle$ в кодовый вектор $\langle h_4(X) h_3(X) h_2(X) h_1(X) \rangle$ с заданными диагностическими свойствами. В представленной на рис. 1 структуре осуществляется преобразование в кодовое слово, принадлежащее композиции равновесных кодов «1 из 4» и «3 из 4». Для проверки принадлежности кодового слова $\langle h_4(X) h_3(X) h_2(X) h_1(X) \rangle$ заданной композиции устанавливается блок TSC . Он снабжен двумя выходами $z^0(X)$ и $z^1(X)$, которые функционируют в парафазной логике: комбинации $\langle 01 \rangle$ и $\langle 10 \rangle$ свидетельствуют об отсутствии ошибок в вычислениях, а комбинации $\langle 00 \rangle$ и $\langle 11 \rangle$ говорят о присутствии ошибки и, косвенно, о наличии неисправности в объекте диагностирования.

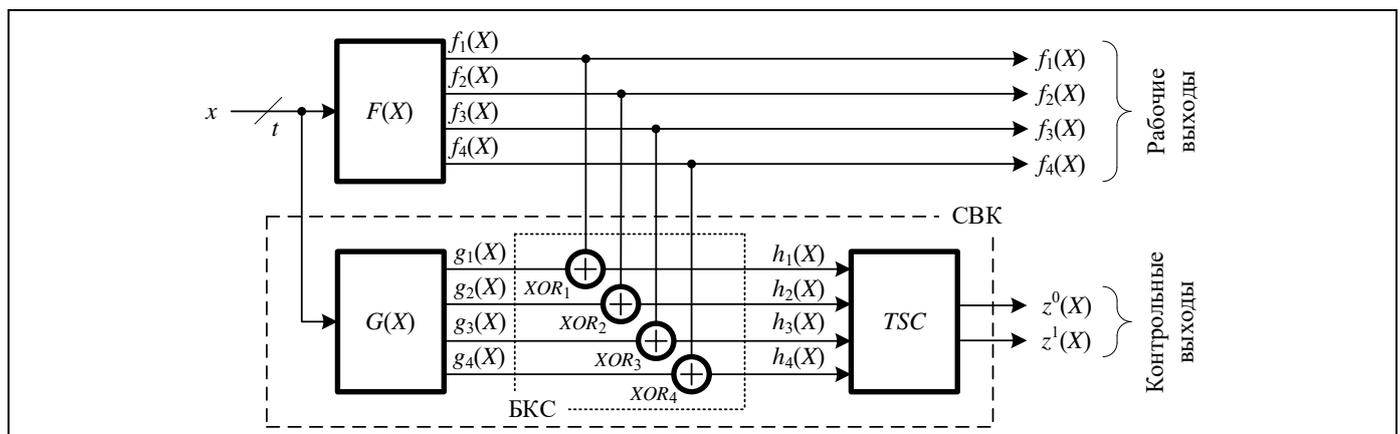


Рис. 1. Структура организации СВК

¹ Аналогичным свойством обладает функция равнозначности, реализуемая элементом $XNOR$. Ее также можно было бы использовать при синтезе СВК на основе ЛКС. Другие элементарные булевы функции можно применять при ЛКС только с рядом ограничений.



На рис. 2 приведена структура простейшего тестера композиции равновесных кодов «1 из 4» и «3 из 4».

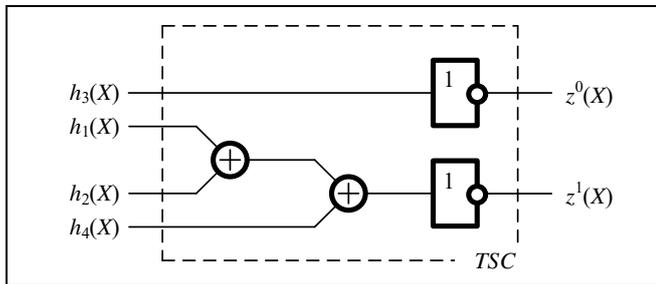


Рис. 2. Структура тестера композиции равновесных кодов «1 из 4» и «3 из 4»

Проверяющий тест T_{Π} для изображенной на рис. 2 структуры тестера содержит комбинации из следующего множества:

$$T_{\Pi} = \{1000, 0010, 1101, 0111\} \cup \{0001, 1011, 0100, 1110\}. \quad (2)$$

При синтезе СВК достаточно обеспечить формирование хотя бы одного из подмножеств из множества (2). Это позволит в процессе эксплуатации самопроверяемого устройства осуществлять полную проверку тестера TSC.

Структура, приведенная на рис. 1, таким образом, будет самопроверяемой.

2. ПРОСТЕЙШИЙ АЛГОРИТМ ПОСТРОЕНИЯ СХЕМЫ ВСТРОЕННОГО КОНТРОЛЯ НА ОСНОВЕ КОМПОЗИЦИИ РАВНОВЕСНЫХ КОДОВ «1 ИЗ 4» И «3 ИЗ 4»

Преобразование вида (1) осуществляется на каждом наборе значений аргументов. Преобразование вектора $\langle f_4(X) f_3(X) f_2(X) f_1(X) \rangle$ в кодировый вектор $\langle h_4(X) h_3(X) h_2(X) h_1(X) \rangle$ можно осуществить восемью разными способами, так как мощность множества кодировых слов композиции равновесных кодов «1 из 4» и «3 из 4» равна $C_4^1 + C_4^3 = 8$. Отсюда следует, что при t аргументах преобразование вида (1) может быть осуществлено 8^{2^t} способами. С учетом необходимости формирования множества T_{Π} для тестера это $8^{2^t} - 3$ способа. Уже при $t = 3$ это дает 16 777 213 способов синтеза самопроверяемых СВК. Среди этих способов можно выбрать наилучший, например, по критерию минимума структурной избыточности самопроверяемого дискретного устройства.

Предложим, однако, здесь простой алгоритм построения СВК на основе ЛКС и композиции

равновесных кодов «1 из 4» и «3 из 4», обеспечивающий формирование множества T_{Π} для тестера, но требующий постпроверки формирования тестовых комбинаций для элементов XOR. Будем исходить из того, что все функции объекта диагностирования определены полностью и на входы поступают все 2^t наборов значений аргументов. Отметим также, что на начальном этапе известны структуры БКС и TSC, а также задана логика работы $F(X)$; значения же функций, реализуемых блоком $G(X)$, не установлены. Основная цель синтеза – получить значения этих функций на каждом наборе значений аргументов, а также обеспечить формирование проверяющих тестов для элементов БКС и TSC.

Алгоритм 1. Синтез СВК для четверки выходов объекта диагностирования:

1. Определить число $\delta = \frac{2^t}{8} = 2^{t-3}$, характери-

зующее количество каждого из используемых слов рассматриваемой композиции равновесных кодов $\langle h_4(X) h_3(X) h_2(X) h_1(X) \rangle$. Это позволит построить СВК, в которой будет обеспечено равномерное формирование кодировых слов композиции равновесных кодов «1 из 4» и «3 из 4» на входах TSC.

2. Рассмотреть наборы значений аргументов в лексикографическом порядке от набора с десятичным эквивалентом 0 до набора с десятичным эквивалентом $2^t - 1$. На наборах значений аргументов с десятичными эквивалентами $0 \dots (2^{t-3} - 1)$ доопределить разряды векторов $\langle h_4(X) h_3(X) h_2(X) h_1(X) \rangle$ до кодирового слова $\langle 0001 \rangle$; на наборах с номерами $2^{t-3} \dots (2 \cdot 2^{t-3} - 1)$ – до кодирового слова $\langle 0010 \rangle$; на наборах с номерами $2 \cdot 2^{t-3} \dots (3 \cdot 2^{t-3} - 1)$ – до кодирового слова $\langle 0100 \rangle$; на наборах с номерами $3 \cdot 2^{t-3} \dots (4 \cdot 2^{t-3} - 1)$ – до кодирового слова $\langle 1000 \rangle$; на наборах с номерами $4 \cdot 2^{t-3} \dots (5 \cdot 2^{t-3} - 1)$ – до кодирового слова $\langle 1110 \rangle$; на наборах с номерами $5 \cdot 2^{t-3} \dots (6 \cdot 2^{t-3} - 1)$ – до кодирового слова $\langle 1101 \rangle$; на наборах с номерами $6 \cdot 2^{t-3} \dots (7 \cdot 2^{t-3} - 1)$ – до кодирового слова $\langle 1011 \rangle$; наконец, на наборах с номерами $7 \cdot 2^{t-3} \dots (8 \cdot 2^{t-3} - 1)$ – до кодирового слова $\langle 0111 \rangle$. Тем самым обеспечивается формирование каждого из кодировых слов композиции равновесных кодов «1 из 4» и «3 из 4» на входах TSC ровно по δ раз.

3. Определить значения функций $g_i(X)$ по известным значениям функций $h_i(X)$ на каждом наборе значений аргументов:

$$g_i(X) = f_i(X) \oplus h_i(X) \Leftrightarrow \Leftrightarrow h_i(X) = f_i(X) \oplus g_i(X), \quad i = \overline{1, 4}. \quad (3)$$

4. Произвести проверку формирования проверяющего теста для каждого из элементов XOR в СВК: в проверяющий тест при канонической ре-

лизации XOR входят все четыре рабочих комбинации {00, 01, 10, 11} [25], каждая из которых должна быть сформирована хотя бы на одном наборе значений аргументов. Если тесты формируются для всех элементов БКС, то перейти к п. 5 алгоритма; в противном случае использовать переопределение значений разрядов в векторе $\langle h_4(X) h_3(X) h_2(X) h_1(X) \rangle$ на выборочном числе наборов значений аргументов в соответствии с методом, изложенным в работе [26].

5. Оптимизировать функции $g_4(X)$, $g_3(X)$, $g_2(X)$, $g_1(X)$ [27].

6. Синтезировать самопроверяемое дискретное устройство в выбранном элементном базисе.

Обратим внимание на п. 2 алгоритма, в котором подразумевается использование всех восьми кодовых слов, принадлежащих заданной композиции. Однако, учитывая выражение (2), можно изменить этот пункт и доопределять векторы $\langle h_4(X) h_3(X) h_2(X) h_1(X) \rangle$ только до четырех кодовых слов заданной композиции, которые обеспечат полную проверку тестера.

Пример 1. Синтезируем СВК для устройства, логика работы которого задана в первых девяти столбцах табл. 1.

В соответствии с п. 1 алгоритма 1 определяем число $\delta = 2^{4-3} = 2$, характеризующее количество каждого из

используемых слов композиции равновесных кодов «1 из 4» и «3 из 4». Затем, согласно п. 2 алгоритма 1, рассматриваем последовательно наборы значений аргументов и фиксируем на них следующие кодовые слова композиции равновесных кодов «1 из 4» и «3 из 4» в векторе $\langle h_4(X) h_3(X) h_2(X) h_1(X) \rangle$: на наборах 0, 1 – слово <0001>, на наборах 2, 3 – слово <0010>, на наборах 4, 5 – слово <0100>, на наборах 6, 7 – слово <1000>, на наборах 8, 9 – слово <1110>, на наборах 10, 11 – слово <1101>, на наборах 12, 13 – слово <1011>, на наборах 14, 15 – слово <0111> (см. столбцы, соответствующие разрядам вектора $\langle h_4(X) h_3(X) h_2(X) h_1(X) \rangle$ в табл. 1). Пользуясь п. 3 алгоритма 1 и формулой (3), получаем значения функций, реализуемых на выходах блока $G(X)$, на каждом наборе значений аргументов (см. столбцы, соответствующие разрядам вектора $\langle g_4(X) g_3(X) g_2(X) g_1(X) \rangle$ в табл. 1). Следуя за п. 4 алгоритма 1, определяем, формируются ли проверяющие тесты для каждого из элементов XOR_4 , XOR_3 , XOR_2 , XOR_1 . Все комбинации, формирующиеся на их входах, приведены в соответствующих столбцах табл. 1. Анализ комбинаций, поступающих на входы каждого элемента коррекции, говорит о том, что проверяющий тест для каждого из них формируется.

Далее следовало бы выполнить п. 5 и 6 алгоритма 1, но они представляются тривиальными: оптимизируются функции $g_4(X)$, $g_3(X)$, $g_2(X)$, $g_1(X)$ и реализуется самопроверяемое дискретное устройство в выбранном элементном базисе. ♦

Таблица 1

Описание работы устройства с четырьмя выходами и СВК для него

№	Наборы значений аргументов				Значения на выходах $F(X)$				Значения на выходах БКС				Значения на выходах $G(X)$				Тестовые комбинации БКС			
	x_4	x_3	x_2	x_1	$f_4(X)$	$f_3(X)$	$f_2(X)$	$f_1(X)$	$h_4(X)$	$h_3(X)$	$h_2(X)$	$h_1(X)$	$g_4(X)$	$g_3(X)$	$g_2(X)$	$g_1(X)$	XOR_4	XOR_3	XOR_2	XOR_1
0	0	0	0	0	1	1	0	0	0	0	0	1	1	1	0	1	11	11	00	01
1	0	0	0	1	0	0	1	0	0	0	0	1	0	0	1	1	00	00	11	01
2	0	0	1	0	1	0	1	0	0	0	1	0	0	0	1	0	10	00	11	00
3	0	0	1	1	0	0	0	0	0	0	1	0	1	0	0	0	01	00	00	00
4	0	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	00	10	00	00
5	0	1	0	1	0	0	1	0	0	1	0	0	0	1	1	0	00	01	11	00
6	0	1	1	0	0	1	0	0	1	0	0	0	0	1	1	0	00	11	01	00
7	0	1	1	1	0	0	0	1	1	0	0	0	0	0	1	1	00	00	01	11
8	1	0	0	0	1	0	1	0	1	1	1	0	1	1	1	0	11	01	11	00
9	1	0	0	1	1	0	0	0	1	1	1	0	1	1	0	0	11	01	00	00
10	1	0	1	0	0	1	1	0	1	1	0	1	1	1	1	0	01	11	11	00
11	1	0	1	1	0	0	0	1	1	1	0	1	1	1	0	0	01	01	00	10
12	1	1	0	0	1	0	1	1	1	0	1	1	0	0	0	0	10	00	10	10
13	1	1	0	1	1	0	1	0	1	0	1	1	0	0	0	1	10	00	10	01
14	1	1	1	0	1	1	1	1	0	1	1	1	1	0	0	0	11	10	10	10
15	1	1	1	1	0	1	0	1	0	1	1	1	0	0	1	0	00	10	01	10



При синтезе СВК для устройства с $n > 4$ выходами используется

Алгоритм 2. Синтез СВК для объекта диагностирования с числом выходов $n > 4$:

1. Упорядочить выходы устройства и разбить их на $q = \left\lceil \frac{n}{4} \right\rceil$ подмножеств (запись $\lceil \dots \rceil$ обозначает целое сверху от вычисляемого значения). Причем, при $n(\bmod 4) = 0$, в каждое из q подмножеств войдут уникальные выходы; если же $n(\bmod 4) \neq 0$, то в последнем подмножестве будут присутствовать $n(\bmod 4)$ уникальных выходов и $4 - n(\bmod 4)$ уже использованных в $q - 1$ подмножестве.

2. Синтезировать СВК по алгоритму 1 для каждого из q подмножеств выходов. При этом для сокращения избыточности отдельные блоки $G(X)$ каждой СВК следует синтезировать совместно.

3. Подключить выходы q СВК к входам самопроверяемого компаратора $qTRC1$, осуществляющего сжатие q парафазных сигналов в один и строящегося на основе $q - 1$ элементарного модуля сжатия парафазных сигналов TRC (*two-rail checker*) [28–30].

Пример 2. Привести структуру самопроверяемого дискретного устройства для организации СВК для исходного устройства с $n = 10$ выходами.

В соответствии с п. 1 алгоритма 2 упорядочиваем выходы и разбиваем их на следующее количество подмножеств: $q = \left\lceil \frac{10}{4} \right\rceil = 3$. Первые два подмножества не пересекаются: I – $\{f_1(X), f_2(X), f_3(X), f_4(X)\}$, II – $\{f_5(X), f_6(X), f_7(X), f_8(X)\}$. Так как $10(\bmod 4) = 2$, в последнее подмножество включаются два выхода из второго подмножества и два оставшихся ранее не использованных выходов: III – $\{f_7(X), f_8(X), f_9(X), f_{10}(X)\}$.

На рис. 3 приведена структура самопроверяемого дискретного устройства в соответствии с данными примера 2. Для ее реализации потребовалось фактически выделить три СВК с отдельными БКС и тестерами. Контрольные выходы отдельных СВК подключены к входам самопроверяемого компаратора, реализованного на основе двух модулей TRC . Так как выходы $f_7(X)$ и $f_8(X)$ использованы дважды, соответствующие функции в СВК отмечены знаками «*» и «**». ♦

Основное преимущество разработанных алгоритмов состоит в том, что они максимально просто позволяют синтезировать СВК для исходных объектов диагностирования, а процедуры, реализуемые в алгоритмах 1 и 2, можно легко автоматизировать.

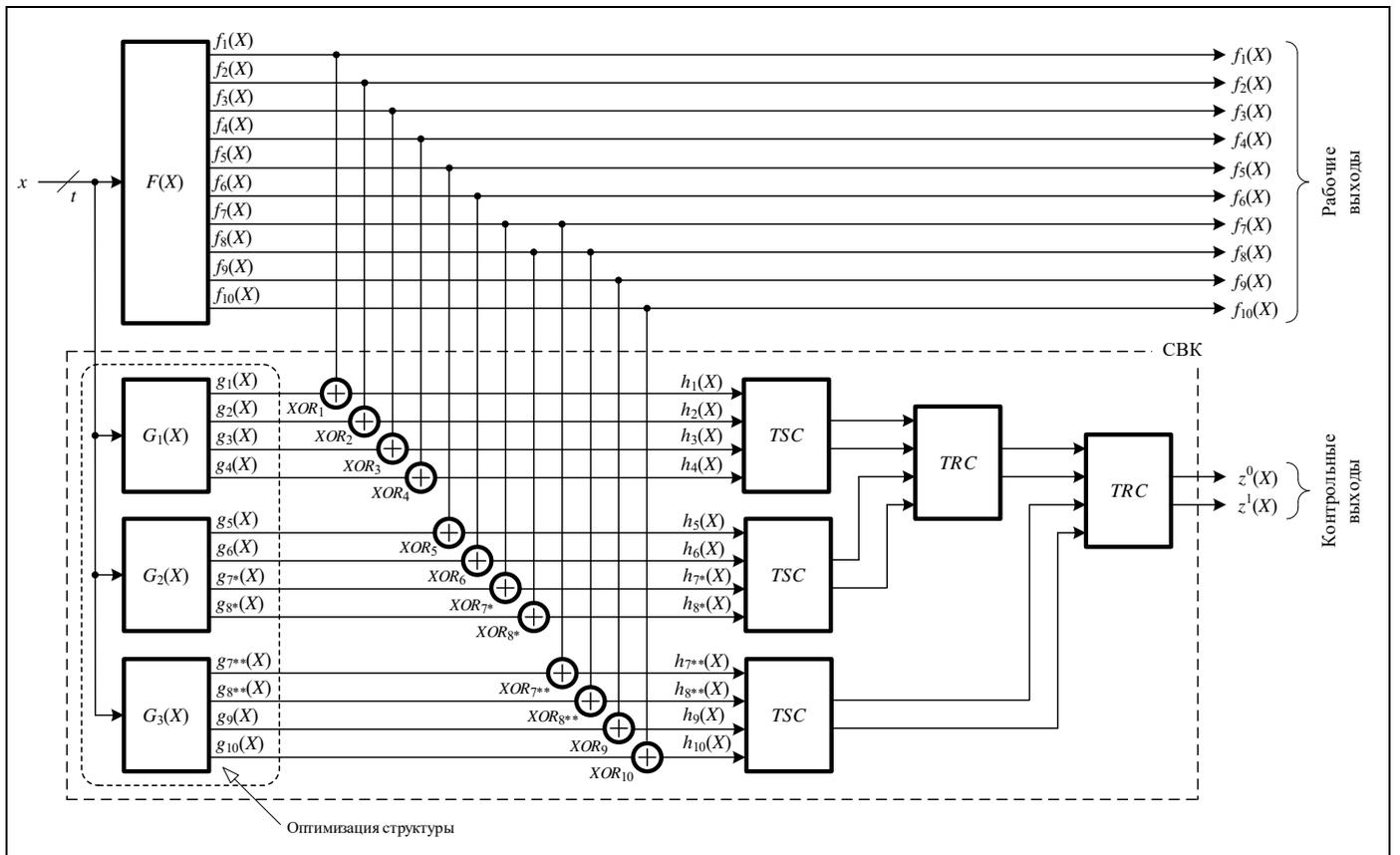


Рис. 3. Структура самопроверяемого дискретного устройства к примеру 2

ровать для использования в системах автоматизированного проектирования самопроверяемых дискретных устройств.

3. РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТОВ

Важной задачей исследования было экспериментальное подтверждение эффективности применения структуры, приведенной на рис. 1, и алгоритмов 1 и 2 при синтезе самопроверяемых дискретных устройств для набора тестовых комбинационных схем MCNC Benchmarks [31, 32]. В ходе экспериментов оценивались показатели структурной избыточности самопроверяемых дискретных устройств, синтезируемых для каждой тестовой схемы. Для сравнения применялись также два метода, в которых подразумевается такое же суммарное количество выходов у блоков $G(X)$, что и у объекта диагностирования: первый – классический метод дублирования, где все блоки $G(X)$ фактически заменены единственной копией объекта диагностирования или устройством, вычисляющим эквивалентные булевы функции [6] (данный метод широко применяется в настоящее время при синтезе самопроверяемых реализаций дискретных устройств [33]), а второй – это метод, основанный на ЛКС и применении равновесного кода «1 из 4» с коррекцией всех сигналов от объекта диагностирования [34].

Для выбранных тестовых схем синтезировались СВК по трем методам. Были выбраны тестовые комбинационные схемы, представленные в формате *.pla. В нем схема задана описанием булевых функций в интервальной форме в виде фактически сжатой таблицы истинности, представляющей собой троичную матрицу с элементарными конъюнкциями [35]. Процедуры обработки тестовых схем, а также разработанные алгоритмы были автоматизированы и добавлены в виде кода в программный модуль DM Coding², что позволило получить описания в формате *.pla для функциональных блоков СВК. Далее с использованием известного интерпретатора SIS [36, 37] и библиотеки функциональных элементов stdcell2_2.genlib были синтезированы самопроверяемые устройства и определены некоторые их параметры. При синтезе

для блоков $G(X)$ использовалась процедура оптимизации булевых функций full_simplify (остальные блоки СВК являются типовыми). Данная процедура позволяет оптимизировать систему булевых функций с применением бинарных диаграмм решений (англ. *binary decision diagram*, BDD) [38]. Затем осуществлялась процедура map -s, позволяющая произвести структурный синтез устройств в заданной библиотеке функциональных элементов и вывести основные их параметры. Для сравнения рассматривался параметр, который характеризует сложность технической реализации устройства, – параметр условной площади, занимаемой устройством на кристалле (в условных единицах библиотеки элементов). Для каждого компонента СВК были определены показатели сложности реализации, а также итоговые значения показателя сложности реализации самопроверяемого устройства (по аналогии с методикой, описанной в работе [39]). Для метода дублирования это величина L_D , для метода на основе ЛКС и кода «1 из 4» – $L_{1/4}$, для метода, описанного в настоящей статье, – $L_{1/4+3/4}$. Для удобства сравнения всех трех методов между собой были рассчитаны следующие относительные показатели:

$$\mu = \frac{L_{1/4+3/4}}{L_D} \cdot 100\%, \quad (4)$$

$$\eta = \frac{L_{1/4+3/4}}{L_{1/4}} \cdot 100\%. \quad (5)$$

Показатели (4) и (5) характеризуют самопроверяемое устройство, синтезированное по предложенному авторами методу, в сравнении с самопроверяемыми устройствами, синтезированными на основе дублирования и кода «1 из 4». Если $\mu > 100\%$, то разработанный в настоящей статье метод эффективнее дублирования, если $\eta > 100\%$, то разработанный метод эффективнее применения кода «1 из 4». Также для удобства анализа эффективности применения нового метода были рассчитаны показатели, характеризующие выигрыш в показателе сложности по сравнению с тем или иным методом:

$$\Delta\mu = \left(1 - \frac{L_{1/4+3/4}}{L_D}\right) \cdot 100\%,$$

$$\Delta\eta = \left(1 - \frac{L_{1/4+3/4}}{L_{1/4}}\right) \cdot 100\%.$$

Значение $\Delta\mu > 0\%$ свидетельствует о выигрыше в показателе сложности реализации перед дублированием; значение $\Delta\eta > 0\%$ – о выигрыше в показателе сложности реализации перед использованием ЛКС и кода «1 из 4».

² Программный модуль DM Coding разработан коллективом научной группы авторов настоящей статьи и канд. техн. наук В. В. Дмитриевым, является наращиваемым и предназначен для постановки экспериментов по разрабатываемым методам технической диагностики. Он написан на языке C# и позволяет реализовывать некоторые функции по анализу тестовых схем в соответствии с разрабатываемыми алгоритмами, в том числе дает возможность получения функционального описания блоков СВК, синтезируемых по различным методам.



В табл. 2 сведены результаты экспериментов для ряда тестовых комбинационных схем. Для каждой тестовой схемы определен показатель сложности реализации $L_{F(X)}$, а также показатели L_D , $L_{1/4}$ и $L_{1/4+3/4}$. Кроме того, для методов, основанных на ЛКС, приведены также данные по значениям показателя сложности реализации блока $G(X)$ – показателя $L_{G(X)}$. При этом для схем с числом выходов $n \geq 5$ блок $G(X)$ получался путем объединения отдельных блоков вычисления значений функций коррекции сигналов каждой из подсхем контроля и синтеза в виде единой схемы. Сравнивая между собой данные показатели для разных методов, основанных на ЛКС, можно оценить влияние показателя сложности реализации блока вычисления функций коррекции на итоговое значение показателя сложности реализации для самопроверяемого дискретного устройства. Отдельно на рис. 4 демонстрируются значения показателей $\Delta\mu$ и $\Delta\eta$ в виде столбчатых диаграмм.

Анализируя данные табл. 2 и рис. 4, можно установить следующее. Для 9-ти тестовых схем из 28-ми оказалось эффективнее использовать метод дублирования. Однако для большинства схем – 19-ти (это составляет примерно 68 % всех схем) – с использованием разработанного метода были получены самопроверяемые дискретные устройства с

меньшими значениями показателей сложности реализации. Причем выигрыш для некоторых схем составил $\Delta\mu > 20\%$. В сравнении с методом, изложенным в работе [34], новый метод дал эффект для 16-ти схем (это составляет примерно 57 % всех схем). Улучшение же по показателю $\Delta\eta$ составило для большинства схем не более 10 %. Это объясняется близостью друг к другу разработанного в настоящем исследовании метода и метода из работы [34] по количеству корректируемых сигналов. Эффект достигался в основном из-за выигрыша в показателе $L_{G(X)}$ при применении того или иного метода. Как следует из рис. 4, для трех тестовых схем получен существенный проигрыш в показателе сложности реализации – свыше 20 % по обоим показателям $\Delta\mu$ и $\Delta\eta$. Это связано со сложностью булевых функций, реализуемых блоками $F(X)$ и $G(X)$ [40]. В отдельно взятых случаях сложность булевых функций, реализуемых блоком $G(X)$, оказалась значительно превышающей сложность булевых функций, реализуемых блоком $F(X)$ (читатель может построчно сравнить данные столбцов $L_{F(X)}$ и $L_{G(X)}$). Но, несмотря на это, для большинства тестовых схем получен положительный эффект при использовании разработанного метода, на основе чего можно судить о его применимости на практике.

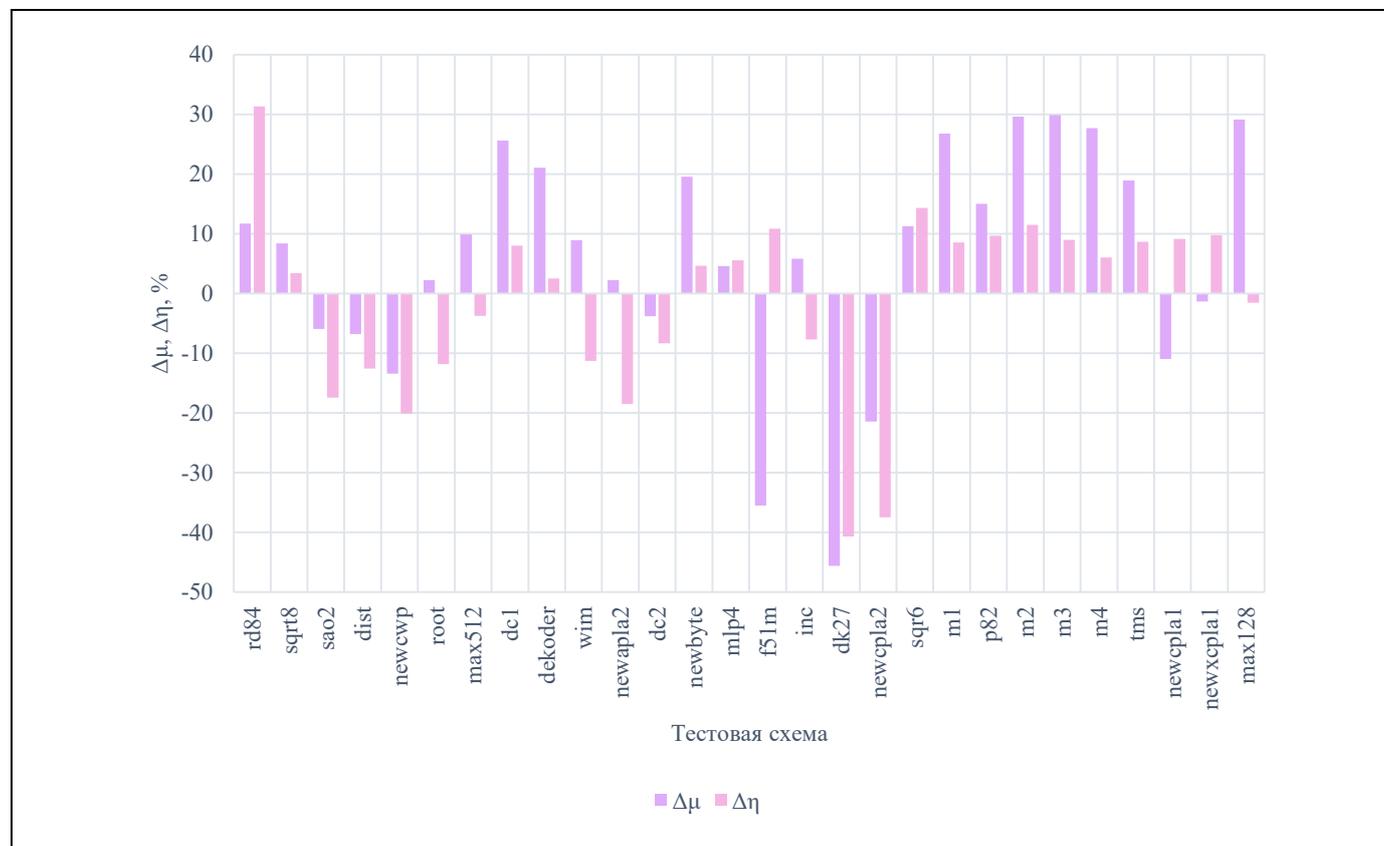


Рис. 4. Значения показателей $\Delta\mu$ и $\Delta\eta$ для различных тестовых схем

Результаты экспериментов

Тестовая схема	n	q	$L_{R(x)}$	L_D	Разработанный метод		Метод из работы [34]		μ	η	$\Delta\mu$	$\Delta\eta$
					$L_{G(x)}$	$L_{1/4+3/4}$	$L_{G(x)}$	$L_{1/4}$				
rd84	4	1	4912	10464	4056	9240	8128	13456	88,303	68,668	11,697	31,332
sqrt8	4	1	1160	2960	1280	2712	1232	2808	91,622	96,581	8,378	3,419
sao2	4	1	2992	6624	3752	7016	2568	5976	105,918	117,403	-5,918	-17,403
dist	5	2	6968	14784	8080	15784	6448	14024	106,764	112,55	-6,764	-12,55
newcwp	5	2	440	1728	784	1960	584	1632	113,426	120,098	-13,426	-20,098
root	5	2	3496	7840	3432	7664	2752	6856	97,755	111,785	2,245	-11,785
max512	6	2	9632	20320	7944	18312	7224	17656	90,118	103,715	9,882	-3,715
dcl	7	2	976	3216	680	2392	632	2600	74,378	92	25,622	8
decoder	7	2	736	2736	688	2160	488	2216	78,947	97,473	21,053	2,527
wim	7	2	712	2688	1000	2448	496	2200	91,071	111,273	8,929	-11,273
newapla2	7	2	600	2464	1072	2408	480	2032	97,727	118,504	2,273	-18,504
dc2	7	2	2424	6112	3184	6344	2440	5856	103,796	108,333	-3,796	-8,333
newbyte	8	2	592	2656	808	2136	384	2240	80,422	95,357	19,578	4,643
mlp4	8	2	7224	15920	7232	15192	7520	16088	95,427	94,431	4,573	5,569
f51m	8	2	2272	6016	5144	8152	5528	9144	135,505	89,151	-35,505	10,849
Inc	9	3	2376	6432	2480	6056	1712	5624	94,154	107,681	5,846	-7,681
dk27	9	3	528	2736	2256	3984	768	2832	145,614	140,678	-45,614	-40,678
newcpla2	10	3	1896	5680	3800	6896	1392	5016	121,408	137,48	-21,408	-37,48
sqrt6	12	3	2648	7600	2896	6744	2672	7872	88,737	85,671	11,263	14,329
m1	12	3	3064	8432	1912	6176	1096	6752	73,245	91,469	26,755	8,531
p82	14	4	2368	7456	2304	6336	1712	7016	84,979	90,308	15,021	9,692
m2	16	4	10096	23328	4656	16416	4288	18544	70,37	88,525	29,63	11,475
m3	16	4	13464	30064	5952	21080	5536	23160	70,117	91,019	29,883	8,981
m4	16	4	18704	40544	8952	29320	8344	31208	72,316	93,95	27,684	6,05
tms	16	4	6784	16704	5096	13544	3928	14832	81,082	91,316	18,918	8,684
newcpla1	16	4	2520	8176	4888	9072	3304	9984	110,959	90,865	-10,959	9,135
newxcpla1	23	6	3760	12112	5920	12272	3216	13600	101,321	90,235	-1,321	9,765
max128	24	6	20192	45184	9232	32016	4528	31536	70,857	101,522	29,143	-1,522



ЗАКЛЮЧЕНИЕ

При синтезе самопроверяемых дискретных устройств на основе ЛКС можно эффективно использовать композицию равновесных кодов «1 из 4» и «3 из 4». Тестер данного кода имеет простую и компактную структуру, для полной проверки которой достаточно четырех комбинаций. Это позволяет довольно просто синтезировать самопроверяемые СВК для дискретных устройств, например, по представленным в статье алгоритмам.

Анализируя результаты экспериментов, полученные с применением использованных программных модулей и продуктов, по показателям структурной избыточности, можно сделать вывод о преимуществах описанного в настоящей статье метода по сравнению с некоторыми известными методами синтеза самопроверяемых дискретных устройств. Более того, как будет показано во второй части статьи, учет свойств композиции равновесных кодов «1 из 4» и «3 из 4» позволяет сократить количество преобразуемых сигналов от объекта диагностирования и тем самым уменьшить количество выходов блоков $G(X)$. Такая процедура модификации структуры организации СВК, изображенной на рис. 1, позволяет добиться еще большего выигрыша в показателях сложности технической реализации самопроверяемых дискретных устройств по сравнению с известными методами. Здесь следует отметить, что для определения особенностей и характеристик разработанного метода также могут использоваться другие метрики, синтезаторы и наборы функциональных элементов.

Достоинства использования ЛКС и композиции равновесных кодов «1 из 4» и «3 из 4» состоят в том, что существует большое число способов синтеза СВК, что позволяет решать оптимизационные задачи и строить самопроверяемые дискретные устройства, например, с наименьшими аппаратными затратами. В качестве недостатка использования разработанного метода следует отметить необходимость учета числа сочетаний искажений на выходах объекта диагностирования при выборе среди них четверок. Здесь существуют два варианта в построении СВК: классический вариант с полным покрытием любых ошибок на выходах объекта диагностирования и вариант с покрытием неисправностей при их проявлении хотя бы на одном наборе значений аргументов. Также может быть использован первый вариант реализации СВК с покрытием любых ошибок на выходах объекта диагностирования с увеличенным числом контролируемых подмножеств выходов по сравнению с ми-

нимально необходимым для разработанного метода, но по показателям сложности технической реализации имеющий преимущество по сравнению с иными методами. О возможности такой реализации говорят некоторые результаты экспериментов и, в частности, анализ выигрышей в показателях сложности.

ЛИТЕРАТУРА

1. *Согомонян Е.С., Слабаков Е.В.* Самопроверяемые устройства и отказоустойчивые системы. – М.: Радио и связь, 1989, 208 с. [*Sogomonyan, E.S., Slabakov, E.V.* Samoproveryaemye ustroystva i otkazoustoichivye sistemy. – Moscow: Radio i svyaz', 1989. – 208 s. (In Russian)].
2. *Mitra, S., McCluskey, E.J.* Which Concurrent Error Detection Scheme to Choose? // Proceedings of International Test Conference. – Atlantic City, NJ, USA, 2000. – P. 985–994. – DOI: 10.1109/TEST.2000.894311
3. *Göessel, M., Ocheretny, V., Sogomonyan, E., Marienfeld, D.* New Methods of Concurrent Checking: Edition 1. – Dordrecht: Springer Science+Business Media B.V., 2008. – 184 p.
4. *Дрозд А.В.* Нетрадиционный взгляд на рабочее диагностирование вычислительных устройств // Проблемы управления. – 2008. – № 2. – С. 48–56. [*Drozhd, A.V.* An Untraditional View on Operational Diagnostics of Computing Devices // Control Sciences. – 2008. – No. 2. – P. 48–56. (In Russian)].
5. *Сапожников В.В., Сапожников В.В.* Самопроверяемые дискретные устройства. – СПб.: Энергоатомиздат, 1992. – 224 с. [*Sapozhnikov, V.V., Sapozhnikov, V.V.* Samoproveryaemye diskretnye ustroystva. – St. Petersburg: Energoatomizdat, 1992. – 224 s. (In Russian)].
6. *Goessel, M., Graf, S.* Error Detection Circuits. – London: McGraw-Hill, 1994. – 261 p.
7. *Сапожников В.В., Сапожников В.В., Ефанов Д.В.* Коды Хэмминга в системах функционального контроля логических устройств. – СПб.: Наука, 2018. – 151 с. [*Sapozhnikov, V.V., Sapozhnikov, V.V., Efanov, D.V.* Kody Khemminga v sistemakh funktsional'nogo kontrolya logicheskikh ustroystv. – St. Petersburg: Nauka, 2018. – 151 s. (In Russian)].
8. *Сапожников В.В., Сапожников В.В., Ефанов Д.В.* Коды с суммированием для систем технического диагностирования. Том 1: Классические коды Бергера и их модификации. – М.: Наука, 2020. – 383 с. [*Sapozhnikov, V.V., Sapozhnikov, V.V., Efanov, D.V.* Kody s summirovaniem dlya sistem tekhnicheskogo diagnostirovaniya. Tom 1: Klassicheskie kody Bergera i ikh modifikatsii. – Moscow: Nauka, 2020. – 383 s. (In Russian)].
9. *Сапожников В.В., Сапожников В.В., Ефанов Д.В.* Коды с суммированием для систем технического диагностирования. Том 2: Взвешенные коды с суммированием. – М.: Наука, 2021, 455 с. [*Sapozhnikov, V.V., Sapozhnikov, V.V., Efanov, D.V.* Kody s summirovaniem dlya sistem tekhnicheskogo diagnostirovaniya. Tom 2: Vzveshennyye kody s summirovaniem. – Moscow: Nauka, 2021. – 455 s. (In Russian)].
10. *Goessel, M., Sapozhnikov, V., Sapozhnikov, V., Dmitriev, A.* A New Method for Concurrent Checking by Use of a 1-out-of-4 Code // Proceedings of 6th IEEE International On-Line Testing Workshop. – Palma de Mallorca, Spain, 2000. – P. 147–152. – DOI: 10.1109/OLT.2000.856627

11. *Dmitriev, A., Saposhnikov, V., Saposhnikov, V., Goessel, M.* New Self-Dual Circuits for Error Detection and Testing // *VLSI Design*. – 2000. – Vol. 11, iss. 1. – P. 1–21. – DOI: 10.1155/2000/84720
12. *Сапожников В.В., Сапожников Вл.В., Дмитриев А.В.* и др. Организация функционального контроля комбинационных схем методом логического дополнения // *Электронное моделирование*. – 2002. – Т. 24, № 6. – С. 52–66. [*Sapozhnikov, V.V., Sapozhnikov, Vl.V., Dmitriev, A.V., et al.* Organization of Functional Control of Combinational Circuits Using the Boolean Complement Method // *Electronic Modeling*. – 2002. – Vol. 24, no. 6. – P. 52–66. (In Russian)].
13. *Гессель М., Морозов А.В., Сапожников В.В., Сапожников Вл.В.* Логическое дополнение – новый метод контроля комбинационных схем // *Автоматика и телемеханика*. – 2003. – № 1. – С. 167–176. [*Gessel, M., Morozov, A.V., Sapozhnikov, V.V., Sapozhnikov, Vl.V.* Logic Complement, a New Method of Checking the Combinational Circuits // *Automation and Remote Control*. – 2003. – Vol. 64, no. 1. – P. 153–161].
14. *Saposhnikov, V.V., Morosov, A., Saposhnikov, Vl.V., Göessel, M.* A New Design Method for Self-Checking Unidirectional Combinational Circuits // *Journal of Electronic Testing: Theory and Applications*. – 1998. – Vol. 12, iss. 1-2. – P. 41–53. – DOI: 10.1023/A:1008257118423
15. *Матросова А.Ю., Митрофанов Е.В.* Синтез легко тестируемых последовательностных схем // *Вестник Томского государственного университета. Управление, вычислительная техника и информатика*. – 2013. – № 2 (23). – С. 140–147. [*Matrosova, A.Yu., Mitrofanov, E.V.* Delay Testable Sequential Circuit Design // *Tomsk State University Journal of Control and Computer Science*. – 2013. – No. 2 (23). – P. 140–147. (In Russian)].
16. *Ефанов Д.В., Елина Е.И.* Исследование способов синтеза схем встроенного контроля на основе логической коррекции сигналов с применением равномерных разделимых кодов // *Микроэлектроника*. – 2024. – Т. 53, № 5. – С. 413–426. – DOI: 10.31857/S0544126924050079. [*Efanov, D.V., Yelina, Y.I.* Investigation of Ways of Synthesizing Concurrent Error-Detection Circuits Based on Boolean Signal Correction Using Uniform Separable Codes // *Russian Microelectronics*. – 2024. – Vol. 53, no. 5. – P. 471–482. – DOI: 10.1134/S1063739724600456].
17. *Ефанов Д.В.* Синтез самопроверяемых вычислительных устройств на основе полной системы особых групп выходов объекта диагностирования // *Известия высших учебных заведений. Приборостроение*. – 2023. – Т. 66, № 5. – С. 355–372. – DOI: 10.17586/0021-3454-2023-66-5-355-372. [*Efanov, D.V.* Synthesis of Self-Checking Computing Devices Based on a Complete System of Special Groups of the Diagnostic Object Outputs // *Journal of Instrument Engineering*. – 2023. – Vol. 66, no. 5. – P. 355–372. – DOI: 10.17586/0021-3454-2023-66-5-355-372 (In Russian)].
18. *Гессель М., Согомонян Е.С.* Построение самотестируемых и самопроверяемых комбинационных устройств со слабозависимыми выходами // *Автоматика и телемеханика*. – 1992. – № 8. – С. 150–160. [*Goessel, M., Sogomonyan, E.S.* Design of Self-Testing and Self-Checking Combinational Circuits with Weakly Independent Outputs // *Automation and Remote Control*. – 1992. – Vol. 53, no. 8. – P. 1264–1272.]
19. *Sogomonyan, E.S., Gössel, M.* Design of Self-Testing and On-Line Fault Detection Combinational Circuits with Weakly Independent Outputs // *Journal of Electronic Testing: Theory and Applications*. – 1993. – Vol. 4, iss. 4. – P. 267–281. – DOI: 10.1007/BF00971975
20. *Freiman, C.V.* Optimal Error Detection Codes for Completely Asymmetric Binary Channels // *Information and Control*. – 1962. – Vol. 5, iss. 1. – P. 64–71. – DOI: 10.1016/S0019-9958(62)90223-1
21. *MacWilliams, F.J., Sloane, N.J.A.* The Theory of Error-Correcting Codes. – Amsterdam: North-Holland, 1977. – 785 p.
22. *Borden, J.M.* Optimal Asymmetric Error Detecting Codes // *Information and Control*. – 1982. – Vol. 53, iss. 1-2. – P. 66–73. – DOI: 10.1016/S0019-9958(82)91125-1
23. *Ефанов Д.В.* Композиции двух равновесных кодов с ортогональными по всем разрядам комбинациями для синтеза самопроверяемых дискретных устройств // *Проблемы управления*. – 2025. – № 3. – С. 49–62. [*Efanov, D.V.* Compositions of Two Constant-Weight Codes with Orthogonal Combinations over All Bits for Self-Checking Discrete Device Design // *Control Sciences*. – 2025. – No. 3. – P. 41–61].
24. *Efanov, D., Osadchy, G., Zueva, M.* Special Aspects of Errors Definition via Sum Codes within Embedded Control Schemas Being Realized by Means of Boolean Complement Method // *Proceedings of 11th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS'2021)*. – Cracow, Poland, 2021. – Vol. 1. – P. 424–431. – DOI: 10.1109/IDAACS53288.2021.9660837
25. *Аксёнова Г.П.* Необходимые и достаточные условия построения полностью проверяемых схем свертки по модулю два // *Автоматика и телемеханика*. – 1979. – № 9. – С. 126–135. [*Aksenova, G.P.* Necessary and Sufficient Conditions for Design of Completely Checkable Modulo 2 Convolution Circuits // *Automation and Remote Control*. – 1979. – Vol. 40, no. 9. – P. 1362–1369].
26. *Ефанов Д.В.* Синтез самопроверяемых комбинационных устройств на основе метода логической коррекции сигналов с применением кодов Боуза–Лина // *Информационные технологии*. – 2023. – Т. 29, № 10. – С. 503–511. – DOI: 10.17587/it.29.503-511. [*Efanov D.V.* Self-Checking Combinational Devices Synthesis Based on the Boolean Signal Correction Method Using Bose–Lin Codes // *Information Technologies*. – 2023. – Vol. 29, no. 10. – P. 503–511. – DOI: 10.17587/it.29.503-511. (In Russian)].
27. *Закревский А.Д., Поттосин Ю.В., Черемисинова Л.Д.* Логические основы проектирования дискретных устройств. – М.: Физматлит, 2007. – 592 с. [*Zakrevskii, A.D., Pottosin, Yu.V., Cheremisinova, L.D.* Logicheskie osnovy proektirovaniya diskretnykh ustroystv. – Moscow: Fizmatlit, 2007. – 592 s. (In Russian)].
28. *Пархоменко П.П., Согомонян Е.С.* Основы технической диагностики. Книга вторая. Оптимизация алгоритмов диагностирования, аппаратные средства. – М.: Энергоатомиздат, 1981. – 320 с. [*Parkhomenko, P.P., Sogomonyan, E.S.* Osnovy tekhnicheskoi diagnostiki. Tom 2: optimizatsiya algoritmov diagnostirovaniya, apparaturnye sredstva. – М.: Ehnergoatomizdat, 1981. – 320 s. (In Russian)].
29. *Lala, P.K.* Self-Checking and Fault-Tolerant Digital Design. – San Francisco: Morgan Kaufmann Publishers, 2001. – 216 p.
30. *Chioktour, V., Kakarountas, A.* Adaptive BIST for Concurrent On-Line Testing on Combinational Circuits // *Electronics*. – 2022. – Vol. 19, no. 11. – P. 1–20. – DOI: 10.3390/electronics11193193



31. *McElvain, K.* IWLS'93 Benchmark Set: Version 4.0: distributed as a part of IWLS'93 benchmark set. – 1993.
32. *Collection of Digital Design Benchmarks.* – URL: <https://ddd.fit.cvut.cz/www/prj/Benchmarks/> (дата обращения: 26.08.2025 г.) [Accessed August 26, 2025].
33. *Hahanov, V., Litvinova, E., Chumachenko, S., et al.* Vector Logic Modeling Self-Tested Circuits // Proceedings of 21st IEEE East-West Design & Test Symposium (EWDTS'2025). – Tbilisi, Georgia, 2025. – DOI: 10.1109/EWDTS67441.2025.11303706
34. *Ефанов Д.В.* Использование равновесного кода «1 из 4» при синтезе самопроверяемых схем встроенного контроля на основе логической коррекции сигналов // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2025. – № 72. – С. 114–133. – DOI: 10.17223/19988605/72/12 [Efanov, D.V. Using the “1-Out-Of-4” Constant-Weight Code in the Synthesis of Self-Checking Concurrent Error-Detection Circuit Based on Boolean Signal Correction // Tomsk State University Journal of Control and Computer Science. – 2025. – No. 72. – P. 114–133. – DOI: 10.17223/19988605/72/12 (In Russian)].
35. *Закревский А.Д., Торопов Н.Р.* Минимизация булевых функций многих переменных в классе ДНФ – итеративный метод и программная реализация // Прикладная дискретная математика. – 2009. – № 1. – С. 5–14. [Zakrevskij, A.D., Toropov, N.R. Minimization of Boolean Functions of Many Variables - Iterative Method and Program Realization // Applied Discrete Mathematics. – 2009. – No. 1. – P. 5–14. (In Russian)].
36. *Sentovich, E.M., Singh, K.J., Moon, C., et al.* Sequential Circuit Design Using Synthesis and Optimization // Proceedings of IEEE International Conference on Computer Design: VLSI in Computers & Processors. – Cambridge, MA, USA, 1992. – P. 328–333. – DOI: 10.1109/ICCD.1992.276282
37. *SIS: A System for Sequential Circuit Synthesis / E.M. Sentovich, K.J. Singh, L. Lavagno, C. Moon, R. Murgai, A. Saldanha, H. Savoj, P. R. Stephan, R.K. Brayton, A. Sangiovanni-Vincentelli.* – Berkeley: Electronics Research Laboratory, Department of Electrical Engineering and Computer Science, University of California, 1992. – 45 p.
38. *Бибило П.Н.* Бинарные диаграммы решений в логическом проектировании. – М.: ЛЕНАНД, 2024. – 560 с. [Bibilo, P.N. Binarnye diagrammy reshenii v logicheskom proektirovanii. – Moscow: LENAND, 2024. – 560 s. (In Russian)].
39. *Ефанов Д.В., Елина Е.И.* Синтез самопроверяемых цифровых устройств на основе логической коррекции сигналов с применением взвешенных кодов Боуза – Лина // Проблемы управления. – 2024. – № 4. – С. 26–43. – DOI: <http://doi.org/10.25728/ru.2024.4.3> [Efanov D.V., Yelina Y.I. Design of Self-Checking Digital Devices with Boolean Signals Correction Using Weight-Based Bose–Lin Codes // Control Sciences. – 2024. – No. 4. – P. 22–36. – DOI: <http://doi.org/10.25728/ru.2024.4.3>].
40. *Коршунов А.Д.* Сложность вычислений булевых функций // Успехи математических наук. – 2012. – Т. 67, № 1 (403). – С. 97–168. – DOI: 10.4213/rm9459. [Korshunov, A.D. Computational Complexity of Boolean Functions // Russian Mathematical Surveys. – 2012. – Vol. 67, no. 1 (403). – P. 93–165. – DOI: <https://doi.org/10.1070/RM2012v067n01ABEH004777>]

Статья представлена к публикации членом редколлегии Л. Ю. Филimonюком.

Поступила в редакцию 14.09.2025,
после доработки 27.11.2025.
Принята к публикации 27.11.2025.

Ефанов Дмитрий Викторович – д-р техн. наук, Санкт-Петербургский политехнический университет Петра Великого, г. Санкт-Петербург; Институт проблем транспорта им. Н. С. Соломенко РАН, г. Санкт-Петербург, ✉ TrES-4b@yandex.ru
ORCID iD: <https://orcid.org/0000-0002-4563-6411>

Елина Есения Игоревна – аспирант, Санкт-Петербургский политехнический университет Петра Великого, г. Санкт-Петербург, ✉ eseniya-elina@mail.ru
ORCID iD: <https://orcid.org/0009-0004-4167-3591>

© 2026 г. Ефанов Д. В., Елина Е. И.



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

**DESIGN OF SELF-CHECKING DISCRETE DEVICES
BASED ON BOOLEAN SIGNALS CORRECTION AND COMPOSITION
OF CONSTANT-WEIGHT CODES OF THE “1-OUT-OF-4” AND “3-OUT-OF-4” TYPES.
PART I: The Design Method with Conversion
of All Signals from the Object under Diagnosis**

D. V. Efanov* and Y. I. Yelina**

***Peter the Great Saint Petersburg Polytechnic University, St. Petersburg, Russia
*Solomenko Institute of Transport Problems, Russian Academy of Sciences, St. Petersburg, Russia

*✉ TrES-4b@yandex.ru, **✉ eseniya-elina@mail.ru

Abstract. It is proposed to use the composition of constant-weight codes of the “1-out-of-4” and “3-out-of-4” types in the design of self-checking discrete devices based on Boolean signals correction. For this composition of constant-weight codes, a simple and compact self-checking checker can be implemented, requiring only four test combinations for a complete check. The structure of a self-checking discrete device is described. The conversion of all four signals from an object under diagnosis is considered when designing a concurrent error-detection circuit for the object. The simplest algorithm for building a self-checking concurrent error-detection circuit based on the Boolean signals correction and the composition of 1-out-of-4 and 3-out-of-4 codes is developed. The results of experiments with combinational benchmarks are provided. They demonstrate the advantage of the proposed approach over the well-known duplication method and the one involving the 1-out-of-4 code together with Boolean signals correction, in terms of structural redundancy of the final self-checking discrete device. The method developed in this paper is promising for self-checking discrete device design with different element bases in various fields of application.

Keywords: self-checking discrete device, concurrent error-detection circuit, Boolean signals correction, composition of constant-weight codes, 1-out-of-4 and 3-out-of-4 codes, structural redundancy, controllable structure.



ПРИМЕНЕНИЕ БОЛЬШИХ ЯЗЫКОВЫХ МОДЕЛЕЙ В СИСТЕМАХ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ.

Ч. 1. Модели объяснения и большие языковые модели

А. А. Кулинич

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

✉ alexkul@rambler.ru

Аннотация. Большие языковые модели оказывают существенное влияние на многие сферы человеческой жизни: на образование, творчество, науку, бизнес. Исследуется применение больших языковых моделей для объяснения альтернатив решений, полученных системой поддержки принятия решений в условиях неопределенности. Рассмотрены классические и прагматические модели объяснения, предложенные философами. Сформулированы цели и задачи объяснения в процессах поддержки принятия решений в условиях неопределенности. Приведены концептуальный анализ функционирования больших языковых моделей и оценка их текущих способностей при решении типовых тестовых задач. Рассмотрены основные техники промптинга (системы запросов к языковой модели), позволяющие настроить языковую модель на генерацию объяснений для альтернатив решений конкретных проблем в предметной области. Рассмотрены техники промптов для поддержки прагматических и классических теорий объяснения альтернатив решений.

Ключевые слова: поддержка принятия решений, модели объяснения, цель объяснения, задачи объяснения, большая языковая модель, техники промптов.

ВВЕДЕНИЕ

Для поддержки принятия решений в социальных, политических, экономических и организационных системах в условиях неопределенности применяют «мягкий» системный анализ. Он основывается на принципе ограниченной рациональности [1], который утверждает, что в силу ограниченных когнитивных ресурсов лиц, создающих модель объекта, можно сформировать упрощенную системную модель с гипотетической структурой, параметры которой могут иметь лингвистические значения [2]. Построение модели в рамках «мягкого» системного анализа основано на изучении литературы, экспертных оценках, т. е. исходными данными являются неструктурированные данные – свободный текст. Модель описывает объект на ограниченном естественном языке, который предлагает эксперт.

Известным математическим аппаратом «мягкого» системного анализа являются когнитивные

карты [3, 4]. Результаты моделирования в такой модели представлены на экспертном ограниченном естественном языке и интерпретируются самим экспертом в терминах его знаний, которых может быть недостаточно для получения нового решения. Интерпретация должна включать знания о предметной области, релеванные решению, но находящиеся за пределами ограниченного экспертного языка и упрощенной модели ситуации. Интерпретируя результаты моделирования, эксперт пытается объяснить и связать модельные процессы с процессами реального мира. В модели ситуации для объяснения результата формально могут быть построены причинно-следственные цепочки вывода результата [5]. Однако эти цепочки рассуждений выражены на ограниченном естественном языке и могут гарантировать строгость логического вывода в рамках принятых допущений и ограничений, но не возможность применения решения в реальном мире. Реальный мир намного разнообразнее и богаче.

В психологии считается, что принятие решений осуществляется интеллектом человека в психической среде, которая называется ментальным пространством. В ментальном пространстве отражены знания, элементы жизненного опыта эксперта, полученные на протяжении всей его жизни. В ментальном пространстве осуществляются мысленные операции – рассуждения, обобщения, интерпретация, объяснение и оценка возможных вариантов решений. Поддержать эти ментальные операции процессов принятия решений в условиях неопределенности можно с помощью систем искусственного интеллекта, в частности больших языковых моделей (англ. *Large Language Model*, LLM).

В настоящее время количество и разнообразие больших языковых моделей стремительно растет. Широкую известность и популярность в практическом применении получили такие большие языковые модели, как GPT-2 [6], GPT-3[7], InstructGPT [8] и GPT-4 [9] от компании OpenAI, BERT [10], RoBERTa, ALBERT [11] (разработчик – Google DeepMind). Активно используются разработки компании Meta AI*, такие как LLaMA (Large Language Model Meta AI) [12]. Очень популярна китайская языковая модель DeepSeek [13]. В России широко используют отечественные большие языковые модели GigaChat-2 от Сбера [14] и Yandex GPT 5.1 от компании Яндекс [15]. Подробный обзор текущего состояния разработок в области больших языковых моделей можно найти в работе [16].

Большие языковые модели применяются для решения самых разных задач в экономике, социальной сфере, медицине, образовании, государственном и муниципальном управлении и др. Приведем некоторые обзоры применения таких моделей: в научных исследованиях [17], для прогнозирования сложных экономических систем в современных условиях [18], в сфере государственного и муниципального управления [19], в области здравоохранения [20], в организациях и банковской сфере [21, 22], в качестве цифровых помощников [23], в бизнес-аналитике и принятии решений [24].

В работе [25] предложена методология использования больших языковых моделей для решения задач государственного и муниципального управления для реферирования и генерации текстового контента. Стратегическое направление в области цифровой трансформации государственного управления в Российской Федерации [26] определяет необходимость внедрения в сферу государ-

ственного управления систем искусственного интеллекта (ИИ), больших данных, Интернета вещей.

В рамках Форсайта по приоритетным направлениям фундаментальных и поисковых исследований в сфере ИИ [27] были выделены следующие приоритетные направления поисковых исследований в этой области: архитектуры и алгоритмы машинного обучения, вычисления и данные для ИИ, фундаментальные и генеративные модели, взаимодействие человека и ИИ, прикладные исследования для науки, образования и социальной сферы.

В центре внимания настоящей статьи находятся актуальные вопросы взаимодействия человека и систем искусственного интеллекта в процессах принятия решений в условиях неопределенности.

Цель статьи заключается в разработке методов и подходов (моделей и алгоритмов) поддержки принятия решений в условиях неопределенности с использованием больших языковых моделей.

Основные задачи, которые необходимо решить для достижения цели, видятся следующим образом.

Вначале необходимо понять, являются ли тексты объяснений, сгенерированные языковой моделью, объяснениями в смысле философских теорий объяснения, которые представляют собой фундаментальное основание методологии науки и научных исследований. Для этого проведем анализ существующих моделей объяснения и выясним их суть.

Вторая задача – это задача настройки большой языковой модели на генерацию корректных объяснений ситуаций для решения задач поддержки принятия решений в условиях неопределенности в тандеме с лицом, принимающим решение (ЛПР).

Третья задача – это оценка качества системы поддержки принятия решений, включающей ЛПР и большую языковую модель в качестве ассистента.

Статья состоит из двух частей. В первой части описываются первые две задачи. В частности, выделены три класса философских моделей объяснения, принятых в методологии науки. Выделенные классы моделей объяснения позволяют понять, является ли текст, сгенерированный большой языковой моделью, объяснением с научной точки зрения, и классифицировать эти тексты в выделенные классы моделей объяснения. Здесь также определены роль, цель и задачи объяснения в процессах принятия решений в условиях неопределенности. Приводятся сведения о текущей производительности больших языковых моделей ведущих иностранных и российских производителей. Это позволяет судить о возможности применения языковых моделей в качестве ассистента ЛПР в процес-

* Компания Meta Platforms, Inc. признана в России экстремистской, ее деятельность запрещена на территории Российской Федерации.



сах принятия решений. Определены возможности настройки большой языковой модели с помощью техник промптов (подсказок языковой модели) на генерацию объяснений в терминах выделенных классов моделей объяснения и задач объяснения, решаемых в условиях неопределенности.

Цель первой части статьи заключается в том, чтобы дать основные понятия и определения, которые будут использованы во второй части. Эта часть будет также полезна тем, кто впервые попытается применить большую языковую модель в своей системе поддержки принятия решений.

Во второй части статьи рассматривается вопрос измерения качества когнитивной системы поддержки принятия решений, включающей человека (ЛПР) и языковую модель. Для определения качества такой системы нужно измерить латентную переменную – удовлетворенность ЛПР объяснением языковой модели. Сформулированы критерии оценки степени удовлетворенности ЛПР при решении задач объяснения, обозначенных в первой части статьи. Предложена модель разумного респондента. Приводится пример оценки объяснений с использованием предложенной модели разумного респондента для двух российских больших языковых моделей (модель Сбера GigaChat 2.0 и модель Яндекса Yandex GPT 5 Pro).

1. МОДЕЛИ ОБЪЯСНЕНИЯ В МЕТОДОЛОГИИ НАУКИ

В философских, социологических, математических словарях и энциклопедиях можно найти различные определения объяснения, начиная от его свойств, структуры и заканчивая его ролью в научном познании. Приведем некоторые обобщенные определения.

В философской энциклопедии [28] объяснение – это рассуждение, посылки которого содержат информацию, достаточную для выведения из нее описания объясняемого явления. Объяснение представляет собой ответ на вопрос: почему данное явление происходит?

Функция объяснения определяется как функция научного познания, раскрытие сущности изучаемого объекта; она осуществляется посредством постижения закона, которому подчиняется данный объект, либо путем установления тех связей и отношений, которые определяют его существенные черты [29].

В методологии науки объяснение – это познавательная процедура, направленная на обогащение и углубление знаний о явлениях реального мира посредством включения этих явлений в структуру определенных связей, отношений и зависимостей,

дающих возможность раскрыть существенные черты данного явления [30].

В философии и методологии науки ставится задача создания единой универсальной теории объяснения, применимой в различных областях человеческой деятельности, например, в ходе получения знаний в рамках таких научных дисциплин, как физика, химия, биология, социология и др. Вариантами создаваемой общей теории объяснения являются различные модели объяснения, предложенные разными исследователями или коллективами исследователей. Разрабатываемые модели объяснения основываются на многочисленных примерах созданных теорий реального мира и в целом эти модели отражают механизмы работы человеческого интеллекта в процессах исследования, а также понимания и обоснования эмпирических наблюдений и фактов [31].

Важно понять, насколько искусственные большие языковые модели способны объяснять реальные объекты или события в процессах принятия решений в условиях неопределенности.

Необходимо отметить, что в настоящее время есть множество тестовых программ (бенчмарков), которые проверяют и оценивают способности языковых моделей в области общих знаний, логических рассуждений, здравого смысла и т. д. При этом актуальны исследование способностей и оценка качества языковой модели как системы объяснения в процессах поддержки принятия решений.

Рассмотрим основные модели объяснения, предложенные в философии науки.

1.1. Дедуктивно-номологическая модель объяснения

Согласно этой модели, научное объяснение состоит из двух основных элементов: экспланандума (*explanandum*), или объясняемого, – предложения, «описывающего явление, которое следует объяснить», и эксплананса (*explanans*), или объяснения, – «предложений, приводимых для объяснения явления» [32]. Чтобы объяснения (эксплананс) объясняли объясняемое (экспланандум), требуется выполнение нескольких условий.

- Объясняемое (экспланандум) должно быть логическим следствием объяснения (эксплананса), и предложения, из которых состоит эксплананс, должны быть истинными [32]; т. е. объяснение должно иметь форму дедуктивного аргумента, в котором экспланандум (объясняемое) является выводом из посылок, составляющих объяснение (эксплананс). В этом заключается «дедуктивный» компонент дедуктивно-номологической модели (ДН-модели).

• Эксплананс должен содержать по меньшей мере один «закон природы», который должен быть необходимой для вывода посылкой в том смысле, что при исчезновении этой посылки выведение экспланандума не будет валидным. Это номологический компонент модели (номологический – это философский термин, который, по сути, означает «законный») [32].

Формально постановка задачи объяснения для этой модели и для всех далее рассматриваемых моделей в самом общем виде выглядит следующим образом.

Модель объяснения – это кортеж $\langle \Lambda, Ex1, O, Ex2 \rangle$, где

• Λ – это множество законов; К. Гемпель [32] выделял несколько классов законов:

○ всеобщие законы, которые относятся ко всем областям знания L_c , – например, законы логики и математики;

○ частные законы отдельных наук L_s – например, законы химии, физики, биологии, социологии, политики и др.;

○ единичные факты L_f , они подчиняются более общим законам.

Таким образом, в ДН-модели множество законов Λ включает три определенных выше подмножества законов, т. е. $\Lambda = \{L_c, L_s, L_f\}$.

Понятие закона многозначно (закон природы, государственный закон и т. д.), поэтому приведем ряд определений, на которые в дальнейшем будем опираться в своих рассуждениях. Законом в философии называется необходимая связь (взаимосвязь, отношение) между событиями, явлениями, а также между внутренними состояниями объектов, определяющая их устойчивость, развитие, стагнацию или разрушение. В философском смысле под законом подразумевают «объективные связи явлений и событий, существующие независимо от того, известны они кому-нибудь или нет» [33].

Закон – это утверждение, выраженное словесно или математически, описывающее объективно существующие соотношения, связи между различными научными явлениями и объектами [28]. Закон предлагается в качестве объяснения фактов и признается на определенном этапе научным сообществом согласующимся с ними. Закон, справедливость которого была установлена не из теоретических соображений, а из опытных данных, называют эмпирическим законом.

• $Ex1$ – объясняемое, т. е. ситуация или явление которое необходимо объяснить.

• $O = \{o_j\}$ – множество фактов, которые характеризуют объясняемую ситуацию $Ex1$.

• $Ex2$ – объяснение.

Определение 1. ДН-объяснением некоторого явления или ситуации $Ex1$ будем называть отображение следующего вида: $DN: L_1(o_1), \dots, L_k(o_k) \rightarrow Ex1$, где $L_i(\cdot)$ – закон, $L_i(\cdot) \in \Lambda$; o_j – факты, они же переменные закона $L_i(\cdot)$, все факты истинны, т. е. вероятность $Pr(o_j) = 1$; DN – процедура дедуктивного вывода из законов L_i и фактов o_j объясняемого явления $Ex1$. ♦

Объяснение, удовлетворяющее определению 1, называют каузально-номологическим объяснением. Это определение предполагает применение в объяснении детерминированного закона. Однако законы из областей квантовой физики, биологии (генетики), а также социологии и других гуманитарных наук имеют вероятностный характер. Поэтому в рамках ДН-объяснений определено дедуктивно-статистическое и индуктивно-статистическое объяснение.

Определение 2. Если в ДН-объяснении при определении объясняющих фактов в качестве закона выступает общий статистический закон $L_i^*(\cdot) \in \Lambda$, такое объяснение называется дедуктивно-статистическим (SN -объяснением). ♦

В этом случае объясняемый объект или ситуация $Ex1$ может включать статистические характеристики случайного процесса (выборки и т. д.), математическое ожидание $M(o_j)$ и среднеквадратическое отклонение $\sigma(o_j)$, и должна быть определена номологическая вероятность объяснения $Ex2(M(o_j), \sigma(o_j))$.

В этом случае отображение объяснения из определения 1 запишем следующим образом:

$$SN: L_1(o_1), \dots, L_i^*(M(o_j), \sigma(o_j)), \dots, \\ L_k(o_k) \rightarrow Ex2(M(o_j), \sigma(o_j)), L_i^*(\cdot) \in \Lambda.$$

Определение 3. Если в ДН-объяснении хотя бы один из фактов получен в результате индуктивного рассуждения (определена субъективная вероятность этого факта), то объяснение называется индуктивно-статистическим (IN -объяснением). ♦

Считается, что индуктивный факт должен иметь определенную субъективную вероятность. Обычно порог вероятности устанавливают на уровне 0,5, т. е. $o_j | Pr(o_j) > 0,5$.

В этом случае отображение объяснения из определения 1 запишем следующим образом:

$$IN: L_1(o_1), \dots, o_j | Pr(o_j) > 0,5, \dots, \\ L_k(o_k) \rightarrow Pr(Ex1), L_i(\cdot) \in \Lambda$$

В отличие от дедуктивно-статистических объяснений, которые формируются на основе большого количества наблюдений случайных величин и могут считаться реальным статистическим законом, индуктивные заключения формируются на



основе малого числа наблюдений и субъективны. В ДН-модели объяснения индуктивные факты не являются надежными без указания контекста их получения и оценки вероятности. Дедуктивно-статистические и индуктивно-статистические объяснения К. Гемпель [32] называет статистически релевантными и, соответственно, индуктивно релевантными.

В ДН-модели объяснения предложены критерии валидности объяснения. Первый критерий – это критерий обобщения, а второй – критерий симметричности объяснения. Оставаясь в рамках математических формулировок ДН-модели, определим обобщение элементов объяснения. В самом общем виде обобщение – это замена объясняемого понятия $Ex1$ и объясняющих фактов (o_1, \dots, o_k) именами классов или категорий, к которым они принадлежат. Обозначим имя класса объясняемого понятия $Kl(Ex1)$, а имена классов фактов через $Kl(o_j)$.

Определение 4. Обобщением ДН-модели объяснения будем называть объяснение, полученное путем замены имени понятия $Ex1$ на имя его класса $Kl(Ex1)$ и замены имен фактов (o_1, \dots, o_k) на имена их классов $(Kl(o_1), \dots, Kl(o_k))$. ♦

Критерий валидности объяснения 1: ДН-объяснение $DN: L_1(o_1), \dots, L_k(o_k) \rightarrow Ex2$ (см. определение 1) валидно (верно), если истинно его обобщенное ДН-объяснение $DN: Kl(L_1(o_1)), \dots, Kl(L_k(o_k)) \rightarrow Kl(Ex2)$. ♦

Формально доказать валидность объяснения на основе обобщения можно, если закон, на основе которого строится объяснение, относится к общим законам L_c и имеется математическая модель объекта или ситуации. В этом случае может существовать строгое математическое или логическое доказательство ДН-объяснения, например, в логике предикатов и др. Однако во многих случаях, к примеру, в области гуманитарных наук, таких моделей просто не существует или они слишком абстрактны. Тогда этот критерий валидности объяснения может быть подтвержден или опровергнут экспертным способом, подстановкой в объяснение имен классов понятий объясняемого и фактов. Этот критерий позволяет утверждать о верности и всеобщности выбранного для объяснения закона.

Другой критерий валидности объяснения – это критерий симметричности логического вывода.

Критерий валидности объяснения 2: ДН-объяснение $DN: L_1(o_1), \dots, L_k(o_k) \rightarrow Ex2$ (см. определение 1) валидно (верно), если истинно обратное ДН-объяснение, т. е. $DN: Ex2 \rightarrow L_1(o_1), \dots, L_k(o_k)$. ♦

Существование обратного вывода объяснения говорит о его непротиворечивости. Формально

проверить этот критерий валидности можно в случаях, когда для объяснения используется известный всеобщий закон, известна модель объекта или ситуации. Для этого можно опираться на математический аппарат теории доказательств [34]. В других случаях возможна экспертная проверка непротиворечивости дедуктивного вывода, приведенного в объяснении.

Модель ДН-объяснения считается классической моделью научного объяснения объектов или ситуаций. Однако эта модель объяснения при условии соблюдения критериев валидности объяснения получается достаточно громоздкой и трудно обозримой. Полное и подробное объяснение, полученное с помощью ДН-модели, считается идеальным научным объяснением, которое называют «скрытой структурой» объяснения.

У философов существует точка зрения, согласно которой подробное объяснение избыточно, и поэтому считаются допустимыми так называемые «наброски объяснения» [32]. Допускается, что неполное объяснение, полученное, например, путем обобщения идеального объяснения, будет содержать объясняющий идеальное объяснение компонент. В данном случае речь идет об объяснении идеального объяснения в понятных терминах, например, «здравого смысла». Здравый смысл (*common sense*) – это совокупность навыков, форм мышления, взглядов на окружающую действительность, выработанных и используемых человеком в повседневной практической деятельности, которые разделяют окружающие люди.

Обычно объяснение считают чем-то, обеспечивающим понимание. В связи с этим одной из задач теории объяснения является выявление тех структурных особенностей объяснений, которые обеспечивают понимание. Понимание – это универсальная операция мышления, связанная с усвоением нового содержания, включением его в систему устоявшихся идей и представлений [35]. Считается, что наброски объяснения в терминах здравого смысла могут обеспечить понимание явления.

Будем считать здравым смыслом некоторые эмпирические законы, принятые обществом (социумом). Добавим в ранее рассмотренное в ДН-модели множество законов Λ законы здравого смысла L_{cs} , $\Lambda = \{L_c, L_s, L_j, L_{cs}\}$.

Тогда объяснение в рамках стратегии скрытой структуры представим следующим образом.

Определение 5. Объяснением $Ex2^*$ скрытой структуры ДН-объяснения $Ex2$ некоторого объекта или ситуации будем называть отображение следующего вида: $DN^*: L_{cs}^1(L_1(o_1)), \dots, L_{cs}^k(L_k(o_k)) \rightarrow Ex2^*$, где $L_{cs}^j(\cdot)$ – законы здравого смысла $L_{cs}^j(\cdot) \in \Lambda$; $Ex2^*$ –

объяснение здравого смысла; DN^* – процедура дедуктивного вывода из законов здравого смысла и фактов объясняемого явления. ♦

Объяснение скрытой структуры можно рассматривать как интерпретацию идеального объяснения ДН-модели. Проверить валидность объяснения скрытой структуры можно, если между основными законами L_c или частными законами L_s и законом здравого смысла L_{cs} определено отношение гомоморфизма, т. е. $\Phi: L_c \rightarrow L_{cs}$ или $\Phi: L_s \rightarrow L_{cs}$. При гомоморфизме элементы и отношения законов L_c, L_s отображаются в элементы и отношения законов здравого смысла L_{cs} при сохранении каузальных отношений. При гомоморфизме обратное отображение не однозначно.

При этом критерии валидности 1 и 2 могут не выполняться. Будем называть такое объяснение эмпирическим объяснением.

Философы считают, что объяснения в терминах законов здравого смысла (объяснения на подобию, аналогии, метафорах и т. д.) играет важную вспомогательную роль при формулировании идеального научного объяснения. Этот факт можно объяснить неоднозначностью обратного отображения $\Phi^{-1}: L_{cs} \rightarrow \{L_{ci}\}$, т. е. из объяснения, основанного на законах здравого смысла (L_{cs}), можно получить множество строгих ДН-объяснений, основанных на общих законах $\{L_{ci}\}$. Это создает возможность выбора лучшей модели объяснения и активизирует мышление исследователя.

1.2. Унификационистская теория объяснения

Следующая модель объяснения основана на унификационистской теории, смысл которой заключается в том, что научное объяснение может быть представлено в виде единого описания различных объектов и явлений. В этой теории определено схематическое предложение следующего вида [36]: «Для всех X , если X есть O и A , то X есть P .» Здесь X – это переменная (объясняемое); O, A, P – это переменные, факты, например свойства объясняемого.

Левая часть этого предложения – посылка ($Prm = (\forall X(O, A))$), а правая – это вывод – объяснение ($Inf = (X \rightarrow P)$). Автор работы [36] вводит понятие схематических аргументов – это последовательности схематических предложений и классификатор, позволяющий определить, какие аргументы являются посылками, а какие выводами и правилами вывода.

Определено также и правило заполнения (замены) абстрактных переменных (X, O, A, P) предмет-

ными переменными (понятиями предметной области). Вводится определение паттерна аргумента, включающего аргументы, их классификацию и инструкцию по заполнению. Считается, что заполненный предметными переменными паттерн является объяснением. Конечно, можно получить множество паттернов объяснения, но какое из них верно?

Автор вводит понятие объяснительного резерва [36] – набора паттернов аргументов, включающего набор убеждений, разделяемых учеными в конкретный период времени. Для доказательства, что конкретный вывод – паттерн объяснения – является здравым или приемлемым объяснением, нужно будет показать, что он принадлежит к объяснительному резерву.

Процедура вывода в данном случае заключается в подстановке в паттерн аргумента значений предметной переменной и проверке заполненного паттерна на соответствие объяснительному резерву. Если такое соответствие есть, то объяснение получено. В противном случае паттерн заполняется новыми данными. Строго говоря, схематическое предложение этой модели является описанием вывода в логике первого порядка. В логике первого порядка в аргументе посылки явным образом присутствует некоторый закон, на основании которого осуществляется вывод. Например, в утверждении «все люди смертны» (это закон), «Сократ – человек» (это факт), «Сократ смертен» (это заключение – вывод) вывод получен на основе закона.

В унификационистской модели объяснения явной ссылки на некоторый закон нет. Ссылка в объяснении на закон обеспечивает важное свойство объяснения – его каузальность. Это вызвало философские споры о корректности такой модели объяснения. Однако автор работы [36] приводит, во-первых, примеры объяснений, которые не содержат каузальных отношений, а во-вторых, он считает, что паттерны аргументов посылок и выводов должны быть обобщением исторического опыта человека и общества в различных сферах деятельности, отражать его культурные традиции. В результате длительного эволюционного процесса в социуме формируются убеждения, включающие в том числе и каузальные отношения. Все эти общественные убеждения включаются в объяснительный резерв предложенной модели и используются для валидации паттерна объяснения. В целом эта модель объяснения не лишена недостатков, но если, соглашаясь с автором этой модели, будем считать общественные убеждения некоторым эмпирическим социальным законом L_{so} и добавим их во множество законов Λ классической модели объяс-



нения, то получим значительное расширение модели объяснения, $\Lambda = \{L_c, L_s, L_f, L_{cs}, L_{so}\}$.

Определение 6. Эмпирическим объяснением некоторого объекта или ситуации $Ex1^{emp}$ будем называть отображение вида $DN^{emp}: L_{so}^1(o_1), \dots, L_{so}^k(o_k) \rightarrow Ex2^{emp}$, где $L_{so}^i()$ – эмпирический социальный закон $L_{so}^i(\cdot) \in \Lambda$; o_j – факты, переменные закона $L_{so}^i(\cdot)$, все факты истинны, вероятность $Pr(o_j) = 1$; $Ex2^{emp}$ – объяснение, основанное эмпирических законах; DN^{emp} – процедура подстановки эмпирического объяснения $Ex2^{emp}$ из множества объяснительного резерва ER , $Ex2^{emp} \in ER$. ♦

Критерии валидности 1 и 2 для эмпирического социального объяснения в данном случае можно применять экспертным способом. Однако из-за эмпиризма закона L_{so} нельзя ничего сказать (даже дать вероятностную оценку объяснения) о достоверности объяснения. В данном случае выполнение этих критериев будет свидетельствовать о корректности и непротиворечивости вывода объяснения, но не о его достоверности. В качестве объяснительного резерва могут выступать разнообразные справочники и энциклопедии.

1.3. Прагматические теории объяснения

Ранее были рассмотрены модели объяснения, которые не учитывали роль человека в процессе объяснения. Существуют прагматические теории объяснения, в которых человек непосредственно включен как в процесс генерации объяснения, так и в процесс потребителя объяснения – в качестве потребителя.

Теория объяснения содержит «прагматические» элементы, если эти элементы требуют обязательной отсылки к фактам об интересах, убеждениях или других психологических характеристиках тех, кто дает объяснение или его получает, и обязательна отсылка к контексту, в котором возникает объяснение [37].

Авторы классической ДН-модели объяснения согласны с тем, что прагматические элементы играют какую-то роль в процессе предоставления или получения объяснений, однако они полагают, что существует непрагматическое ядро объяснения, описание которого является главной задачей теории объяснения [32].

Современные прагматические теории объяснения основаны на конструктивном эмпиризме, изложенном в работе американского философа Б. ван Фраассена [37]. Он считает, что цель науки состоит в построении «эмпирически адекватных» теорий, дающих истинные или верные описания доступных наблюдению явлений, а научная дея-

тельность является скорее конструированием, чем открытием: конструирование моделей, которые должны быть адекватны явлению, а не открытие истины, имеющей отношение к ненаблюдаемому [37]. «Цель науки – дать нам теории, которые являются эмпирически адекватными и принятие теории включает, как веру, только то, что она эмпирически адекватна» [37]. Под «эмпирической адекватностью» Б. ван Фраассен имел в виду совпадение эмпирических проявлений теоретической модели явления и самого явления.

Теория конструктивного эмпиризма дополняет теорию научного реализма, основанную на строгих логических выводах, законах и доказательствах. Однако, как считает Б. ван Фраассен, строгость логических выводов в теории научного реализма основывается на недоказанных постулатах и аксиомах, что ограничивает ее применение в гуманитарных сферах. При построении теории конструктивного эмпиризма автор ссылается на модель семиотики – науки о знаках и знаковых системах, в которой наблюдаемое явление представлено на трех уровнях: синтаксическом, семантическом, прагматическом [38].

Строгое логическое объяснение, полученное в рамках теории научного реализма, можно представить как синтаксическую абстрактную структуру. В конструктивном эмпиризме осуществляется семантическая (смысловая) интерпретация таких логических структур.

В конструктивном эмпиризме человек – это наблюдатель проявлений действительности, конструктор эмпирической модели действительности и ее валидатор. Следовательно, «принятие» теории означает всего лишь убеждение в ее эмпирической адекватности [37]. Вопросы применения конструктивного подхода для исследования социальных систем рассмотрены в работе [39].

В рамках теории конструктивного эмпиризма предложена модель объяснения. Эта модель объяснения представляется в виде тройки:

$$Q = \langle T_k, X, RL \rangle,$$

где Q – вопрос; T_k – тема или контекст; $X = \{X_1, \dots, X_n\}$ – контрастные классы ответов, по сути альтернативы ответов; RL – отношение релевантности темы и альтернативного ответа.

Поясним эту модель. Объяснение – это не просто набор суждений, а всегда ответ на вопрос «почему?», который всегда возникает в определенном контексте и определяется тремя факторами. Первым фактором является тема – то, о чем вопрошается (T_k). Вторым фактором будет выступать контрастный класс, представляющий собой класс, состоящий из множества альтернативных теме вы-

сказываний. Третий фактор – отношение релевантности между темой вопроса и контрастным классом, устанавливающее, что может выступить в качестве объяснения.

Пусть для конкретной темы T_k и заданного множества альтернативных ответов X может быть образовано множество пар – тема и альтернативный ответ – путем прямого произведения $T_k \times X$.

Определение 7. Для пары $(T_k, X_j) \in T_k \times X$, $X_j \in X$, существует отношение релевантности $RL \subseteq T_k \times X$, если субъект верит и убежден, что X_j отвечает на вопрос Q в контексте T_k . В этом случае объясняемое Q и объяснение X_j эмпирически адекватны. ♦

Высказывание X_j является релевантным вопросу Q тогда и только тогда, когда существует отношение RL для пары (T_k, X_j) . Установление отношения релевантности не может быть задано однозначно. Например, рассмотрим вопрос $Q =$ «Почему кровь циркулирует в теле?» и два возможных варианта ответа на него: $X_1 =$ «потому что сокращение сердца заставляет кровь двигаться по артериям»; $X_2 =$ «чтобы доставить кислород ко всем тканям организма».

Видно, что есть две альтернативы ответов (два контрастных класса), но вне контекста постановки вопроса нельзя установить отношение релевантности между темой вопроса Q и контрастным классом X [37].

Вопрос «почему?» фиксирует некую проблему и задает определенный контекст. При этом ответ на него включает и теоретический контекст – научное объяснение. Для данного контекста вопроса эмпирически адекватным будет ответ $X_1 =$ «потому что сокращение сердца заставляет кровь двигаться по артериям». Здесь имеет место каузальное отношение релевантности. Однако если поменять контекст вопроса $Q =$ «Зачем кровь циркулирует в теле?», то эмпирически адекватным будет ответ $X_2 =$ «чтобы доставить кислород ко всем тканям организма», с функциональным отношением релевантности.

Считается, что «научное объяснение относится не к сфере чистой науки, а к сфере применения науки. Оно представляет собой использование науки для удовлетворения некоторых наших желаний; это всегда конкретные желания, возникающие в конкретном контексте, но они всегда являются желаниями получения определенной дескриптивной информации... Точное содержание желания и оценка того, насколько хорошо оно удовлетворено, разнятся от контекста к контексту» [37].

Обобщая рассмотренные модели объяснения, отметим основные отличия прагматических моде-

лей объяснения от семейства ДН-моделей. Напомним, что семейство ДН-моделей объяснения (в литературе их называют классическими) разрабатывалось в рамках теории научного реализма, и такие модели характеризуются следующими свойствами:

- объяснение основывается на законах (детерминированных, статистических, эмпирических),
- логический вывод объясняемого через наблюдаемые факты основан на законах и должен быть корректным и истинным (достоверным),
- объяснение отражает объективные законы природы и (или) общества и не зависит от психологических особенностей субъекта (его интересов, убеждений, желаний, оценок и т. д.).

Прагматические теории объяснения основаны на теории конструктивного эмпиризма и характеризуются следующими свойствами:

- субъект включен в модель объяснения. Объяснение формируется с учетом его психологических особенностей и контекста, который он формулирует в вопросе;
- ответ формируется на основе субъективной эмпирической адекватности объяснения. Утверждение об эмпирической адекватности объяснения гораздо слабее утверждения истинности (достоверности) объяснения в моделях ДН-объяснений. Однако эмпирическая адекватность объяснения позволяет удовлетворить исследовательские потребности субъекта;
- объяснение содержит научно обоснованные утверждения на уровне семантики – смысла объясняемого, а не логического вывода в терминах, доступных ограниченному кругу узких специалистов, как это делается в рамках теории научного реализма;

- объяснение направлено на удовлетворение потребностей, желаний конкретного субъекта в интересном ему контексте для получения дополнительной релевантной информации и расширения его мировоззрения – ментального пространства.

Эти два класса моделей образуют конкурирующие теории объяснения, но в процессах поддержки принятия решений могут дополнять друг друга.

Отметим, что в литературе приводятся модели объяснения для разных предметных областей, например модели объяснения в социологии, технике, математике, модели объяснения сознания и др. Выше были рассмотрены некоторые философские модели объяснения из методологии научного познания. Другие модели научного объяснения можно найти в философской литературе, например в обзорной работе [9]. Задача приведенного здесь краткого обзора заключается в том,



чтобы выделить наиболее известные, качественно отличающиеся классы моделей объяснения с тем, чтобы в дальнейшем можно было экспертным путем классифицировать тексты, сгенерированные большой языковой моделью, в один из классов научного объяснения. Это позволит оценить качество объяснения языковой модели и, следовательно, ее полезность для решения задач поддержки принятия решений.

1.4. Цель объяснения в поддержке принятия решений

Отметим, что, литература, посвященная различным конкурирующим моделям объяснения, достаточно обширна. Здесь исследователи по-разному трактуют само понятие объяснения, однако достаточно мало уделяют внимания вопросу о целях объяснения – тому, для чего используются объяснения [31]. Очевидно, что цель объяснения определяется целью научного исследования, которое зависит от объекта исследования. Часто предлагаемыми в самом общем виде целями исследования считаются: доказательство, исследование структуры, принципа функционирования, прогноз развития, управление сложным объектом и др. Цель объяснения также зависит от области исследования; например, в практике обучения математике выделяют такие виды объяснения: объяснение – обоснование, объяснение решения задачи, объяснение – раскрытие смысла и объяснение – верификация.

Однако объединяющим началом разных моделей объяснения в разных предметных областях принято считать связь объяснения и понимания. Считается, что объяснение обеспечивает понимание. Под пониманием подразумевается универсальная операция мышления, связанная с усвоением нового содержания (новое содержание включено в объяснение), включением его в систему устойчивых идей и представлений (по сути, в систему знаний субъекта) [35].

Фактически объяснение – это описание основных свойств исследуемого объекта или ситуации с возможной привязкой их к известным законам, которое запускает универсальную операцию мышления – понимание, обеспечивающее встраивание этой новой информации в уже имеющиеся у субъекта знания.

В работе [40] автор, связывая понимание и объяснение, считает, что тот, «кто понимает объяснение, постоянно и легко переходит от реального объекта к его «идеальной» модели (именно «идеальная» модель объекта создается абстракциями) и обратно», может повторить все познавательные операции объясняющего и осознать, почему про-

изведены именно эти операции и в такой последовательности. Таким образом, объяснение обеспечивает связь реального мира с ментальными процессами мышления человека и направлено на формирование абстрактной идеализированной ментальной модели объекта или ситуации. Затем ментальная модель может быть выражена на строгом математическом языке и стать законом.

В этом случае в аспекте поддержки принятия решений можно несколько иначе сформулировать цель объяснения. Во-первых, объяснение направлено на формирование информационной среды об объекте или ситуации, включающей параметры объекта или явления, структуру, принадлежность к определенному классу (классификацию) и др. в ментальной сфере субъекта. В этом случае работают разнообразные отношения релевантности объяснения, отвечающие потребностям субъекта, и нестрогая эмпирическая адекватность объяснения.

Во-вторых, объяснение направлено на формирование или выявление каузальных отношений, строгий логический вывод на основе известных или вновь открытых законов или закономерностей для принятия решения по управлению объектом или ситуацией.

Таким образом, цель объяснения в процессах поддержки принятия решений имеет две составляющие:

- исследовательскую, направленную на формирование информационной среды для принятия решений и выработки альтернатив решений;

- практическую, направленную на принятие, обоснование и реализацию альтернатив решения по управлению объектом или ситуацией.

В этом случае конкурирующие модели объяснений – классическая (ДН-модель) и прагматическая дополняют друг друга. Если рассматривать процесс принятия решений как исследовательский процесс, то можно определить два этапа этого процесса. Первый этап – исследовательский – заключается в серии последовательных вопросов и ответов-объяснений, он позволяет сформировать необходимую информационную среду для выработки альтернатив принятия решений. На этом этапе уместны прагматические модели объяснения, не ограничивающие желания и потребности исследователя и позволяющие получить эмпирически адекватные объяснения альтернатив решения.

На втором этапе осуществляется выбор лучшей альтернативы с использованием ДН-модели объяснения, позволяющей обосновать альтернативы ссылками на известные законы и строгий логический вывод.

2. БОЛЬШИЕ ЯЗЫКОВЫЕ МОДЕЛИ И ОЦЕНКА ИХ СПОСОБНОСТЕЙ

Большие языковые модели представляют собой нейронные сети глубокого обучения, которые обучаются на гигантских объемах текстовых данных. В их основе лежит архитектура трансформера, которая включает набор нейронных сетей, состоящих из энкодера и декодера. Трансформер – это архитектура нейронной сети, направленная на решение задачи обработки естественного языка и многих других задач машинного обучения [41]. Энкодер – это элемент архитектуры трансформера, преобразующий корпус текста в векторные представления, в которых каждое слово представлено в виде вектора вероятностей его совместного употребления с другими словами корпуса текста, при этом сохраняется информация о структуре и взаимосвязях слов. Декодер использует закодированную информацию для генерации ответа или выполнения предсказаний. Он «раскодирует» полученные данные, создавая новые последовательности текста, учитывая предыдущие слова, контекст и вероятности совместного употребления слов в этом контексте.

Нейронная сеть большой языковой модели представляется в виде многослойной структуры, в которой каждый слой состоит из множества искусственных нейронов, связанных с нейронами соседних слоев. Число слоев в разных больших языковых моделях отличается и может достигать нескольких десятков или больше. В результате обучения нейронной сети большим корпусом текста формируется словарь токенов – отдельных букв или их групп, слов, фраз, предложений, для которых определена вероятность их совместного употребления в определенном контексте. При кодировании входной информации и передачи ее от слоя к слою используется механизм внимания, позволяющий выявить важные элементы структуры языка. Механизмы внимания могут выделять различные типы синтаксических отношений между словами, фактически выделяя синтаксическую структуру предложения [42]. В результате обучения образуется векторное пространство контекстных векторов токенов, для которых определена (косинусная) мера близости токенов. При декодировании для ответа на вопрос из векторного пространства выбираются токены (слова, фразы, предложения), близкие к теме вопроса, из которых выстраиваются предложения – ответ на вопрос. Выбор из векторного пространства токенов, близких к токенам вопроса, позволяет сконструировать

ответ на поставленный вопрос из слов, часто встречающихся в контексте вопроса.

Таким образом, большая языковая модель формирует статистический ответ на поставленный вопрос, не работая на уровне семантики, не анализируя смысл вопроса и ответа. Смысл ответа зависит от качества корпуса текста. Если в корпус включались тексты из проверенных источников, то ответ будет по смыслу верен и понятен.

Если языковая модель не понимает вопроса или была обучена неполными или ошибочными данными, то она пытается угадать ответ, опираясь на имеющиеся синтаксические шаблоны, что может привести к ложным ответам. Такое поведение большой языковой модели называется галлюцинацией. Галлюцинация – это феномен больших языковых моделей, когда они генерируют фактически неверные или вымышленные данные, которые не основаны на реальной информации. Такое поведение свойственно и людям, когда в условиях дефицита информации или времени для изучения проблемы, они формируют случайные или вымышленные ответы, далекие от реальности.

Проблема галлюцинаций больших языковых моделей особенно актуальна в приложениях, в которых достоверная информация критически важна. В настоящее время разработаны различные метрики оценки склонности языковых моделей к галлюцинациям и методы уменьшения галлюцинаций языковых моделей, связанные: с созданием корректных корпусов текста (ручная проверка, чистка текста); формированием корректных запросов к языковой модели, подсказкой правильного ответа; дообучением языковой модели; поиском верной релевантной информации во внешних источниках информации.

Необходимо отметить, что галлюцинации языковой модели в случаях решения исследовательских и творческих задач могут стимулировать интуицию человека и привести к оригинальному решению. Напомним, что эвристические методы решения творческих задач, такие как мозговой штурм, синектика, не отвергают абсурдные и контрпродуктивные альтернативы решений, считая их стимуляторами интуиции, способной привести к новому и оригинальному решению.

Несмотря на этот недостаток, языковые модели находят применение благодаря их способностям решать многие задачи, представленные на естественном языке, которые ранее решал только человек.

Большие языковые модели оцениваются с помощью специальных программ – бенчмарков, которые измеряют их основные качества: объем знаний, точность ответов, надежность и др.



Бенчмарк использует определенные наборы данных, метрики и задачи оценки для тестирования языковой модели, что позволяет сравнивать различные модели и измерять их точность. Выделяют бенчмарки на проверку знаний, логического мышления, понимания прочитанного текста, здравого смысла и др. [43]. Разработано большое количество бенчмарков; рассмотрим некоторые из них.

Бенчмарки на проверку знаний тестируют модели в различных областях. Они оценивают, насколько эффективно модель может вспоминать информацию из разных сфер, таких как физика, география и т. д. Известный бенчмарк MMLU (*Measuring Massive Multitask Language Understanding*) создан для проверки уровня фактических знаний модели по различным темам, таким как гуманитарные науки, социальные науки, история, компьютерные науки и даже право. Он включает 57 вопросов и 15 тыс. задач, направленных на то, чтобы убедиться в высоких способностях языковой модели. На этом бенчмарке языковая модель GPT-4-omni правильно ответила на 88,7 % заданных ей вопросов.

Бенчмарки на проверку логического мышления тестируют способности модели «думать» пошагово и делать логические выводы.

Используются бенчмарки оценки математических способностей языковых моделей. Например, тест GSM8K состоит из 8,5 тыс. задач по математике для средней школы. Решение этих задач требует выполнения моделью нескольких шагов последовательности элементарных вычислений. Языковые модели, специально обученные для математического рассуждения, показывают хорошие результаты на этом бенчмарке – например, модели GPT-4 достигают 96,5 % точности.

Бенчмарк вопросов и ответов (англ. *Graduate-Level Google-Proof Q&A Benchmark*, GPQA) уровня аспирантуры оценивает логическое мышление языковой модели, используя набор данных всего из 448 вопросов. Этот сложный тест, разработанный экспертами из областей биологии, физики и химии, языковая модель GPT-4-omni проходит, набирая лишь 53,6 % точности, в то время как аспиранты достигают 65 %.

Исследования по разработке рассуждающих языковых моделей ведутся многими основными конкурирующими производителями языковых моделей. Так, в 2024 г. компания OpenAI выпустила новую языковую модель OpenAI-o1, которая демонстрирует отличные результаты в сложных рас-

суждениях, превосходя людей в тестах по математике, кодированию и естественным наукам. На отборочных экзаменах Международной математической олимпиады (англ. *International Mathematical Olympiad*, IMO) эта модель правильно решила 83 % задач, тогда как ее предшественница GPT-4o дала лишь 13 % верных ответов. Разработчики из OpenAI утверждают, что при решении сложных контрольных задач по физике, химии и биологии модель демонстрирует результаты, сопоставимые с результатами аспирантов.

Бенчмарки на понимание прочитанного текста тестируют способности модели интерпретировать естественный язык и генерировать соответствующие ответы. Тестирование заключается в получении ответов на вопросы по текстам, что позволяет оценить понимание, способность делать выводы и улавливать и не забывать важные детали.

Одним из бенчмарков для тестирования понимания прочитанного является тест DROP (*Discrete Reasoning Over Paragraphs*), который ставит перед моделями задачу выполнения рассуждений на основе анализа содержимого абзацев. Этот бенчмарк включает 96 тыс. вопросов для проверки способностей языковой модели к рассуждению. Вопросы DROP содержат информацию, которая требует от моделей выполнения математических операций, таких как сложение, вычитание и сравнение, на основе информации, разбросанной по всему тексту. Отвечая на эти сложные вопросы, языковая модель GPT-4 достигает точности 80 %, в то время как люди дают 96 % правильных ответов на наборе данных DROP.

Бенчмарки на проверку общих знаний (здравого смысла) оценивают способность модели к обобщенным знаниям о мире. Такие наборы тестов обычно включают вопросы, требующие для правильного ответа обширных энциклопедических знаний. Тестирование здравого смысла в языковых моделях оценивает способность модели делать суждения и выводы, соответствующие человеческому мышлению. Люди формируют целостное представление о мире через практический опыт, а языковые модели обучаются на огромных наборах данных, не понимая контекста.

Бенчмарк HellaSwag (*Harder Endings, Longer Contexts, and Low-Shot Activities for Situations with Adversarial Generations*) предназначен для проверки способности модели предсказывать правдоподобное продолжение определенного сценария. Те-

стирования методами HellaSwag показывают, что современные модели, такие как GPT-4, достигли уровня точности, близкого к человеческой.

Бенчмарк IFEval (*Inference and Fidelity Evaluation*) позволяет оценить как точность, так и качество сгенерированного текста. Сначала модель оценивается по показателю Inference – способности генерировать текст на большом объеме данных. Затем происходит оценка качества (Fidelity) сгенерированного текста. Оценка включает в себя проверку на соответствие сгенерированного текста ожидаемому результату, а также оценку степени сохранения смысла и структуры текста. Далее вычисляется итоговая оценка IFEval, которая учитывается как способность модели генерировать текст, так и качество этого текста. Чем выше оценка IFEval, тем лучше модель справляется с задачей генерации текста.

Бенчмарк HumanEval – это эталонный набор данных, предназначенный для объективной оценки качества кода, генерируемого моделями искусственного интеллекта на основании текстового описания задачи. Бенчмарк состоит из 164 задач по программированию, написанных вручную специально для этого набора данных, чтобы гарантировать их отсутствие в обучающих выборках моделей. Все задачи сформулированы на языке Python и представлены в виде фрагментов кода с описанием.

В табл. 1 представлены оценки производительности больших языковых моделей от ведущих разработчиков: GigaChat 2 MAX от российской компании Сбер; Qwen 2.5 72B от ведущей платформы электронной коммерции Alibaba; Llama 3.3 70B (разработчик – компания Meta AI); GPT-4o (разработчик – OpenAI); DeepSeek-V3 от китайской компании DeepSeek, принадлежащей фонду HighFlyer; Yandex GPT5 Lite Instruct от российской компании Яндекс. Тестирование проводилось на бенчмарках следующих категорий: общие знания, математика, работа с кодом, качество генерации текста. Данные актуальны [44, 45] на начало

2025 г. Цифры в таблице показывают процент решенных задач.

Тестирование способностей языковых моделей с помощью бенчмарков позволяет объективно оценить и сравнить качество моделей разных производителей, понять текущий уровень и динамику развития языковых моделей. Производители языковых моделей считают, что способности этих моделей приближаются к способностям человека, что позволяет им быть ассистентом в системах поддержки принятия решений.

Однако решение практических задач, например, по принятию решений в сложных экономических, политических или социальных ситуациях требует от ЛПП комплексного сочетания всевозможных способностей человеческого интеллекта. Поэтому автоматизированные метрики сами по себе не могут охватить весь спектр оценки языковой модели, особенно когда дело касается субъективных аспектов понимания и генерации языка. Здесь человеческая оценка является гораздо более точной.

В этом случае целесообразно для оценки привлекать экспертов или группу экспертов, способных дать, например, на основе опроса, точную и надежную оценку способностей больших языковых моделей [43].

В случае экспертного оценивания объяснений языковых моделей возникают некоторые проблемы. Дело в том, что необходимо оценить взаимодействие человека и большой языковой модели. Языковая модель выдает статистически правдоподобное объяснение, не понимая его смысла, а человек пытается понять объяснение и встроить его в собственную систему знаний. Понимание человеком объяснения языковой модели зависит от уровня его знаний, и поэтому разные люди, получая один и тот же ответ языковой модели, могут дать ему разные оценки.

В этом случае возникает задача разработки метода оценки индивидуальной (не групповой) удовлетворенности человека объяснениями ситуации

Таблица 1

Производительность больших языковых моделей от ведущих разработчиков

Категория	Название бенчмарка	GigaChat 2 MAX	Qwen 2.5 72B	Llama 3.3 70B	GPT-4o	DeepSeek-V3	Yandex GPT5 Lite Instruct
Общие знания	MMLU (RU)	80,46	78,30	65,08	80,00	73,74	70,0
	MMLU (EN)	86,00	83,85	78,57	88,70	85,24	75,8
Математика	GSM8K	95,68	95,07	92,87	95,00	94,99	87,9
	MATH	77,26	78,74	62,80	76,60	85,48	82,0
Работа с кодом	HumanEval	87,20	86,60	86,00	84,00	91,46	71,8
Качество генерации текста	IFEVAL (RU)	83,62	84,27	75,12	80,24	84,37	76,9
	IFEVAL (EN)	89,99	90,43	90,83	88,51	92,21	72,6



принятия решения, которые он получает от большой языковой модели. Решение данной задачи будет рассмотрено во второй части этой статьи.

3. МОДЕЛИ ОБЪЯСНЕНИЯ И БОЛЬШИЕ ЯЗЫКОВЫЕ МОДЕЛИ

Рассмотренные ранее модели объяснения образуют два класса моделей: классические (дедуктивно-номологические), основанные на строгом логическом выводе и известных объективных законах или закономерностях, и прагматические, учитывающие цели и желания исследователя или объясняющего, а также контекст вопроса, требующего объяснения.

В этих моделях сам человек формулирует объяснение, основываясь на наблюдаемых фактах и знаниях. Очевидно, что знания человека-исследователя, необходимые для формулировки ответа-объяснения, в каждом классе моделей объяснения будут разными. Бенчмарки показывают хорошие способности языковых моделей в разных областях. Способна ли языковая модель дать объяснение, которое удовлетворило бы ЛПР?

Знания языковой модели определяются содержанием корпуса текста, которым ее обучили. Тогда для работы языковой модели в классе классических моделей объяснения ее необходимо обучить известными законами логики, математики, физики, химии, биологии и т. д. В этом случае модель с большой вероятностью будет генерировать строгие логические объяснения, обоснованные известными законами. Достоверность таких выводов может быть проверена, а сами выводы и объяснения можно использовать для управления ситуацией.

Для работы языковой модели, функционирующей в классе прагматических моделей объяснения, ее нужно обучать общими знаниями, законами здравого смысла, традициями и др. Тогда будут получаться эмпирически адекватные объяснения, основанные, согласно теории конструктивного эмпиризма, на вере исследователя. Такие объяснения полезны для расширения ментального пространства ЛПР, повышения его осведомленности и стимулирования интуиции для генерации нетривиальных творческих решений.

Обучение языковых моделей для получения классических объяснений или прагматических объяснений с «чистого листа» – это очень затратные процедуры. Во-первых, это сбор, подготовка, очистка обучающего корпуса релевантного задаче текста очень больших размеров. Подготовку корпусов текстов, как правило, выполняют специаль-

ные коллективы, нанятые по аутсорсингу. В словарях современных языковых моделей количество токенов исчисляется миллиардами. Считается, что только у очень больших языковых моделей появляются эмерджентные свойства, такие как, например, способность к рассуждению. Во-вторых, для обучения нейронной сети требуются очень большие вычислительные мощности – суперкомпьютеры – и, соответственно, большие временные и энергетические затраты. Обучение языковой модели с «чистого листа» доступно крупным компаниям, которые позволяют пользоваться уже обученными моделями другим разработчикам для решения их собственных задач.

Уже обученные языковые модели, имеющиеся в открытом доступе, можно дообучить собственными текстами, ориентируя их работу на решение собственных конкретных задач. В этом случае этапы подготовки корпуса текста и обучение модели также присутствуют, но уже с меньшими затратами.

Большая языковая модель – это огромная нейронная сеть с огромным количеством связей между токенами. Каждый вопрос к сети активирует определенную цепочку токенов, формируя ответ. Вариантов активизации цепочек токенов тоже огромное количество, а, следовательно, количество правдоподобных альтернативных ответов также велико.

В настоящее время активно развивается альтернативный способ настройки большой языковой модели под собственные задачи с помощью так называемых промптов (англ. *Prompt Engineering*, PE) [46, 47]. Промпт – это подсказка или последовательность подсказок языковой модели. Промпты представляются в виде алгоритма или сценария решения пользовательской задачи, создаются пользователем и могут встраиваться в языковую модель. Таким образом, промпты управляют работой языковой модели для получения лучшего результата.

Техники промптов и модели объяснения

Рассмотрим несколько наиболее популярных техник создания промптов.

Zero-shot. Это техника создания промпта без использования примеров входных или выходных данных, используемых для дообучения языковой модели. Это обычный простой вопрос языковой модели, в ответ на который она может сообщить информацию, для получения которой не требуется многошаговая инструкция.

Few-shot. Это техника создания промпта с примерами, т. е. используются несколько приме-

ров входных и выходных данных, чтобы дообучить модель. Языковая модель, обучаясь на этих примерах, выдает ответы, следуя их шаблонам. Например, так можно дообучить модель для точного определения тональности (эмоциональной «окраски») анализируемого текста. Few-shot-промптинг может использоваться в качестве техники для обеспечения контекстного дообучения языковой модели. В примерах задается контекст, в котором рассчитывают получить ответ.

Role based. Это техника создания промпта, в основе которой лежит задание роли языковой модели, в которой она будет выступать при генерации ответа на поставленный вопрос. Например, можно попросить модель ответить на некоторый вопрос с точки зрения экономиста, финансиста, юриста и т. д. Вопрос один и тот же, а ответы будут сгенерированы в контекстах разных предметных областей [48].

Chain-Of-Thought (CoT). Это техника создания промпта, который заставляет языковую модель думать поэтапно, шаг за шагом. В данной технике пользователь задает языковой модели алгоритм решения задачи. Это делает размышления модели близкими к человеческим. Как правило, сложная задача декомпозируется на подзадачи, последовательное решение которых повышает точность общего решения. Применение этой техники в больших языковых моделях существенно повышает точность решения математических и логических задач известного математического бенчмарка GSM8K.

Авторы этой техники констатируют повышение качества решения лингвистических задач и задач общих знаний и здравого смысла в случае применения этого промпта. Эта популярная техника в разных модификациях широко используется в рассуждающих языковых моделях [49].

Chain-of-Verification. Это техника создания промпта, которая дополняет Chain-Of-Thought, заставляя модель проверять все предыдущие шаги перед тем, как сделать следующий шаг, делая рассуждающие языковые модели более надежными [50].

Chain-of-Note. Это техника создания промпта, который заставляет языковую модель делать так называемые «заметки» в процессе решения задачи, поясняя каждый последовательный шаг. Эта техника позволяет обнаружить галлюцинации языковой модели [51].

Chain-of-Knowledge. Это техника создания промпта, в который включаются проверенные знания о предметной области, что позволяет языковой модели использовать их для решения задачи. В

отличие от техники Chain-Of-Thought-промптинга, в этой технике языковая модель в ходе рассуждения опирается на известные факты, законы или закономерности, которые пользователь включает в промпт, и выстраивает из них логическую цепочку, приводящую к конкретному и обоснованному ответу. Например, для решения задачи по физике пользователь включает в промпт законы физики, на которые языковая модель будет опираться при генерации решения задачи. Корректность ответа будет выше, чем в простом Chain-Of-Thought [52].

Tree of Thoughts (ToT) [53]. Эта техника применяется для решения сложных исследовательских задач или задач стратегического планирования, когда традиционные или простые методы создания промптов оказываются недостаточными. В работе [53] предложен промпт Tree of Thoughts, который обобщает метод цепочки мыслей (CoT). Эта техника предлагает обращать внимание на исследование ответов, которые служат промежуточными шагами для общего решения проблем с помощью языковых моделей. Метод ToT поддерживает построение дерева ответов (дерева мыслей), в узлах которого расположены промежуточные тексты ответов. На построенном дереве предлагается применять алгоритмы поиска, например, поиска в ширину и поиска в глубину, для получения решения и его анализа.

ReAct Prompting. В работе [54] предложена техника ReAct, в которой для генерации цепочек рассуждений использовалась языковая модель в интерактивном режиме с человеком с целью формирования релевантных задаче действий. Генерация цепочек рассуждений позволяет модели создавать, отслеживать и обновлять планы действий, а также обрабатывать ошибочные ситуации. Действия позволяют взаимодействовать с внешними источниками информации, такими как базы знаний.

Промпт ReAct позволяет языковой модели взаимодействовать с внешними источниками для получения дополнительной информации, что приводит к более надежным и достоверным ответам. ReAct улучшает интерпретируемость и надежность языковой модели. В целом авторы обнаружили, что наилучшим подходом является использование ReAct в сочетании с цепочкой мыслей (CoT), что позволяет использовать как внутренние знания, так и внешнюю информацию, полученную в процессе рассуждения.

Промпты здравого смысла (Commonsense). В настоящее время большой исследовательский интерес представляют промпты для поддержки дедуктивных рассуждений на основе знаний здраво-



го смысла. Дело в том, что большие языковые модели обучаются на больших текстовых корпусах, а отвечая на вопросы, дают статистически лучшие ответы. Такие ответы верные, но они, как правило, не содержат элементы дедуктивного вывода, которые для человека кажутся очевидными и представляются в виде неявных знаний. Для получения от языковой модели ответа, содержащего объяснение с дедуктивным выводом, разрабатываются промпты, подсказки языковой модели для извлечения осмысленного ответа. Генерация такой подсказки языковой модели основана на использовании внешних баз знаний здравого смысла. Это статические базы знаний ConceptNet [55], ATOMIC [56] и динамическая база знаний здравого смысла CoMeT [54], в которой знания здравого смысла формируются из контекста вопроса к языковой модели и представляются в виде графа рассуждений. Эксперименты с этой динамической базой знаний [57] показали рост производительности языковых моделей на бенчмарках здравого смысла. В работе [58] представлен метод генерации знаний из самой языковой модели и предоставления этих знаний в качестве дополнительных входных данных (подсказки) при ответе на вопрос. Этот метод не требует доступа к структурированной базе знаний здравого смысла, но при этом повышает производительность современных больших языковых моделей. В работе [59] рассмотрены вопросы генерации языковой моделью ответов здравого смысла на основе внутреннего диалога с языковой моделью. В этом методе не используются внешние базы знаний, но оценки производительности языковой модели на бенчмарках здравого смысла с таким промптом показали рост. Однако, как считают авторы работы [59], недостаток больших языковых моделей – это отсутствие у них интроспекции, т. е. знаний о своих знаниях.

Выше были рассмотрены лишь некоторые техники промптов. В настоящее время множество разработчиков предлагают пользовательские промпты, которые могут встраиваться в уже обученные языковые модели и не требуют дополнительного дообучения языковой модели на предметной области. Это представляется недорогим вариантом подстройки языковой модели для решения пользовательских задач.

В поддержке принятия решений существуют эвристические приемы (методы) анализа ситуаций в условиях неопределенности, которые представляются как методики поиска решений в сложных ситуациях. Это, например, метод «пять почему», направленный на выявление причин проблемы; метод «Фишбоун», позволяющий построить иерархическую модель причинно-следственных

связей; метод гирлянд и ассоциаций, направленный на генерацию творческих решений; SWOT-анализ, позволяющий определить стратегию развития организации, и др. Все эти методики прошли апробацию в практике принятия решений, а алгоритмы их работы можно рассматривать как алгоритм работы промпта большой языковой модели.

Промпт можно представить как управляющую программу, которая позволяет извлечь в результате выполнения последовательных шагов из языковой модели необходимую для принятия решений информацию.

Рассмотрим, как перечисленные выше техники промптов могут помочь языковой модели в реализации рассмотренных моделей объяснения.

Для прагматических теорий объяснения, ориентированных на удовлетворение желаний и интересов пользователя, подойдут техники, позволяющие организовать диалог с языковой моделью для расширения ментального пространства и осведомленности пользователя. Напомним, что в прагматических моделях объяснения важно, чтобы между темой вопроса и объяснением языковой модели существовало отношение релевантности и убеждение ЛПР, что объяснение эмпирически адекватно.

Для поддержки прагматических моделей объяснения можно применить техники промптов, приведенные в табл. 2.

Отметим, что техника промпта Few-shot позволяет реализовать унификационистскую модель объяснения [36]. Напомним, что в этой модели задается шаблон, в который подставляется объяснение, выбираемое из объяснительного ресурса объяснения. Объяснительный ресурс – это мнение видных ученых, обычаи, традиции, сведения, которые могут быть представлены в справочниках, энциклопедиях и др.

Для классических теорий объяснения (ДН-моделей [32]), направленных на получение и обоснование объяснения в виде логического вывода решения, основанного на известных законах и закономерностях, подойдут техники промпта, представленные в табл. 3.

Ранее было сказано о двух целях объяснения в поддержке принятия решений. Это исследовательская цель, которая достигается использованием языковой модели с промптами, реализующими получение прагматических объяснений, повышающих осведомленность эксперта, и практическая цель, которая достигается с использованием языковой модели с промптами для классических теорий объяснения, помогающими получить и обосновать строгий логический вывод решения и его применение для управления ситуацией.

Таблица 2

Роль промпта в прагматической модели объяснения

Техника промпта	Роль промпта в прагматической модели объяснения
Zero-shot	Позволяет получить ответ на любой вопрос
Few-shot	Позволяет получить ответ в виде заданного пользователем шаблона
Role based	Позволяет получить ответы на один и тот же вопрос в разных контекстах
Chain-Of-Thought (CoT)	Позволяет представить сложную задачу в виде связанных подзадач и решить их последовательно
Chain-of-Verification	Расширяет технику цепочки мыслей (CoT), позволяет верифицировать рассуждения языковой модели и обнаружить галлюцинации
Chain-of-Note	Расширяет технику цепочки мыслей (CoT), позволяет повысить качество рассуждений языковой модели, дополняя их пояснениями
Tree of Thoughts (ToT)	Позволяет построить дерево рассуждений вместо цепочки мыслей в методе (CoT); это позволяет найти лучшее альтернативное объяснение
ReAct Prompting	Допускает интерактивный режим пользователя и языковой модели. Пользователь может задействовать внешние данные для решения исследовательских задач и задач планирования

Таблица 3

Роль промпта в классической модели объяснения

Техника промпта	Роль промпта в классической модели объяснения
Chain-of-Knowledge	Позволяет дообучить языковую модель моделью знаний, включающую описание законов или закономерностей конкретной предметной области. Это позволяет повысить достоверность логического вывода языковой модели в исследуемой предметной области
ReAct Prompting	Организует интерактивный режим работы пользователя и языковой модели с возможностью обращения к внешним источникам проверенных данных. Это позволяет использовать описание законов из внешних источников данных, что может повысить достоверность логического вывода
Commonsense	Позволяет получить «набросок» классического объяснения в терминах здравого смысла. Такое объяснение выше было названо объяснением скрытой структуры дедуктивно-номологического объяснения

Применение комбинаций существующих техник промптов или разработка собственного промпта позволяют настроить большую языковую модель на решение задач объяснения альтернатив решений как в парадигме классической теории объяснений, так и в прагматической теории объяснений.

Вопросы разработки и тестирования пользовательских промптов для решения исследовательской и практической задач поддержки принятия решений в настоящей работе не рассматриваются.

ЗАКЛЮЧЕНИЕ

Рассмотрены вопросы применения больших языковых моделей для генерации и объяснения альтернатив решений, полученных системой поддержки принятия решений в условиях неопределенности. Рассмотрены классические (дедуктивно-номологические) и прагматические модели объяснения, предложенные философами. Сформулированы цели и задачи объяснения в процессах поддержки принятия решений в условиях неопределенности. Цель объяснения в системе поддержки

принятия решений заключается в формировании информационной среды принятия решений и сводится к решению исследовательской задачи, позволяющей сформулировать альтернативы решений, и практической задачи, направленной на обоснование и реализацию лучшей альтернативы. Приведены концептуальный анализ функционирования больших языковых моделей, оценка их способностей при решении типовых тестовых задач, определяющие текущий уровень их способностей. Рассмотрены основные техники промптинга (системы запросов к языковой модели) позволяющие настроить языковую модель на генерацию объяснений для решения исследовательской и практической задач поддержки принятия решений в условиях неопределенности.

В первой части статьи сформулированы основные понятия и определения, которые далее будут использованы во второй части статьи. Вторая часть статьи посвящена вопросам измерения и оценки удовлетворенности лица, принимающего решение, объяснениями больших языковых моделей. Оценки удовлетворенности проводятся для исследовательской и практической задач объясне-



ния в условиях неопределенности. Будут проанализированы объяснения двух российских больших языковых моделей, которые будут отнесены к выделенным классам моделей объяснения.

ЛИТЕРАТУРА

1. *Simon, H.A.* Rationality as Process and as Product of Thought / In: *Decision Making: Descriptive, Normative, and Prescriptive Interactions*, ed. by D.E. Bell, H. Raiffa, A. Tversky. – Cambridge: Cambridge University Press, 1988. – P. 58–77.
2. *Checkland, P.B.* Systems Thinking, Systems Practice. – New York: Wiley, 1981. – 330 p.
3. *Axelrod, R.* The Structure of Decision: Cognitive Maps of Political Elites. – Princeton: Princeton University Press, 1976.
4. *Kosko, B.* Fuzzy Thinking: The New Science of Fuzzy Logic. New York: Hyperion, 1993. – 336 p.
5. *Кулинич А.А.* Верификация когнитивных карт на основе объяснения прогнозов // Управление большими системами. – 2010. – Вып. 30.1. – С. 453–469. [*Kulinich, A.A.* Cognitive Maps Verification Based on Processes Explanation // Large-Scale System Control. – 2010. – No. 30.1. – P. 453–469. (In Russian)]
6. *GPT-2* нейросеть от OpenAI. Быстрый старт. – URL: <https://habr.com/ru/articles/440564/> (дата обращения: 16.02.2026). [*GPT-2 nejroset' ot OpenAI. Bystryj start.* – URL: <https://habr.com/ru/articles/440564/> (accessed February 16, 2026). (In Russian)]
7. *Brown, T., Mann, B., Ryder, N., et al.* Language Models Are Few-Shot Learners // arXiv:2005.14165. – 2020. – DOI: <https://doi.org/10.48550/arXiv.2005.14165>
8. *Ouyang, L., Wu, J., Jiang, X., et al.* Training Language Models to Follow Instructions with Human Feedback // arxiv.org/abs/2203.02155. – 2022. – DOI: <https://doi.org/10.48550/arXiv.2203.02155>
9. *Brown, A.* GPT-4 Is OpenAI's Most Advanced System, Producing Safer and More Useful Responses // International Journal of Architectural Computing. – 2024. – Vol. 22, no. 3. – P. 275–276. – DOI:10.1177/14780771241280148
10. *Devlin, J., Chang, M.-W., Lee, K., Toutanova, K.* Bert: Pre-Training of Deep Bidirectional Transformers for Language Understanding // arXiv:1810.04805. – 2018. – DOI: <https://doi.org/10.48550/arXiv.1810.04805>
11. *Lan, Z., Chen, M., Goodman, S., et al.* Albert: A Lite Bert for Self-Supervised Learning of Language Representations // arXiv:1909.11942. – 2019. – DOI: <https://doi.org/10.48550/arXiv.1909.11942>
12. *Touvron, H., Lavril, T., Izacard, G., et al.* Llama: Open and Efficient Foundation Language Models // arXiv:2302.13971. – 2023. – DOI: <https://doi.org/10.48550/arXiv.2302.13971>
13. *Bi, X., Chen, D., Chen, G., et al.* Deepseek LLM: Scaling Open-Source Language Models with Longtermism // arXiv:2401.02954. – 2024. – DOI: <https://doi.org/10.48550/arXiv.2401.02954>
14. *GigaChat-2.* – URL: <https://giga.chat/> (дата обращения: 14.02.2026). [*GigaChat-2.* – URL: <https://giga.chat/> (Accessed February 14, 2026). (In Russian)]
15. *Yandex GPT 5.1.* – URL: <https://ya.ru/ai/gpt> (дата обращения: 14.02.2026). [*Yandex GPT 5.1.* – URL: <https://ya.ru/ai/gpt> (accessed February 14, 2026). (In Russian)]
16. *Визильтер Ю.В.* Приоритетные направления исследований и ключевые тенденции развития технологий ИИ // Труды Двадцать второй Национальной конференции по искусственному интеллекту с международным участием, КИИ–2025. – Т. 1. – Санкт-Петербург, 2025 г. – С. 7–25. – DOI: 10.15622/rcai.2025.001 [*Vizil'ter, Yu.V.* Prioritetnye napravleniya issledovaniy i klyuchevye tendentsii razvitiya tekhnologii II // Trudy Dvadtsat' vtoroy Natsional'noy konferentsii po iskusstvennomu intellektu s mezhdunarodnym uchastiem, KII–2025. – Vol. 1. – St. Petersburg, 2025. – S. 7–25. – DOI: 10.15622/rcai.2025.001] (In Russian)
17. *Брагин А.В., Бахтизин А.Р., Макаров В.Л.* Большие языковые модели четвертого поколения как новый инструмент в научной работе // Искусственные общества. – 2023. – Т. 18, № 1. – DOI: 10.18254/S207751800025046-9 [*Bragin, A., Bakhtizin, A., Makarov, V.* Large Fourth-Generation Language Models as a New Tool in Scientific Research // Artificial Societies. – 2023. – Vol. 18, no. 1. – DOI: 10.18254/S207751800025046-9 (In Russian)]
18. *Бахтизин А.Р.* Вопросы прогнозирования в современных условиях // Экономическое возрождение России. – 2023. – № 2 (76). – С. 53–62. [*Bakhtizin, A.R.* The Challenges of Forecasting under Current Conditions // The Economic Revival of Russia. – 2023. – No. 2 (76). – P. 53–62.] (In Russian)]
19. «Коммерсантъ» узнал о тестировании российских языковых моделей для «Госуслуг» // Forbes. – 02.02.2024. – URL: https://www.forbes.ru/tekhnologii/505447-kommersant-uznal-o-testirovanii-rossijskih-azykovykh-modelej-dla-gosuslug?ysc_lid=lzfs8wut7z80399139 (дата обращения: 14.02.2026). [*Kommersant*] uznal o testirovanii rossijskih yazykovykh modelej dlya «Gosuslug» // Forbes. – February 2, 2024. – URL: https://www.forbes.ru/tekhnologii/505447-kommersant-uznal-o-testirovanii-rossijskih-azykovykh-modelej-dla-gosuslug?ysc_lid=lzfs8wut7z80399139 (accessed February 14, 2026). (In Russian)]
20. *Гусев А.* Обзор Российских систем искусственного интеллекта для здравоохранения. – URL: <https://webiomed.ru/blog/obzor-rossijskikh-sistem-iskusstvennogo-intellekta-dlia-zdravookhraneniia/> (дата обращения: 14.02.2026). [*Gusev, A.* Obzor Rossijskikh sistem iskusstvennogo intellekta dlya zdravookhraneniya. – URL: <https://webiomed.ru/blog/obzor-rossijskikh-sistem-iskusstvennogo-intellekta-dlia-zdravookhraneniia/> (accessed February 14, 2026). (In Russian)]
21. *Shrestha, Y.R., Ben-Menahem, S., von Krogh, G.* Organizational Decision Making Structures in the Age of Artificial Intelligence // California Management Review. – 2019. – Vol. 61, no. 4. – P. 66–83.
22. *Гасанов, Е.* Decision Intelligence: искусственный интеллект с человеческим лицом // IT World. – 22.02.2022. – URL: <https://www.it-world.ru/cionews/1q746pou69z40kokwskckk4gkg4c0og.html> (дата обращения: 14.02.2026). [*Gasanov, E.* Decision Intelligence: Artificial Intelligence with a Human Face // IT World. – March 22, 2022. – URL: <https://www.it-world.ru/cionews/1q746pou69z40kokwskckk4gkg4c0og.html> (accessed February 14, 2026). (In Russian)]
23. *Виртуальные помощники* // TAdviser. – 06.08.2022. – URL: <https://www.tadviser.ru/index.php> (дата обращения 15.02.2026) [*Virtual'nye pomoshchniki* // TAdviser. – August 6, 2022. – <https://www.tadviser.ru/index.php> (accessed February 15, 2026). (In Russian)]
24. *ИИ в аналитике: что за пределами BI?* // TAdviser. – 04.08.2022. – URL: <https://www.tadviser.ru/index.php> (дата обращения: 15.02.2026). [*II v analitike: chto za predelami BI?* // TAdviser. – August 4, 2022. – URL: <https://www.tadviser.ru/index.php> (accessed February 15, 2026). (In Russian)]

25. Дудихин В.В., Кондрашов П.Е. Методология использования больших языковых моделей для решения задач государственного и муниципального управления по интеллектуальному реферированию и автоматическому формированию текстового контента // Государственное управление. – 2024. – № 105. – С. 169–179. [Dudikhin, V.V., Kondrashov, P.E. Methodology of Using Large Language Models to Solve Tasks of State and Municipal Government for Intelligent Abstracting and Automatic Generation of Text Content // Public Administration. – 2024. – No. 105. – P. 169–179.] (In Russian)]
26. Стратегическое направление в области цифровой трансформации государственного управления: утв. распоряжением Правительства Российской Федерации от 22.10.2021 № 2998-р. [Strategicheskoe napravlenie v oblasti cifrovoj transformacii gosudarstvennogo upravleniya: utv. rasporyazheniem Pravitel'stva Rossijskoj Federacii ot 22.10.2021 № 2998-r. (In Russian)]
27. Дмитрий Чернышенко провел стратегическую форсайт-сессию по фундаментальным исследованиям в сфере искусственного интеллекта // Официальный сайт Правительства Российской Федерации. – 30.05.2024. – URL: <http://government.ru/news/51726/> (дата обращения 16.02.2026). [Dmitrii Chernyshenko provel strategicheskuyu forsait-sessiyu po fundamental'nym issledovaniyam v sfere iskusstvennogo intellekta // Ofitsial'nyu sait Pravitel'stva Rossijskoj Federatsii. – May 30, 2024. – URL: <http://government.ru/news/51726/> (accessed February 16, 2026). (In Russian)]
28. *Философия*: Энциклопедический словарь / под ред. А.А. Ивина. – М.: Гардарики, 2004. [Filosofiya: Enciklopedicheskij slovar' / pod red. A.A. Ivina. – M.: Gardariki, 2004. (In Russian)]
29. *Философский энциклопедический словарь* / гл. редакция: Л.Ф. Ильичев, П.Н. Федосеев, С.М. Ковалев, В.Г. Панов. – М.: Советская энциклопедия, 1983. [Filosofskij enciklopedicheskij slovar' / gl. redakciya: L.F. Il'ichyov, P.N. Fedoseev, S.M. Kovalyov, V.G. Panov. – M.: Sovetskaya enciklopediya, 1983. (In Russian)]
30. *Новая философская энциклопедия*: В 4 т. / под ред. В.С. Степина. – М.: Мысль, 2001. [Novaya filosofskaya enciklopediya: Vol. 4 / pod red. V.S. Stepina. – M.: Mysl', 2001. (In Russian)]
31. Woodward, J. Scientific Explanation // In: The Stanford Encyclopedia of Philosophy. Ed. by E.N. Zalta. – Stanford: Stanford University, 2017. – URL: <https://plato.stanford.edu/archives/fall2023/entries/scientific-explanation/> (дата обращения 15-02-2026). [Accessed February 15, 2026.]
32. Hempel, C. Aspects of Scientific Explanation and Other Essays in the Philosophy of Science. – New York: Free Press, 1965.
33. *Закономерность (закон)* // Большая российская энциклопедия: [в 35 т.] / гл. ред. Ю.С. Осипов. – М.: Большая российская энциклопедия, 2004–2017. [Zakonomernost' (zakon) // Bol'shaya rossijskaya enciklopediya: [Vol. 35.] / gl. red. YU. S. Osipov. – Moscow: Bol'shaya rossijskaya enciklopediya, 2004–2017. (In Russian)]
34. Драгалин А. Г. Математический интуиционизм. Введение в теорию доказательств. – М.: Наука, 1979. – 256 с. – (Математическая логика и основания математики). [Dragalin, A.G. Mathematical Intuitionism. Introduction to Proof Theory / Translated by E. Mendelson. – Providence, Rhode Island: AMS, 1988.]
35. *Понимание* / А.А. Ивин, А.Л. Никифоров. Словарь по логике. – М.: Туманит, изд. центр ВЛАДОС, 1997. [Ponimanie / A.A. Ivin, A.L. Nikiforov. Slovar' po logike. – M.: Tumanit, izd. centr Vlados, 1997. (In Russian)]
36. Kitcher, P. Explanatory Unification and the Causal Structure of the World // In: Scientific Explanation. Ed. by P. Kitcher and W. Salmon. – Minneapolis: University of Minnesota Press, 1989. – P. 410–505.
37. Van Fraassen, B. The Scientific Image. – Oxford: Oxford University Press, 1980.
38. Morris, C. Foundations of the Theory of Signs // In: Writings on the General Theory of Signs. Ed. by T. Sebeok. – The Hague: Mouton, 1971.
39. Плотинский Ю.М. Модели социальных процессов: Учебное пособие для высших учебных заведений. Изд. 2-е, перераб. и доп. – М.: Логос, 2001. – 296 с. [Plotinskij, Yu.M. Modeli social'nyh processov: Uchebnoe posobie dlya vysshih uchebnyh zavedenij. Izd. 2 – e, pererab. i dop. – Moscow: Logos, 2001. – 296 s. (In Russian)]
40. Сохор А. М. Объяснение в процессе обучения: Элементы дидактической концепции. – М.: Педагогика, 1988. – 128 с. [Sohor, A.M. Ob"yasnenie v processe obucheniya: Elementy didakticheskoy koncepcii. – Moscow: Pedagogika, 1988. – 128 s. (In Russian)]
41. Vaswani, A., Shazeer, N., Parmar, N., et al. Attention is All You Need // arXiv:1706.03762. – 2017. – DOI: <https://doi.org/10.48550/arXiv.1706.03762>
42. Marecek, D., Rosa, R. Extracting Syntactic Trees from Transformer Encoder Self-attentions // Proceedings of the 2018 EMNLP Workshop BlackboxNLP: Analyzing and Interpreting Neural Networks for NLP. – Brussels, Belgium, 2018. – P. 347–349. – URL: <https://aclanthology.org/W18-5444.pdf/> (дата обращения 15.02.2026). [Accessed February 15, 2026.]
43. Barskaya, I. Benchmarks For LLMs // Unite.AI. – August 28, 2024. – URL: <https://www.unite.ai/benchmarks-for-llms/> (дата обращения 15.02.2026). [Accessed February 15, 2026.]
44. Сбер представил GigaChat 2.0 — новое поколение нейросети. – URL: <https://t-j.ru/news/sber-gigachat-2/> (дата обращения 15.02.2026). [Sber predstavil GigaChat 2.0 — novoe pokolenie nejroseti. – URL: <https://t-j.ru/news/sber-gigachat-2/> (accessed February 15, 2026). (In Russian)]
45. YandexGPT-5-Lite-8B-instruct. – URL: <https://huggingface.co/yandex/YandexGPT-5-Lite-8B-instruct>. (дата обращения 15.02.2026). [Accessed February 15, 2026. (In Russian)]
46. Sahoo, P., Singh, A.K., Saha, S., et al. A Systematic Survey of Prompt Engineering in Large Language Models: Techniques and Applications // arXiv:2402.07927. – 2024. – DOI: <https://doi.org/10.48550/arXiv.2402.07927>
47. Prompt Engineering Guide. – URL: <https://www.promptingguide.ai/> (дата обращения: 16.12.2023). [Accessed December 16, 2023].
48. Wang, Z., Peng, Z., Que, H., et al. RoleLLM: Benchmarking, Eliciting, and Enhancing Role-Playing Abilities of Large Language Models // arXiv:2310.00746. – 2024. – DOI: <https://doi.org/10.48550/arXiv.2310.00746>
49. Wei, J., Wang, X., Schuurmans, D., et al. Chain of Thought Prompting Elicits Reasoning in Large Language Models // arXiv:2201.11903. – 2022. – DOI: <https://doi.org/10.48550/arXiv.2201.11903>
50. Dhuliawala, S., Komeili, M., Xu, J., et al. Chain-of-Verification Reduces Hallucination in Large Language Models //



- arXiv:2309.11495. – 2023. – DOI: <https://doi.org/10.48550/arXiv.2309.11495>
51. Yu, W., Zhang, H., Pan, X., et al. Chain-of-Note: Enhancing Robustness in Retrieval-Augmented Language Models // arXiv:2311.09210. – 2023. – DOI: <https://doi.org/10.48550/arXiv.2311.09210>
52. Liu, J., Lui, A., Lu, X., et al. Generated Knowledge Prompting for Commonsense Reasoning // arXiv:2110.08387. – 2021. – DOI: <https://doi.org/10.48550/arXiv.2110.08387>
53. Yao, S., Yu, D., Zhao, J., et al. Tree of Thoughts: Deliberate Problem Solving with Large Language Models // arXiv:2305.10601. – 2023. – DOI: <https://doi.org/10.48550/arXiv.2305.10601>
54. Yao, S., Zhao, J., Yu, D., et al. React: Synergizing Reasoning and Acting in Language Models // arXiv:2210.03629. – 2023. – DOI: <https://doi.org/10.48550/arXiv.2210.03629>
55. Speer, R., Chin, J., Havasi, C. Conceptnet 5.5: An Open Multilingual Graph of General Knowledge // arxiv.org/abs/1612.03975. – 2017. – DOI: <https://doi.org/10.48550/arXiv.1612.03975>
56. Sap, M., Le Bras, R., Allaway, E., et al. Atomic: An Atlas of Machine Commonsense for If-Then Reasoning // Proceedings of the AAAI Conference on Artificial Intelligence. – Honolulu, 2019. – Vol. 33. – P. 3027–3035.
57. Bosselut, A., Rashkin, H., Sap, M., et al. COMET: Commonsense Transformers for Automatic Knowledge Graph Construction // Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. – Florence, Italy, 2019. – P. 4762–4779.
58. Bosselut, A., Le Bras, R., Choi, Y. Dynamic Neuro-Symbolic Knowledge Graph Construction for Zero-Shot Commonsense Question Answering // Proceedings of the AAAI Conference on Artificial Intelligence. – Vancouver, Canada, 2021. – P. 4923–4931.
59. Shwartz, V., West, P., Le Bras, R., et al. Unsupervised Commonsense Question Answering with Self-talk // arXiv:2004.05483. – 2020. – DOI: <https://doi.org/10.48550/arXiv.2004.05483>

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

*Поступила в редакцию 08.07.2025,
после доработки 17.11.2025.*

Принята к публикации 16.12.2025.

Кулинич Александр Алексеевич – канд. техн. наук, Институт проблем управления им. В.А. Трапезникова РАН, г. Москва,
✉ alexkul@rambler.ru
ORCID ID: <https://orcid.org/0000-0002-4751-205X>

© 2026 г. Кулинич А. А.



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

APPLICATION OF LARGE LANGUAGE MODELS IN DECISION SUPPORT SYSTEMS. PART I: Explanation Models and Large Language Models

A. A. Kulinich

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

✉ alexkul@rambler.ru

Abstract. Large language models (LLMs) significantly influence many spheres of life: education, creativity, science, and business. This paper considers the use of LLMs to explain alternative solutions obtained by a decision support system under uncertainty. Classical and pragmatic models of explanation proposed by philosophers are discussed. The goals and tasks of explanation in decision support processes under uncertainty are formulated. The operation of LLMs is conceptually analyzed, and their current capabilities in solving typical test tasks are assessed. The main techniques of prompting (a system of queries to a language model) are considered; with these techniques, a language model can be tuned to generate explanations for alternative solutions to particular tasks in a subject area. Finally, prompt techniques for supporting pragmatic and classical theories of explaining alternative solutions are considered.

Keywords: decision support, explanation models, explanation goal, explanation tasks, large language model (LLM), prompt techniques.

МЕТОДЫ РЕШЕНИЯ ЗАДАЧИ ОПТИМИЗАЦИИ ПОТОКА ВОЗДУШНЫХ СУДОВ НА ПОСАДКУ

Е. Л. Кулида*, В. Г. Лебедев**, Н. А. Егоров***

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

*✉ elena-kulida@yandex.ru, **✉ lebedev-valentin@yandex.ru, ***✉ negorov@bk.ru

Аннотация. Рассматривается задача оптимизации последовательности и времен посадки воздушных судов, актуальная для увеличения пропускной способности взлетно-посадочных полос аэропортов. Приведены краткий обзор основных подходов к решению задачи, математическая постановка задачи, основные ограничения и целевые функции. Изложено описание точного метода решения и методов приближенного решения задачи при помощи генетических алгоритмов с двумя различными операторами скрещивания, а также эвристического алгоритма. Время получения точного решения экспоненциально растет с ростом размерности задачи (количества воздушных судов), что делает проблематичным его применение на практике. Приближенные методы позволяют получить не всегда оптимальное, но достаточно хорошее решение в реальном времени при посадке воздушных судов. Приведено описание разработанного программного комплекса моделирования, на основе которого проведено сравнительное исследование эффективности основных реализованных методов. Предлагается использование глубокого обучения с подкреплением для решения задачи.

Ключевые слова: пропускная способность взлетно-посадочных полос аэропортов, минимальное безопасное время между посадками, критерии оптимальности, библиотека методов.

ВВЕДЕНИЕ

Ограничения пропускной способности взлетно-посадочных полос аэропортов – известное узкое место при организации воздушного движения, которое является одной из главных причин задержки рейсов. Задержки рейсов, особенно перед посадкой, когда воздушное судно (ВС) находится в полете, приводят к дополнительным расходам авиакомпаний, увеличению расхода топлива, загрязнению атмосферы, недовольству пассажиров. Для увеличения пропускной способности существующих взлетно-посадочных полос аэропортов одной из основных является задача оптимизации потока ВС на посадку. Минимальное безопасное время между посадками последовательных ВС зависит от типов (весов) этих ВС. Пропускная способность взлетно-посадочных полос аэропортов зависит от последовательности ВС на посадку, поскольку ми-

нимальное безопасное время посадки тяжелого ВС после легкого существенно меньше, чем время посадки легкого ВС после тяжелого вследствие образующегося турбулентного вихревого следа, создаваемого ведущим ВС. Широко применяемый на практике метод «первым пришел – первым обслужен» (англ. *First Come First Served*, FCFS) редко обеспечивает оптимальную последовательность с точки зрения пропускной способности взлетно-посадочной полосы или средней задержки ВС.

Задача оптимизации потока ВС на посадку включает планирование последовательности и времени посадки находящихся в зоне терминала ВС с учетом эксплуатационных ограничений с целью минимизации заданных критериев. Решение этой задачи актуально для инструментов поддержки принятия решений, помогающих диспетчерам управлять потоком прибывающих ВС, таких как CenterTRACON (CTAS), разработанный NASA и FAA [1], и его европейский аналог ArrivalMANager



(AMAN) [2]. Существует версия Extended AMAN, рассматривающая область до 500 миль от аэропорта, что позволяет заблаговременно начать оптимизировать последовательность ВС, чтобы уменьшить заторы, шум и расход топлива в воздушном пространстве вблизи аэропортов [3].

Задача оптимизации потока ВС на посадку известна давно и решается уже несколько десятилетий, начиная с 1980-х гг., но поиск новых решений продолжается до настоящего времени, поскольку успешность применения решений в режиме реального времени, когда ВС готовятся к посадке в зоне аэропорта, требует как можно более высокой скорости получения решения [4–6].

Различаются статическая и динамическая постановки задачи оптимизации потока ВС на посадку. В статической постановке информация о наборе ВС, заходящих на посадку, предполагается известной [7–10]. В динамической постановке набор ВС с течением времени изменяется, так как некоторые ВС совершают посадку, а новые ВС появляются в зоне аэропорта [11, 12]. Однако подавляющая часть исследований посвящена решению статической задачи, поскольку динамическая задача решается на основе решения статической с использованием скользящего горизонта и обновлением решений при изменениях набора ВС [12–15].

Одним из главных требований при организации воздушного движения является обеспечение безопасности. Обязательными требованиями с точки зрения обеспечения безопасности для задачи оптимизации потока ВС на посадку являются ограничения безопасного разделения между ВС, согласно правилам Международной организации гражданской авиации, и требование посадки ВС в течение заданного временного окна, обусловленного летно-техническими характеристиками ВС и имеющимся на ВС топливом. На практике на условия эшелонирования могут влиять погодные условия, конфигурация взлетно-посадочных полос, структура маршрутов вылета и захода на посадку и т. д. Могут вводиться дополнительные ограничения, такие как ограниченное смещение ВС относительно позиции в порядке регистрации в зоне аэропорта [7, 16–18], ограничения приоритета, временные отрезки закрытия взлетно-посадочной полосы [8] и т. д. В литературе исследуются различные постановки задачи, в которых рассматривается одна [11, 19, 20] или несколько взлетно-посадочных полос [8, 9, 21–24], используемых только для посадки или как для посадки, так и для взлетов [10, 25, 26]. В последнее время появились работы по интегрированному планированию взаи-

мосвязанных процессов прилета-вылета и движения ВС по рулежным дорожкам [27].

Целевые функции также могут быть различными. Чаще всего целевая функция связана с минимизацией отклонения от целевого времени приземления – осуществляется минимизация средней, максимальной или суммарной задержки рейсов [3, 21, 23, 28–30], максимизация пропускной способности взлетно-посадочной полосы [31], минимизация расходов авиакомпании [32–34], минимизация экологических издержек за счет оптимизации расходов топлива и минимизация выбросов вредных веществ в окружающую среду [32, 35, 36]. Часто при постановке задачи рассматриваются несколько противоречащих друг другу целей, т.е. ставится задача многокритериальной оптимизации [37–39]. Для нескольких взлетно-посадочных полос важной является сбалансированность количества приземляющихся на разные полосы самолетов.

Область исследований задачи динамично развивается от точных методов решения до современных подходов на основе обучения с подкреплением. Задача оптимизации потока ВС на посадку формализована как NP-трудная задача на основе смешанного целочисленного программирования [14, 34, 40–42]. Однако поиски точного решения возможны лишь для задачи небольшой размерности (для небольшого количества ВС), поскольку с ростом количества ВС время расчета решения экспоненциально растет и не является удовлетворительным для практического применения. К основным методам решения задачи относятся динамическое программирование, метод ветвей и границ, эвристические и метаэвристические методы [43]. В большом количестве работ для решения задачи используется динамическое программирование [15–18], метод ветвей и границ [7, 25]. Однако важным недостатком этих подходов является то, что они требуют слишком больших вычислительных затрат для решения крупномасштабных задач. Для получения качественных приближенных решений за приемлемое время разработано значительное количество разнообразных эвристических и метаэвристических алгоритмов [44], таких как генетические алгоритмы [45, 46], имитация отжига [20], поиск с запретами [14], оптимизация муравьиной колонии [38, 47–49], оптимизация роя частиц [50], оптимизации серого волка [51], алгоритм поиска воробьев [52] и т.д. Модели целочисленного программирования и метаэвристические методы, использующие начальное решение для повышения эффективности вычислений, сравниваются в работе [24].

Таблица 1

Минимальная допустимая задержка между приземлениями ВС (с)

		Тип следующего приземляющегося ВС			
		1	2	3	4
Тип предыдущего приземляющегося ВС	1	96	200	181	228
	2	72	80	70	110
	3	72	100	70	130
	4	72	80	70	90

Сравнение эффективности таких подходов к решению задачи затрудняется тем, что основным средством исследования является моделирование, при этом исследователи используют разные доступные им наборы данных, полученные из разных аэропортов по всему миру, например, Мальпенса [13], Линате [14, 17], столичного аэропорта Пекина [11], Карфагена [9], Фьюмичино [13, 14], Гатвик [26], Доха [10], Куньмин Чаншуй [52] и т. д. В связи с этим предлагаемые решения чаще всего сравниваются с используемым в настоящее время подходом «первым пришел – первым обслужен» [28, 29, 42, 47–49].

В последнее время большой потенциал для получения быстрых решений представляют методы обучения с подкреплением [53–55]. Однако в настоящее время из-за низкого уровня доверия к результатам методы искусственного интеллекта не имеют широкого применения в системах организации воздушного движения, в которых критически важна безопасность, традиционно обеспечиваемая человеком. В работе [56] делается вывод о том, что такие методы нуждаются в дальнейшем исследовании и необходима разработка методов объяснимого искусственного интеллекта для их принятия конечными пользователями.

1. АЛГОРИТМЫ РЕШЕНИЯ ЗАДАЧИ

1.1. Постановка и точное решение задачи

Задача формирования оптимальной последовательности посадки ВС формулируется в виде задачи линейного или квадратичного смешанно-целочисленного программирования в зависимости от выбора целевой функции.

Введем обозначения: P – количество ВС, ожидающих посадки; E_i – самое раннее возможное время приземления i -го ВС, $i = \overline{1, P}$; L_i – самое позднее возможное время приземления i -го ВС, $i = \overline{1, P}$; T_i – оптимальное время приземления i -го ВС, $i = \overline{1, P}$; C_i – класс турбулентности i -го ВС, $i = \overline{1, P}$, ВС разбиваются на классы в соответствии с категорией турбулентности, минимальные допустимые задержки назначаются классам ВС; $S_{c_i c_j}$ – минимальная задержка между посадкой ВС класса c_j после посадки ВС класса c_i , $i, j = \overline{1, P}$, $i \neq j$; x_i – назначенное время приземления i -го ВС, $i = \overline{1, P}$.

В табл. 1 представлен пример матрицы $S_{c_i c_j}$ для четырех типов ВС: 1 – Boeing 747, 2 – Boeing 727, 3 – Boeing 707, 4 – McDonnell Douglas DC 9.

Задача формулируется с использованием $\sim P^2$ дополнительных булевых переменных δ_{ij} , определяющих очередность следования ВС на посадку:

$$\delta_{ij} = \begin{cases} 1, & \text{если ВС } i \text{ приземляется раньше ВС } j, \\ 0, & \text{в противном случае,} \end{cases}$$

$$i, j = \overline{1, P}, i \neq j.$$

Основные ограничения:

- Ограничение временного окна посадки $[E_i, L_i]$:

$$E_i \leq x_i \leq L_i, i = \overline{1, P}. \quad (1)$$

- Ограничение порядка приземления – либо ВС i приземляется раньше ВС j ($\delta_{ij} = 1$), либо ВС j приземляется раньше ВС i ($\delta_{ji} = 1$):

$$\delta_{ij} + \delta_{ji} = 1, i, j = \overline{1, P}, i \neq j. \quad (2)$$

- Ограничения минимального разделения между ВС можно представить в виде:

$$x_j \geq x_i + S_{c_i c_j} \delta_{ij} - (L_i - E_j) \delta_{ji}, \forall (i, j) \in U. \quad (3)$$

Задача заключается в минимизации целевой функции при заданных ограничениях. Наиболее часто используются:

- Минимизация стоимости отклонения от оптимального времени посадки – кусочно-линейная целевая функция:

$$\min_X \sum_{i=1}^P (g_i \alpha_i + h_i \beta_i), X = \{x_i\},$$

где g_i и h_i – штрафы i -го ВС за опережение и задержку фактического времени приземления относительно оптимального. Переменные α_i , β_i определяются соотношениями

$$\alpha_i = \max(0, T_i - x_i), i = \overline{1, P};$$

$$\beta_i = \max(0, x_i - T_i), i = \overline{1, P}.$$



• Минимизация суммы квадратов отклонений назначенных времен посадок от оптимальных – квадратичная целевая функция (минимизируются также максимальные и средние отклонения от оптимального времени посадки):

$$\min_X \sum_{i=1}^P (T_i - x_i)^2, X = \{x_i, i = \overline{1, P}\}.$$

• Минимизация времени для приземления всей группы ВС:

$$\min(\max[x_1, \dots, x_N] - \min[x_1, \dots, x_N]).$$

Точное решение задачи можно получить при помощи программного пакета CPLEX. Точные решения не применимы на практике из-за большого времени счета, но они позволяют оценить эффективность приближенных методов решения задачи.

1.2. Приближенное решение задачи при помощи генетических алгоритмов

Решение задачи представляется в виде двух синхронизированных векторов $R = \{Y, X\}$, где $Y = \{y_1, \dots, y_P\}$ – перестановка номеров ВС, определяющих последовательность посадок; $X = \{x_{y_1}, \dots, x_{y_P}\}$ – вектор упорядоченных по возрастанию времен посадок, в котором ВС следуют в порядке, определяемом вектором Y .

Реализация генетического алгоритма для решения задачи оптимизации потока ВС на посадку начинается с создания начальной популяции, содержащей N случайных решений.

Векторы перестановок номеров посадок Y формируются случайным образом, затем для известного вектора $Y = \{y_1, \dots, y_P\}$ рассчитывается вектор времен посадок $X = \{x_{y_1}, \dots, x_{y_P}\}$ с минимальными делениями между последовательными ВС по формулам

$$\begin{aligned} x_{y_1} &= \max(T_1, E_{y_1}), \\ x_{y_i} &= \min\left(\max(x_{y_{i-1}} + S_{C_{y_i}, C_{y_{i-1}}}, E_{y_i}), L_{y_i}\right), \quad (4) \\ i &= \overline{2, P}. \end{aligned}$$

Необходимое и достаточное условие выполнения ограничений (1)–(3) задачи:

$$W(R) = \sum_{i=1}^P \sum_{\substack{j=1, \\ j \neq i, x_j > x_i}}^P \max(0, S_{C_{y_i}, C_{y_j}} - (x_j - x_i)) = 0. \quad (5)$$

Если условие не выполняется, то решение недопустимо.

Решения $R^1 = \{Y^1, X^1\}$ и $R^2 = \{Y^2, X^2\}$ сравниваются на основе значения функции $W(R)$ и значения целевой функции $F(R)$ следующим образом:

- если $W(R^1) < W(R^2)$, то решение R^1 лучше, чем R^2 ;
- если $W(R^1) = W(R^2) = 0$, $F(R^1) < F(R^2)$, то решение R^1 лучше, чем R^2 ;
- если $W(R^1) = W(R^2) = 0$, $F(R^1) = F(R^2)$, то решение R^1 эквивалентно решению R^2 .

Затем выполняется итерационный процесс поиска решения задачи.

Шаг этого процесса заключается в формировании новой популяции при помощи генетических операторов. Все решения текущей популяции разбиваются на пары, над каждой парой выполняются генетические операторы скрещивания и мутации. В результате применения генетических операторов для каждой пары решений создается новая пара дочерних решений, т. е. количество решений удваивается. Для всех решений рассчитываются значения $W(R)$ и $F(R)$, затем отбираются решения для новой популяции, состоящей из N решений.

Процесс завершается после заданного числа итераций, из последней популяции выбирается решение с минимальным значением целевой функции.

Определяющее значение для успешного решения задачи при помощи генетического алгоритма имеет оператор скрещивания. Проблема заключается в том, что стандартные операторы скрещивания неприменимы к перестановкам, поскольку в результате их использования может получиться вектор, не являющийся перестановкой. Разработаны специальные операторы скрещивания для перестановок. Но они существенно сложнее и требуют значительных вычислений.

Поскольку решение поставленной задачи состоит из двух векторов $R = \{Y, X\}$ для преодоления этой проблемы можно применить следующий искусственный прием [57]. Для решений

$$R^1 = \{Y^1, X^1\} \text{ и } R^2 = \{Y^2, X^2\}$$

оператор одноточечного скрещивания с параметром k применяется к векторам

$$X^1 = \{x_{y_1}^1, \dots, x_{y_p}^1\} \text{ и } X^2 = \{x_{y_1}^2, \dots, x_{y_p}^2\}.$$

Полученные векторы

$$Z^1 = \{x_{y_1}^1, \dots, x_{y_k}^1, x_{y_{k+1}}^2, \dots, x_{y_p}^2\},$$

$$Z^2 = \{x_{y_1}^2, \dots, x_{y_k}^2, x_{y_{k+1}}^1, \dots, x_{y_p}^1\}$$

упорядочиваются по возрастанию. После переупорядочивания последовательности ВС на посадку соответствующие векторы

$$\hat{Z}^1 = \{z_{y_1}^1, \dots, z_{y_p}^1\} \text{ и } \hat{Z}^2 = \{z_{y_1}^2, \dots, z_{y_p}^2\}$$

отображаются в векторы

$$\hat{Y}_1 = \{y_1', \dots, y_p'\} \text{ и } \hat{Y}_2 = \{y_1'', \dots, y_p''\}.$$

Векторам \hat{Y}_1 и \hat{Y}_2 по формулам (4) сопоставляются векторы \hat{X}_1 и \hat{X}_2 , получаются два дочерних решения

$$\hat{R}_1 = \{\hat{Y}_1, \hat{X}_1\} \text{ и } \hat{R}_2 = \{\hat{Y}_2, \hat{X}_2\}.$$

Для оценки эффективности предложенного оператора скрещивания был также реализован более сложный матричный оператор скрещивания, разработанный для решения этой задачи в работе [45]. Генетический алгоритм на основе этого оператора требует существенно большего времени счета для выполнения большого количества матричных операций, но при этом выигрыша в эффективности получаемого решения нет.

1.3. Эвристический алгоритм улучшения решения

Известным свойством генетических алгоритмов является быстрое улучшение начального (как правило, случайного) решения на первых итерациях и замедление скорости сходимости по мере приближения к оптимальному решению. Для ускорения сходимости вблизи оптимального решения хорошие результаты показал эвристический алгоритм итеративного улучшения приближенного решения [58].

Текущим становится исходное решение $R_1 = \{Y_1, X_1\}$.

Первый шаг итеративного алгоритма заключается в последовательных $P-1$ сравнениях решений $R_j = \{Y_j, X_j\}$ и $R_{j+1} = \{Y_{j+1}, X_{j+1}\}$ для двух перестановок номеров ВС, отличающихся порядком двух рядом стоящих ВС, в соответствии с формулами

$$Y_j = \{y_1, \dots, y_j, y_{j+1}, \dots, y_p\},$$

$$Y_{j+1} = \{y_1, \dots, y_{j+1}, y_j, \dots, y_p\},$$

$$j = \overline{1, P-1}.$$

Соответствующие векторы X_j и X_{j+1} используются для сравнения решений. Текущим решением \hat{R} при каждом сравнении становится лучшее из решений R_j и R_{j+1} .

Второй шаг итеративного алгоритма заключается в последовательных $P-2$ сравнениях решений $R_j = \{Y_j, X_j\}$, $R_{j+1} = \{Y_{j+1}, X_{j+1}\}$, $R_{j+2} = \{Y_{j+2}, X_{j+2}\}$ для трех перестановок номеров ВС, отличающихся порядком трех рядом стоящих ВС в соответствии с формулами

$$Y_j = \{y_1, \dots, y_j, y_{j+1}, y_{j+2}, \dots, y_p\},$$

$$Y_{j+1} = \{y_1, \dots, y_{j+1}, y_{j+2}, y_j, \dots, y_p\},$$

$$Y_{j+2} = \{y_1, \dots, y_{j+2}, y_j, y_{j+1}, \dots, y_p\},$$

$$j = \overline{1, P-2},$$

и в выборе лучшего из трех решений в качестве текущего решения \hat{R} .

Как показали вычислительные эксперименты, предлагаемый эвристический алгоритм хорошо улучшает приближенное решение, полученное после небольшого числа шагов генетического алгоритма.

2. ИССЛЕДОВАНИЕ ЭФФЕКТИВНОСТИ АЛГОРИТМОВ РЕШЕНИЯ ЗАДАЧИ

Сложным и неоднозначным является вопрос эффективности применения различных подходов к выбору метода решения конкретной задачи. Этот выбор зависит от интенсивности потока ВС на посадку, имеющегося на решение задачи времени, а также от целей и ограничений при принятии решений. Ни один известный метод не может одинаково успешно решать все задачи оптимизации. Для разных групп задач могут лучше работать разные методы [5].

Для исследования эффективности алгоритмов построения оптимальных очередей ВС на посадку авторами был разработан инструментальный программный комплекс моделирования. Были реализованы система имитационного моделирования потока ВС на посадку [59], алгоритмы смешанного целочисленного программирования для точного решения задачи с различными целевыми функциями с помощью метода CPLEX [60], два генетических алгоритма с разными операторами скрещивания, эвристический и гибридный алгоритмы [61]. Все это позволило:



- моделировать задачи с заданным количеством ВС на посадку,
- применять различные алгоритмы оптимизации и анализировать результаты,
- визуализировать получаемые результаты,
- формировать статистические данные на основе заданного количества тестов.

На рис. 1 представлено основное окно программы, в котором задаются параметры задачи, выбираются методы решения и визуализируются результаты получаемых решений.

Рассматриваются четыре типа ВС, обозначаемые буквами U – ультралегкие, L – легкие, M – средние, H – тяжелые.

На рис. 1 представлены три разных решения для одного из тестов. Первое решение соответствует порядку посадок «первым пришел – первым обслужен» (Method: SortT), второе решение получено с помощью генетического алгоритма (Method: GA), третье – точное решение, полученное при помощи CPLEX (Method: CPLEX).

- Для каждого решения R рассчитываются
- статус

$$Status(R) = \begin{cases} \text{True, если } W(R) = 0, \\ \text{False, в противном случае;} \end{cases}$$

- максимальное опережение оптимального времени посадки

$$\max D - (R) = \max_{i=1, P} (0, T_i - x_i);$$

- максимальное отставание от оптимального времени посадки

$$\max D + (R) = \max_{i=1, P} (0, x_i - T_i);$$

- значение целевой функции

$$F(R) = \sum_{i=1}^P (T_i - x_i)^2;$$

- среднее отклонение от оптимального времени посадки

$$\text{avg}D(R) = \sum_{i=1}^P |T_i - x_i| / P;$$

- время занятия взлетно-посадочной полосы

$$\text{EndTime}(R) = \max[x_1, \dots, x_N] - \min[x_1, \dots, x_N].$$

В ходе вычислительных экспериментов генерировались тесты – последовательности очередей ВС на посадку с разным количеством ВС. Затем полученная задача оптимизации потока ВС решалась, и рассчитывалась эффективность решений разными методами на основе большого количества тестов.

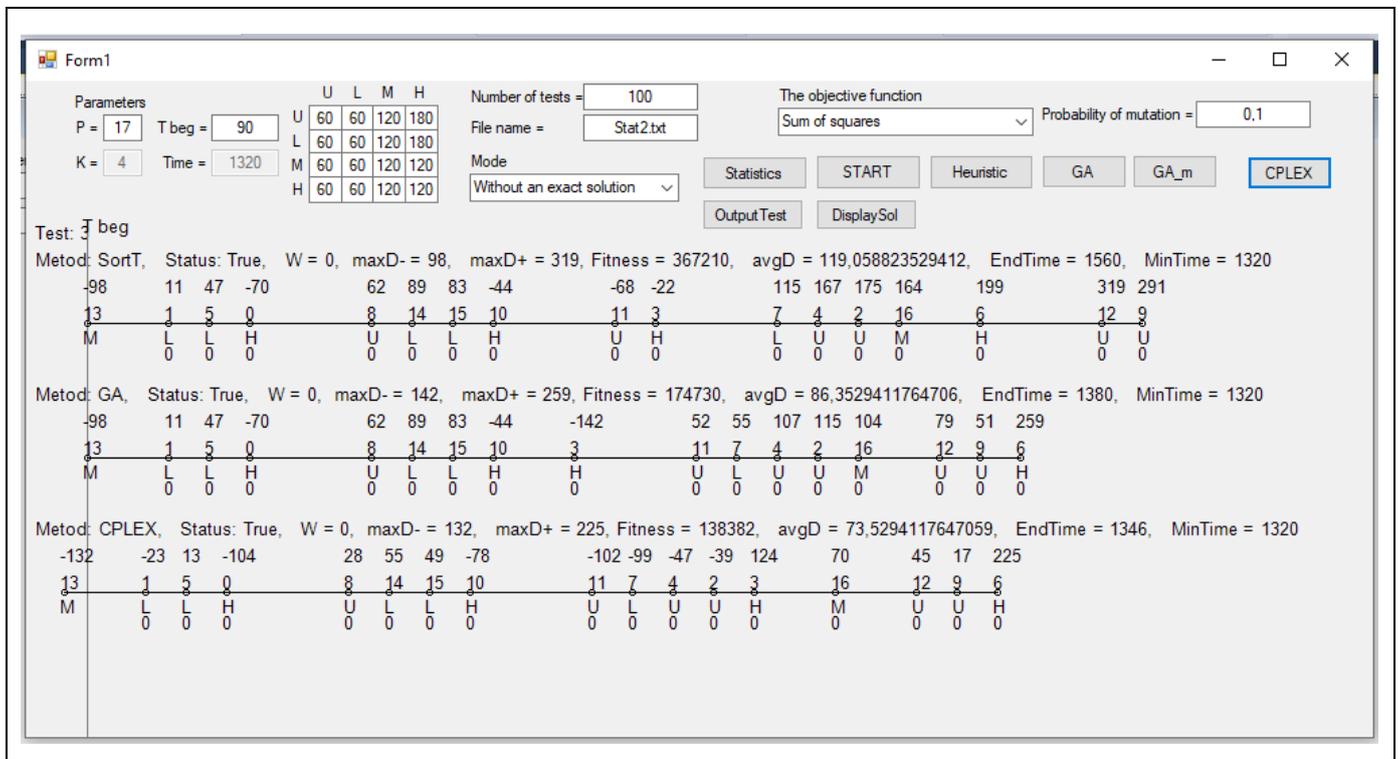


Рис. 1. Визуализация решения задачи разными алгоритмами

Тесты генерировались следующим образом.

Тип ВС определяется при помощи функции генерации случайного целого числа из заданного диапазона $1 \div K$: $C_i = rnd.Next(K)$, $i = \overline{1, P}$.

Времена $0 \leq E_i \leq T_i \leq L_i \leq Time$, $i = \overline{1, P}$, для всех ВС формируются при помощи датчика случайных вещественных чисел. Здесь:

- $T_i = rnd.Next(T_0, T_0 + Time)$ – оптимальное время прибытия i -го ВС, случайное число в диапазоне $(T_0, T_0 + Time)$, $i = \overline{1, P}$. Интенсивность потока ВС варьируется параметром $Time$, который выбирается близким к минимальному времени, необходимому для посадки ВС в соответствии в формулами (4) и (5).

- $E_i = T_0 - t - rnd.Next(2t)$ – самое раннее возможное время приземления i -го ВС, $i = \overline{1, P}$; t – ориентировочная минимальная длительность окна посадки, например $t = 300$ с; $rnd.Next(2t)$ – случайный разброс времени окна посадки в диапазоне $0, 2t$.

- $L_i = T_0 + t + rnd.Next(2t)$ – самое позднее возможное время приземления i -го ВС, $i = \overline{1, P}$.

Программное средство моделирования описано в работе [62].

Сравнение полученных значений позволяет сделать вывод о том, что по всем минимизируемым параметрам приближенное решение в среднем существенно превосходит решение с исходным порядком посадок ВС «первым пришел, первым обслужен», но хуже оптимального решения.

Сравниваются значения целевой функции для решения в порядке «первым пришел, первым обслужен» R и сгенерированного нового решения \bar{R} . Назовем выигрышем в эффективности решения процентное уменьшение значения целевой функции нового решения \bar{R} относительно значения целевой функции решения R , рассчитываемое по формуле

$$\lambda = \frac{F(R) - F(\bar{R})}{F(R)} \cdot 100 \%$$

В табл. 2 приведен средний по 500 тестам выигрыш в эффективности решений λ для задач с 17, 50 и 100 ВС, полученных при применении генетического алгоритма (ГА) и затем улучшенных при

помощи эвристического алгоритма (ГА + ЭА). Для задачи небольшой размерности (с 17-ю ВС) приводится также максимальный возможный выигрыш в эффективности для оптимального решения.

Таблица 2

Выигрыш в эффективности при применении разных алгоритмов

Количество ВС	Оптимальное решение	ГА	ГА + ЭА
17	≈ 67,04 %	≈ 46,2 %	≈ 56,4 %
50	–	≈ 51,5 %	≈ 70,1 %
100	–	≈ 53,5 %	≈ 72,4 %

3. ОБУЧЕНИЕ С ПОДКРЕПЛЕНИЕМ – ПЕРСПЕКТИВНЫЙ ПОДХОД К РЕШЕНИЮ ЗАДАЧИ

В последние годы значимые результаты для задач принятия решений получены во многих областях, в том числе в авиации, при помощи алгоритмов обучения с подкреплением [63, 64]. Метод обучения с подкреплением заключается в поиске оптимального управления марковским процессом принятия решений методом проб и ошибок при взаимодействии агента со средой. Выбор действий агентом при различных состояниях среды оценивается с помощью немедленных вознаграждений. Цель агента заключается в максимизации долгосрочных вознаграждений [65].

Предлагается использовать этот подход для решения задачи оптимизации посадок ВС, для которой критически важным является время решения, поскольку при обучении с подкреплением значительная часть времени вычислений переносится на этап обучения, а обученная программа-агент генерирует решения чрезвычайно быстро. Были проведены вычислительные эксперименты для изучения возможности применения глубокого обучения с подкреплением для решения задачи оптимизации потока ВС на посадку.

Состояние среды в момент времени t определяется следующим образом:

$$S_t = \{Y(t), X(t), C(t), D(t)\},$$

где $Y(t) = (y_1(t), \dots, y_N(t))$ – последовательность номеров ВС на посадку; $X(t) = (x_1(t), \dots, x_N(t))$ – времена посадок в порядке возрастания;



$C(t) = (c_{y_1(t)}, \dots, c_{y_N(t)})$ – вектор типов ВС в порядке посадок; $D(t) = (d_{y_1(t)}, \dots, d_{y_N(t)})$ – вектор отклонений от оптимальных времен посадок.

Задача оптимизации заключается в определении последовательности $Y = (y_1, \dots, y_N)$ и времен $X = (x_1, \dots, x_N)$ посадок для группы из N ВС в течение заданного интервала времени $[T_0, T_k]$, минимизирующих целевую функцию при выполнении ограничений (1)–(3).

Минимизируемая целевая функция имеет вид:

$$F(X) = \sum_{i=1}^N (T_i - x_i)^2.$$

На рис. 2 представлен пример исходной информации для задачи. Матрица времен разделения представлена в двух видах: матрица времен разделения между самолетами из разных классов и мат-

рица времен разделения между конкретными типами самолетов.

Эпизод обучения начинается с некоторого произвольного состояния S_0 . На каждом временном шаге t при состоянии S_t агент выполняет некоторое действие $A_t \in A$, в результате которого среда переходит в состояние S_{t+1} , при этом агент получает немедленное вознаграждение R_t .

Немедленное вознаграждение за эпизод определяется большим штрафом, если не удалось получить расписание времен посадок всех самолетов, удовлетворяющее заданным ограничениям, и штрафом, равным сумме квадратов отклонений от целевых времен посадок, в противном случае:

$$R_t = \begin{cases} -5000, & \text{корректное расписание не получено} \\ -\sum_{i=1}^N (T_i - x_i)^2, & \text{в противном случае.} \end{cases}$$

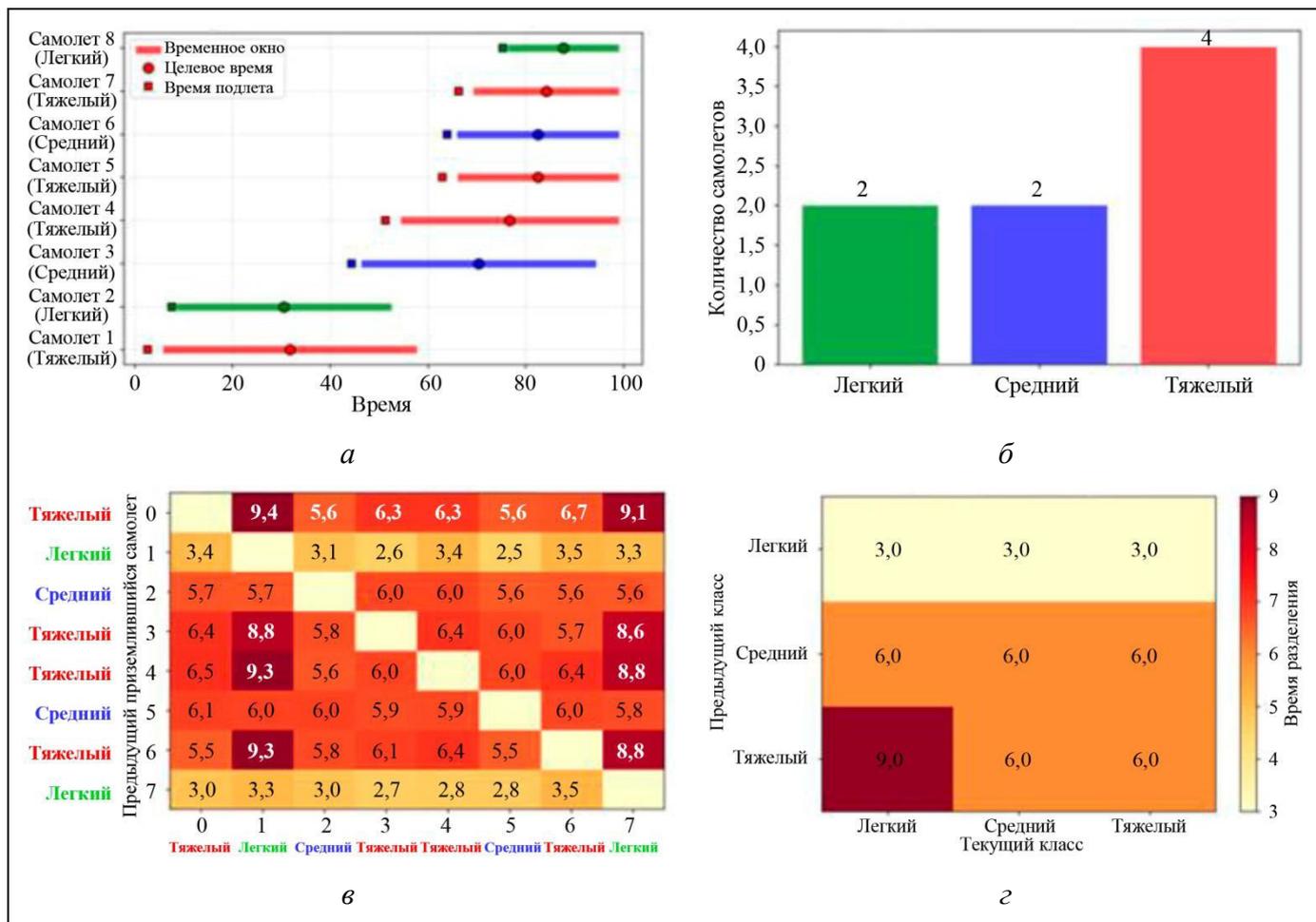


Рис. 2. Исходные данные для задачи: а – временные окна посадки самолетов; б – распределение самолетов по классам; в – несимметричная матрица времен разделения (строка → столбец: предыдущий → текущий); г – матрица разделения между классами (предыдущий → текущий)

После завершения эпизода наборы $[S_t, A_t, R_t, S_{t+1}]$ сохраняются в памяти (буфере воспроизведения) (рис. 3). При обучении агента из буфера воспроизведения выбираются мини-пакеты накопленных данных (рис. 4).

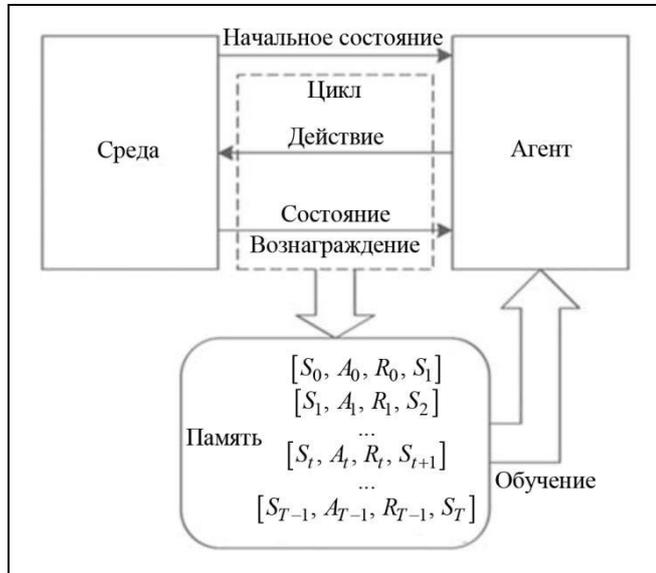


Рис. 3. Формирование данных для обучения агента

В состоянии S_t агент выбирает действие $A_t \in A$ в соответствии с функцией выбора действия a в состоянии s – стратегии $\pi(s, a)$; целью стратегии является максимизация долгосрочного вознаграждения.

При Q -обучении [66] функция ценности действий $Q(S_t, A_t)$ аппроксимирует оптимальную функцию ценности действий независимо от применяющейся стратегии. На каждом временном шаге значения функции ценности действия обновляются по формуле

$$Q(S_t, A_t) \leftarrow Q(S_t, A_t) + \alpha \left[R_{t+1} + \gamma \max_{a \in A} Q(S_{t+1}, a) - Q(S_t, A_t) \right],$$

где α – скорость обучения, γ – коэффициент дисконтирования.

После обновления функции ценности действия обновляется стратегия $\pi(s, a)$. Q -обучающий агент выполняет ϵ -жадный выбор действия. Действие A_t , максимизирующее значение функции $Q(S_t, A_t)$ в состоянии S_t , выбирается с вероятностью $1-\epsilon$, с вероятностью ϵ выбирается случайное действие. Сначала устанавливается максимальное значение $\epsilon = \epsilon_{\max}$, затем в процессе обучения ϵ уменьшается до тех пор, пока не достигнет минимального значения $\epsilon = \epsilon_{\min}$. Таким образом реализуется итерационный процесс, в ходе которого последовательно уточняется функция $Q(s, a)$, и стратегия $\pi(s, a)$ постоянно совершенствуется относительно уточненной функции $Q(s, a)$.

При глубоком Q -обучении (см. рис. 4) для аппроксимации функции Q используется глубокая

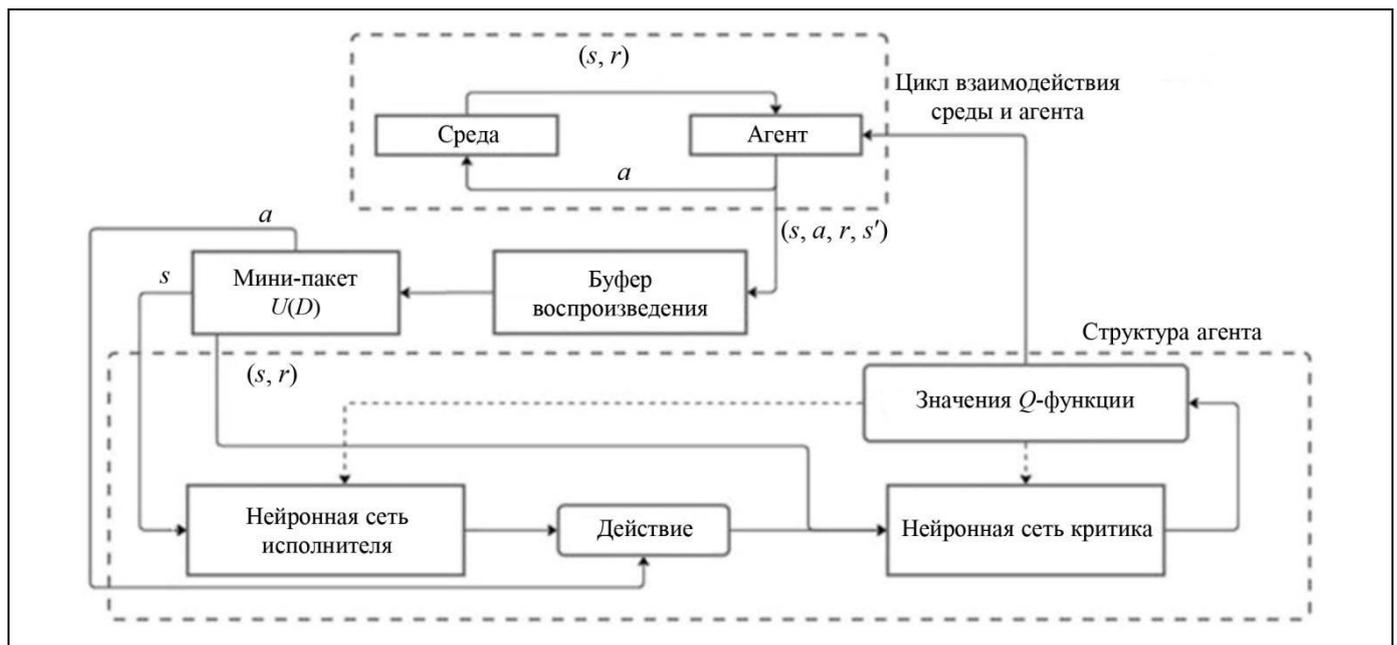


Рис. 4. Схема алгоритма глубокого Q -обучения



нейронная сеть и рассматривается функция $Q(s, a, \theta)$, где θ – параметры нейронной сети [67]. Цель обучения сети заключается в определении параметров θ'_t :

$$Y'_t = R_{t+1} + \max_{a_{t+1}} Q(s_{t+1}, A_{t+1}; \theta'_t).$$

Для определения наилучшего действия в соответствии с функцией $Q(s, a, \theta)$ используется вторая нейронная сеть, цель обучения которой заключается в определении параметров θ''_t :

$$Y''_t = R_{t+1} + \gamma Q(s_{t+1}, \arg \max_{a_{t+1}} Q(s_{t+1}, a_{t+1}; \theta'_t); Q''_t).$$

График наград при последовательных эпизодах обучения агента с параметрами $\alpha = 0,0005$, $\gamma = 0,98$, $\epsilon_{\max} = 1,0$, $\epsilon_{\min} = 0,05$, $\epsilon_{\text{decay}} = 0,999$ представлен на рис. 5.

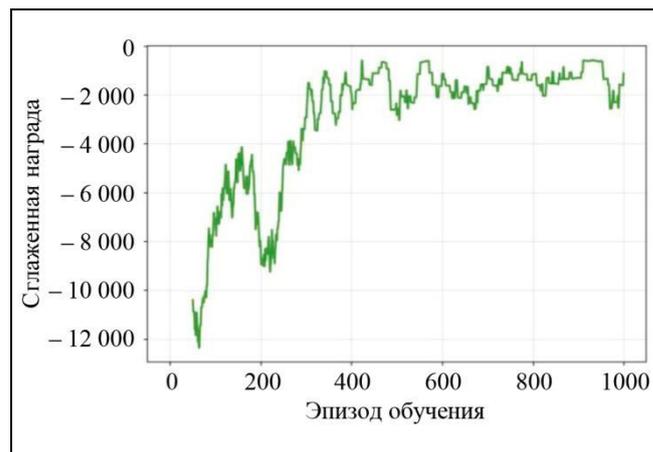


Рис. 5. График обучения агента

На рис. 6 показано отклонение времен посадок от целевого времени для тестового примера, представленного на рис. 2, в расписании, составленном обученным агентом.

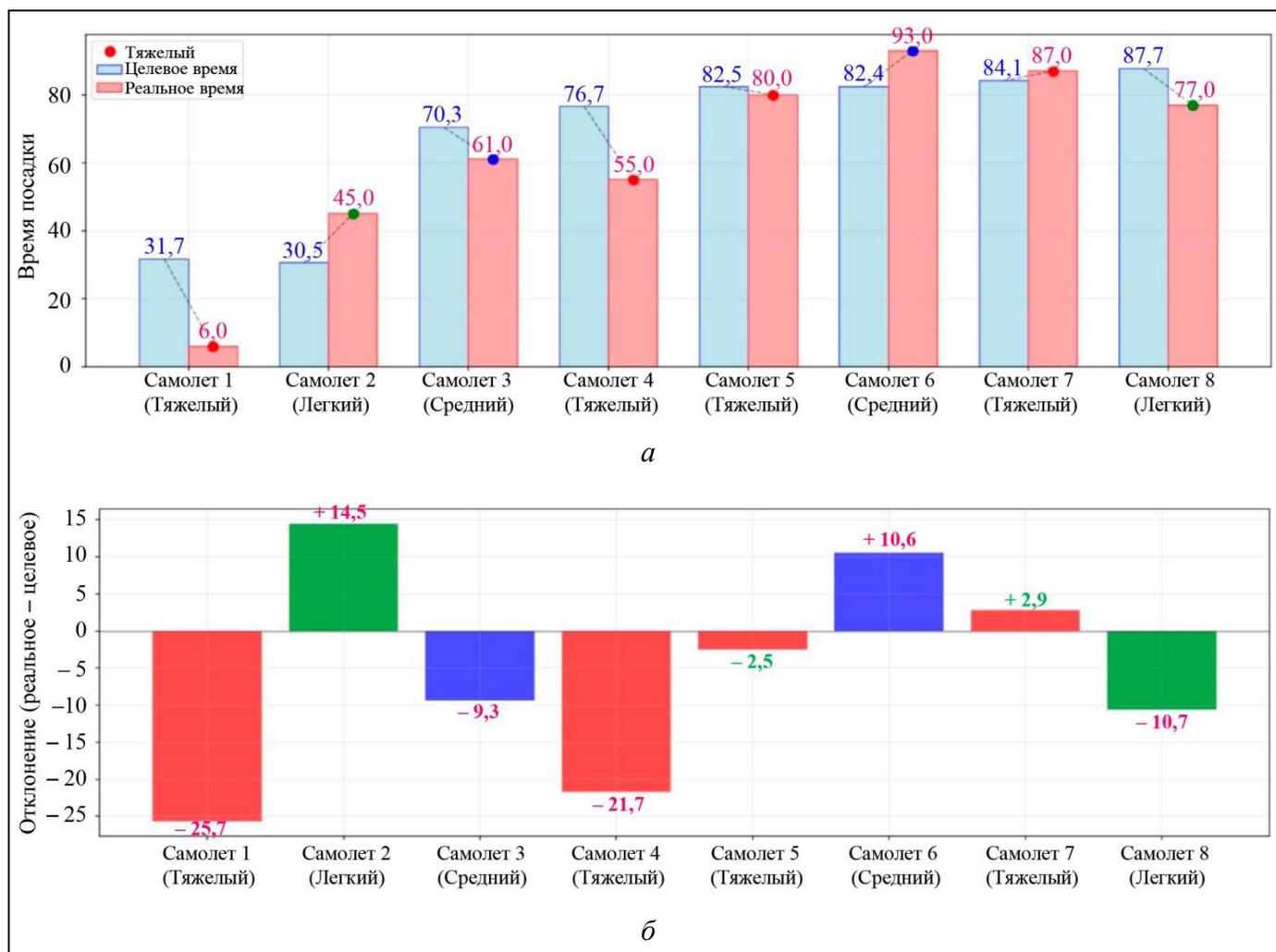


Рис. 6. Отклонения времен посадок от целевого времени при решении задачи обученным агентом

В изученной литературе встречаются упоминания о применении метода для решения этой задачи, однако описаний конкретных алгоритмов в открытых источниках найти не удалось. Подход на основе обучения с подкреплением имеет ряд неоспоримых достоинств по сравнению с другими алгоритмами:

– существенно более быстрое решение при помощи обученного агента;

– возможность использовать необходимую целевую функцию, переобучив агента с новой функцией награды;

– возможность решать задачу без разделения самолетов на 3-4 класса, задавая матрицу времен разделений между посадками для конкретных типов самолетов.

Дальнейшие исследования этого подхода предполагаются продолжить. Проведенные эксперименты с использованием одного из первых алгоритмов обучения с подкреплением – глубокого Q-обучения – показывают возможность применения обучения с подкреплением для решения задачи оптимизации потока ВС на посадку. Они показывают, что программа-агент способна научиться составлять расписание посадок ВС с соблюдением необходимых ограничений, не используя для этого предварительно запрограммированный алгоритм, а в процессе многократно повторяющихся эпизодов взаимодействия со средой моделирования, на основе получаемой на каждом шаге взаимодействия награды, с помощью которой среда оценивает успешность действий агента.

ЗАКЛЮЧЕНИЕ

Оптимизация потока ВС на посадку является одной из важнейших проблем при управлении воздушным движением, поскольку оптимальное планирование посадок позволяет повышать эффективность использования взлетно-посадочных полос и уменьшать задержки рейсов. Разработан ряд методов решения этой задачи на основе смешанного целочисленного программирования, динамического программирования, метаэвристических подходов, таких как генетические алгоритмы, оптимизация муравьиной колонии, алгоритм оптимизации серого волка и т. д. В проведенном исследовании на большом количестве тестов сравнивалась эффективность генетического и гибридного (генетического и эвристического) алгоритмов. Результаты исследований показали, что применение этих алгоритмов позволяет существенно улучшить значение целевой функции, а во многих случаях даже получить ее оптимальное значение. Эти методы не такие затратные по времени вычислений, как точ-

ные методы, но все-таки требуют значительного объема вычислений в то время, когда ВС находятся в полете вблизи аэропорта, и решение необходимо постоянно пересчитывать. Ситуация осложняется тем, что основным способом оценки эффективности эвристических алгоритмов является моделирование, при этом помимо использования аппаратуры с разными параметрами и разного уровня мастерства программистов, используются доступные наборы данных, полученные в различных условиях в аэропортах по всему миру.

Перспективным для более быстрого решения задачи представляется получивший широкое распространение в последнее время подход обучения с подкреплением. Особенностью этого подхода является перенесение значительной части вычислений на этап обучения нейронных сетей (агентов), а затем обученный агент генерирует решение достаточно быстро. Для оценки эффективности такого подхода необходимы дальнейшие исследования.

Несмотря на многолетние усилия многочисленных ученых, эффективные алгоритмы решения задачи оптимизации последовательности и времен посадки ВС пока не найдены. Наилучшие разработанные алгоритмы имеют как несомненные достоинства, так и очевидные недостатки и показывают разную эффективность в различных ситуациях, при различных целевых функциях и ограничениях. Возможным решением проблемы может стать использование библиотеки методов для применения наилучшего доступного решения в конкретных условиях задачи оптимизации последовательности и времен посадки ВС.

ЛИТЕРАТУРА

1. *Erzberger, H., Davis, T.J., Green, S.* Design of Center-TRACON Automation System // Proc. AGARD Guidance and Control Symposium on Machine Intelligence in Air Traffic Management. – Berlin, Germany, 1993.
2. *Liang, M.* Aircraft Route Network Optimization in Terminal Maneuvering Area: Ph. D. thesis. – Toulouse, France: Université Paul Sabatier, 2018. – URL: <https://theses.hal.science/tel-01703861/document>.
3. *Montlaur, A., Delgado, L.* Delay Assignment Optimization Strategies at Pre-Tactical and Tactical Levels // Fifth SESAR Innovation Days. – Bologna, Italy, 2015. – P. 1–8.
4. *Ikli, S., Mancel, C., Mongeau, M., et al.* The Aircraft Runway Scheduling Problem: A Survey // Computers & Operations Research. – 2021. – Vol. 132, no. 3. – DOI: 10.1016/j.cor.2021.105336
5. *Shirini, K., Aghdasi, H.S., Saeedvand, S.* A Comprehensive Survey on Multiple-Runway Aircraft Landing Optimization Problem // International Journal of Aeronautical and Space Sciences. – 2024. – No. 25. – P. 1574–1602. – DOI: 10.1007/s42405-024-00747-z



6. *Княжеский А.Ю., Баушев С.В.* Современное состояние и перспективы развития систем планирования использования воздушного пространства. Ч. 2 // *Crede Experto: транспорт, общество, образование, язык.* – 2025. – № 3 (46). – DOI: 10.51955/2312-1327_2025_3_87. [*Knyazhsky, A.Yu., Baushev, S.V.* Current State and Prospects for Development of Systems for Planning Airspace Management. Part 2 // *Crede Experto: Transport, Society, Education, Language.* – 2025. – No. 3 (46). – DOI 10.51955/2312-1327_2025_3_87. (In Russian)]
7. *Eun, Y., Hwang, I., Bang, H.-C.* Optimal Arrival Flight Sequencing and Scheduling Using Discrete Airborne Delays // *IEEE Transactions on Intelligent Transportation Systems.* – 2010. – Vol. 11, no. 2. – P. 359–373. – DOI: 10.1109/TITS.2010.2044791
8. *Phirouzabadi, A.M., Aminmayeri, M., Mahmoudian, M.* Aircraft Landing Scheduling Based on Unavailability of Runway Constraint Through a Time Segment Heuristic Method // *International Journal of Informatics and Communication Technology (IJ-ICT).* – 2013. – Vol. 2, no. 3. – P. 175–182. – DOI: 10.11591/ij-ict.v2i3.5284
9. *Messaoud, M.B., Ghedira, K., Harizi, R.* The Multiple Runway Aircraft Landing Problem: A Case Study for Tunis Carthage Airport // *Proceedings of IEEE International Conference on Systems, Man, and Cybernetics (SMC).* – Banff, Canada, 2017. – DOI: 10.1109/SMC.2017.8123051
10. *Ghoniem, A., Sherali, H.D., Baik, H.* Enhanced Models for a Mixed Arrival-Departure Aircraft Sequencing Problem // *INFORMS Journal on Computing.* – 2014. – Vol. 26, no. 3. – P. 514–530. – DOI: 10.1287/ijoc.2013.0581
11. *Ji, X.-P., Cao, X.-B., Du, W., Tang, K.* An Evolutionary Approach for Dynamic Single Runway Arrival Sequencing and Scheduling Problem // *Soft Computing.* – 2017. – Vol. 21. – P. 7021–7037. – DOI: 10.1007/s00500-016-2241-8
12. *Bennell, J.A., Mesgarpour, M., Potts, C.N.* Dynamic Scheduling of Aircraft Landing // *European Journal of Operational Research.* – 2017. – Vol. 258, no. 1. – P. 315–327. – DOI: 10.106/j.ejor.206.08.05
13. *Samà, M., D'Ariano, A., Pacciarelli, D.* Rolling Horizon Approach for Aircraft Scheduling in the Terminal Control Area of Busy Airports // *Procedia Social and Behavioral Sciences.* – 2013. – No. 80. – P. 531–552. – DOI: 10.1016/j.sbspro.2013.05.029
14. *Furini, F., Kidd, M.P., Persiani, C.A., Toth, P.* Improved Rolling Horizon Approaches to the Aircraft Sequencing Problem // *Journal of Scheduling.* – 2015. – Vol. 18, no. 5. – P. 435–447. – DOI: 10.1007/s10951-014-0415-8
15. *Lieder, A., Stolletz, R.* Scheduling Aircraft Take-offs and Landings on Interdependent and Heterogeneous Runways // *Transportation Research. Part E: Logistics and Transportation Review.* – 2016. – No. 88. – P. 167–188. – DOI: 10.1016/j.tre.2016.01.015
16. *Balakrishnan, H., Chandran, B.G.* Algorithms for Scheduling Runway Operations Under Constrained Position Shifting // *Operations Research.* – 2010. – Vol. 58, no. 6. – P. 1650–1665. – DOI: 10.1287/opre.1100.0869
17. *Furini, F., Kidd, M.P., Persiani, C.A., Toth, P.* State Space Reduced Dynamic Programming for the Aircraft Sequencing Problem with Constrained Position Shifting // In: *Lecture Notes in Computer Science*, vol. 8596. Ed. by P. Foulhoux, L.E. Neves Gouveia, A.R. Mahjoub, V.T. Paschos. *Proceedings of International Symposium on Combinatorial Optimization.* – Lisbon, Portugal, 2014. – P. 267–279. – DOI: 10.1007/978-3-319-09174-7-23
18. *Chandran, B., Balakrishnan, H.* A Dynamic Programming Algorithm for Robust Runway Scheduling // *Proceedings of 2007 American Control Conference.* – New York, USA, 2007. – DOI: 0.09/ACC.2007.4282922
19. *Vadlamani, S., Hosseini, S.* A Novel Heuristic Approach for Solving Aircraft Landing Problem with Single Runway // *Journal of Air Transport Management.* – 2014. – No. 40. – P. 144–148. – DOI: 0.1016/j.jairtraman.2014.06.009
20. *Faye, A.* A Quadratic Time Algorithm for Computing the Optimal Landing Times of a Fixed Sequence of Planes // *European Journal of Operational Research.* – 2018. – Vol. 270, no. 3. – P. 1148–1157. – DOI: 10.1016/j.ejor.2018.04.021
21. *Hancerliogullari, G., Rabadi, G., Al-Salem, A.H., Kharbeche, M.* Greedy Algorithms and Metaheuristics for a Multiple Runway Combined Arrival-Departure Aircraft Sequencing Problem // *Journal of Air Transport Management.* – 2013. – No. 32. – P. 39–48. – DOI: 10.1016/j.jairtraman.2013.06.001
22. *Ghoniem, A.F., Farhadi, F., Reihaneh, M.* An Accelerated Branch-and-Price Algorithm for Multiple-runway Aircraft Sequencing Problems // *European Journal of Operational Research.* – 2015. – Vol. 246, no. 1. – P. 34–43. – DOI: 10.1016/j.ejor.2015.04.019
23. *Salehipour, A.* An Algorithm for Single-and Multiple-Runway Aircraft Landing Problem // *Mathematics and Computers in Simulation.* – 2020. – Vol. 175, no. 6. – P. 179–191. – DOI: 10.1016/j.matcom.2019.10.006
24. *Santos, L., Fuchigami, H.* An Optimization Study of the Time-Index and Arc-Time-Indexed Models and the Iga Metaheuristic for the Aircraft Sequencing Problem // *Revista de Gestão Social e Ambiental.* – 2025. – Vol. 19, no. 4. – DOI: 10.24857/rgsa.v19n4-032
25. *D'Ariano, A., D'Urgolo, P.* Optimal Sequencing of Aircrafts Take-off and Landing at a Busy Airport // *Proceedings of 13th International IEEE Conference on Intelligent Transportation Systems.* – Funchal, Portugal, 2010. – P. 1569–1574. – DOI: 10.1091/ITSC.2010.5625114
26. *Rodríguez-Díaz, A., Adenso-Díaz, B., González-Torre, P.L.* Minimizing Deviation from Scheduled Times in a Single Mixed Operation Runway // *Computers Operations Research.* – 2016. – No. 78. – P. 193–202. – DOI: 10.1016/j.cor.2016.09.014
27. *Xia, C., Wen, Y., Hu, M., et al.* Microscopic-Level Collaborative Optimization Framework for Integrated Arrival-Departure and Surface Operations: Integrated Runway and Taxiway Aircraft Sequencing and Scheduling // *Aerospace.* – 2024. – Vol. 11, no. 12. – DOI: 10.3390/aerospace11121042
28. *Salehipour, A., Ahmadian, M.M.* Heuristics for Flights Arrival Scheduling at Airports // *International Transactions in Operational Research.* – 2020. – Vol. 29, no. 1. – DOI: 10.1111/itor.12901
29. *Awasthi, A., Kramer, O., Lässig, J.* Aircraft Landing Problem: an Efficient Algorithm for a Given Landing Sequence // *Proceedings of 2013 IEEE 16th International Conference on Computational Science and Engineering.* – Sydney, Australia, 2013. – DOI: 0.109/CSE.2013.14
30. *Mokhtarimousavi, S., Rahami, H., Kaveh, A.* Multi-Objective Mathematical Modeling of Aircraft Landing Problem on a Runway in Static Mode, Scheduling and Sequence Determination Using NSGA-II // *International Journal of Optimization in Civil Engineering.* – 2015. – Vol. 5, no. 1. – P. 21–36.

31. *Prakash, R., Piplani, R., Desai, J.* An Optimal Data-Splitting Algorithm for Aircraft Scheduling on a Single Runway to Maximize Throughput // *Transportation Research. Part C: Emerging Technologies.* – 2018. – No. 95. – P. 570–581. – DOI: 10.1016/j.trc.2018.07.031
32. *Chen, H., Solak, S.* Lower Cost Arrivals for Airlines: Optimal Policies for Managing Runway Operations Under Optimized Profile Descent // *Production and Operations Management.* – 2015. – Vol. 24, no. 3. – P. 402–420. – DOI: 10.1111/poms.12244
33. *Rezaei, H.* An Iterative Bidding Approach Applied to Cost Reduction in the Context of Aircraft Landing Problem // *Proceedings of 2018 IEEE 22nd International Conference on Computer Supported Cooperative Work in Design (CSCWD).* – Nanjing, China, 2018. – DOI: 10.1109/cscwd.2018.8465389
34. *Serhan, D., Lee, H., Yoon, S.W.* Minimizing Airline and Passenger Delay Cost in Airport Surface and Terminal Airspace Operations // *Journal of Air Transport Management.* – 2018. – No. 73. – P. 120–133. – DOI: 10.1016/j.jairtraman.2018.07.001
35. *Sölveling, G., Solak, S., Clarke, J.-P., Johnson, E.L.* Scheduling of Runway Operations for Reduced Environmental Impact // *Transportation Research. Part D: Transport and Environment.* – 2011. – Vol. 16, no. 2. – P. 110–120. – DOI: 10.1016/j.trd.2010.09.004
36. *Tian, Y., Wan, L., Han, K., Ye, B.* Optimization of Terminal Airspace Operation with Environmental Considerations // *Transportation Research. Part D: Transport and Environment.* – 2018. – No. 63. – P. 872–889. – DOI: 10.1016/j.trd.2018.06.018
37. *Zhang, J., Zhao, P., Zhang, Y., et al.* Criteria Selection and Multi-Objective Optimization of Aircraft Landing Problem // *Journal of Air Transport Management.* – 2020. – No. 82. – DOI: 10.1016/j.jairtraman.2019.101734
38. *Tang, K., Wang, Z., Cao, X., et al.* A Multi-Objective Evolutionary Approach to Aircraft Landing Scheduling Problems // *Proceedings of 2008 IEEE Congress on Evolutionary Computation (IEEE World Congress on Computational Intelligence).* – Hong Kong, China, 2008. – DOI: 10.1109/cec13288.2008
39. *Wang, S., Yue, Z., Zhang, Z.-H., Yu, H.* Multi-Objectives Optimization on Flights Landing Sequence at Busy Airport // *Journal of Transportation Systems Engineering and Information Technology.* – 2012. – Vol. 12, no. 4. – P. 135–142. – DOI: 10.1016/s1570-6672(11)60218-3
40. *Beasley, J.E., Krishnamoorthy, M., Sharaiha, Y.M., Abramson, D.* Scheduling Aircraft Landings – the Static Case // *Transportation Science.* – 2000. – Vol. 34, no. 2. – P. 180–197. – DOI: 10.1287/trsc.34.2.80.12302
41. *Спирidonov А.А., Кумков С.С.* Поддержание порядка движения судов в проблеме безопасного объединения потоков воздушных судов // *Вестник Удмуртского университета. Математика. Механика. Компьютерные науки.* – 2022. – Т. 32. Вып. 3. – С. 433–446. – DOI: 35634/vm220306. [*Spiridonov, A.A., Kumkov, S.S.* Keeping Order of Vessels in Problem of Safe Merging Aircraft Flows // *Vestnik Udmurtskogo Universiteta. Matematika. Mekhanika. Komp'uternye Nauki.* – 2022. – Vol. 32, no. 3. – P. 433–446. – DOI: 35634/vm220306]
42. *Rogovs, S., Nikitina, V., Gerdt, M.* A Novel Mixed-Integer Programming Approach for the Aircraft Landing Problem // *Frontiers in Future Transportation.* – 2022. – No. 3. – DOI: 10.3389/ffutr.2022.968957
43. *Bennell, J.A., Mesgarpour, M., Potts, C.N.* Airport Runway Scheduling // *Annals of Operations Research.* – 2011. – Vol. 9, no. 2. – P. 115–138. – DOI: 10.1007/s10288-011-0172-x
44. *Вересников Г.С., Егоров Н.А., Кулида Е.Л., Лебедев В.Г.* Методы построения оптимальных очередей воздушных судов на посадку. Ч. 2. Методы приближенного решения // *Проблемы управления.* – 2018. – № 5. – С. 2–13. – DOI: 10.25728/pu.2018.5.1 [*Veresnikov, G.S., Egorov, N.A., Kulida, E.L., Lebedev, V.G.* Methods for Solving of the Aircraft Landing Problem. II. Approximate Solution Methods // *Automation and Remote Control.* – 2019. – Vol. 80, no. 8. – P. 1502–1518.]
45. *Hu, X.B., Di Paolo, E.* Binary-Representation-Based Genetic Algorithm for Aircraft Arrival Sequencing and Scheduling // *IEEE Transactions on Intelligent Transportation Systems.* – 2008. – Vol. 9, no. 2. – P. 301–310. – DOI: 10.1109/TITS.2008.922884
46. *Hu, X.B., Paolo, E.D.* An Efficient Genetic Algorithm with Uniform Crossover for Air Traffic Control // *Computers & Operations Research.* – 2009. – Vol. 36, no. 1. – P. 245–259. – DOI: 10.1016/j.cor.2007.09.005
47. *Hu, B.* An Efficient Ant Colony Algorithm Based on Wake-Vortex Modeling Method for Aircraft Scheduling Problem // *Journal of Computational and Applied Mathematics.* – 2016. – Vol. 317, no. 2. – P. 157–170. – DOI: 10.1016/j.cam.2016.11.043
48. *Zhan, Z.-H., Zhang, J., Li, Y., et al.* An Efficient Ant Colony System Based on Receding Horizon Control for the Aircraft Arrival Sequencing and Scheduling Problem // *IEEE Transactions on Intelligent Transportation Systems.* – 2010. – Vol. 11, no. 2. – P. 399–412. – DOI: 10.1109/TITS.2010.2044793
49. *Feng, X.R., Feng, X.J., Liu, D.* The Application of Ant Colony Optimization Algorithm in the Flight Landing Scheduling Problem // *Applied Mechanics and Materials.* – 2013. – Vols. 411–414. – P. 2698–2703. – DOI: 10.4028/www.scientific.net/AMM.411-414.2698
50. *Girish, B.* An Efficient Hybrid Particle Swarm Optimization Algorithm in a Rolling Horizon Framework for the Aircraft Landing Problem // *Applied Soft Computing.* – 2016. – No. 44. – P. 200–221. – DOI: 10.1016/j.asoc.2016.04.011
51. *Teymori, M., Taghizadeh, H., Honarmand, M.A., Pourmahmond, J.* A New Improved Gray Wolf Optimization Algorithm to Solve the Aircraft Landing Problem at Mashhad Shahid Hasheminejad International Airport // *International Journal of Nonlinear Analysis and Application.* – 2022. – Vol. 13, no. 2. – P. 435–445. – DOI: 10.22075/ijnaa.2021.23255.2510
52. *Zhao, W., Liang, T.* Optimization of Terminal Area Arrival Flight Sorting Based on an Improved Sparrow Search Algorithm // *Science Progress.* – 2024. – Vol. 107, no. 1. – P. 1–17. – DOI: 10.1177/00368504241238078
53. *Кулида Е.Л., Лебедев В.Г.* Методы решения задач планирования и регулирования потоков воздушного движения Ч. 2. Применение методов глубокого обучения с подкреплением // *Проблемы управления.* – 2023. – № 2. – С. 3–18. – DOI: 10.25728/pu.2023.2.1. [*Kulida, E.L., Lebedev, V.G.* Methods for Solving Some Problems of Air Traffic Planning and Regulation. Part II: Application of Deep Reinforcement Learning // *Control Sciences.* – 2023. – No. 2. – P. 2–14. – DOI: 10.25728/cs.2023.2.1]
54. *Соседов В.А., Кулида Е.Л., Макаревич М.В.* Решение задачи оптимизации последовательности и времен посадок воздушных судов при помощи обучения с подкреплением // *Труды 17-й Междунар. конф. «Управление развитием крупномасштабных систем» (MLSD'2024).* – Москва, 2024. – С. 1056–1062. [*Sosedov, V., Makarevich, M., and Kulida, E.* Solving the Aircraft Landing Problem Using Deep Q-Networks Approach // *Proceedings of 2024 17th International Conference*



- on Management of Large-Scale System Development (MLSD). – Moscow, Russia, 2024. – P. 1–5. – DOI: 10.1109/MLSD61779.2024.10739636]
55. *Brittain, M., Wei, P.* Autonomous Aircraft Sequencing and Separation with Hierarchical Deep Reinforcement Learning // Proceedings of the International Conference for Research in Air Transportation (ICRAT2018). – Barcelona, Spain, 2018.
56. *Degas, A., Islam, M.R., Hurter, C., et al.* A Survey on Artificial Intelligence (AI) and eXplainable AI in Air Traffic Management: Current Trends and Development with Future Research Trajectory // Applied Sciences. – 2022. – Vol. 12, no. 3. – Art. no. 1295. – DOI: 10.3390/app12031295
57. *Кулида Е.Л.* Генетический алгоритм решения задачи оптимизации последовательности и времен посадок воздушных судов // Автоматика и телемеханика. – 2022. – № 3. – С. 156–168. – DOI: 31857/S0005231022030114. [*Kulida, E.L.* Genetic Algorithm for Solving the Problem of Optimizing Aircraft Landing Sequence and Times // Automation and Remote Control. – 2022. – Vol. 83, no. 3. – P. 426–436.]
58. *Кулида Е.Л., Лебедев В.Г., Егоров Н.А.* Исследование эффективности алгоритма оптимизации потока воздушных судов на посадку // Проблемы управления. – 2019. – № 6. – С. 63–69. – DOI: 10.25728/ru.2019.6.7. [*Kulida, E.L., Lebedev, V.G., Egorov, N.A.* Study of the Effectiveness of the Algorithm to Optimize the Flow of Aircraft on Landing // Control Sciences. – 2019. – No. 6. – P. 63–69. – DOI: 10.25728/ru.2019.6.7. (In Russian)]
59. *Кулида Е.Л.* Инструментальное средство для исследования алгоритмов построения оптимальных очередей воздушных судов на посадку // Свидетельство о государственной регистрации программы для ЭВМ № 2018615221 РФ. Зарег. 03.05.2018. [*Kulida, E.L.* An Instrumental Tool for Studying Algorithms for Constructing Optimal Aircraft Landing Sequences // Certificate of State Registration of a Computer Program no. 2018615221, Russian Federation. Registered May 3, 2018. (In Russian)]
60. *Егоров Н.А.* Библиотека точных методов построения оптимальных очередей воздушных судов на посадку для 4-х критериев оптимальности, реализованная в виде .Net сборки // Свидетельство о государственной регистрации программы для ЭВМ № 2019616970 РФ. Зарег. 03.06.2019. [*Egorov, N.A.* A Library of Precise Methods for Constructing Optimal Aircraft Landing Sequences by Four Optimality Criteria, Implemented as a .Net Assembly // Certificate of State Registration of a Computer Program no. 2019616970, Russian Federation. Registered June 3, 2019. (In Russian)]
61. *Кулида Е.Л.* Реализация генетического алгоритма решения задачи оптимизации последовательности и времен посадок воздушных судов // Свидетельство о государственной регистрации программы для ЭВМ № 2022616075 РФ. Зарег. 18.04.2022. [*Kulida, E.L.* Implementation of a Genetic Algorithm for Optimizing the Sequence and Timing of Aircraft Landings // Certificate of State Registration of a Computer Program no. 2022616075, Russian Federation. Registered April 18, 2022 (In Russian)]
62. *Кулида Е.Л.* Система анализа алгоритмов оптимизации последовательности и времен посадок прибывающих воздушных судов / Труды 13-й Международной конференции «Управление развитием крупномасштабных систем» (MLSD'2020, Москва). – Москва, 2020. – С. 1528–1534. [*Kulida, E.L.* Analysis of Algorithms for Solving the Aircraft Landing Problem // Proceedings of 2020 13th International Conference on Management of Large-Scale System Development (MLSD). – Moscow, Russia, 2020. – P. 1–4. – DOI: 10.1109/MLSD49919.2020.9247839]
63. *Razzaghi, P., Tabrizian, A., Guo, W., et al.* A Survey on Reinforcement Learning in Aviation Application // Engineering Applications of Artificial Intelligence. – 2024. – Vol. 136, no. 3. – DOI: 10.1016/j.engappai.2024.108911
64. *Иванова П.И., Печенежский В.К., Чуви́ковская Е.К.* Применение технологий машинного обучения и нейронной сети для выбора оптимальных траекторий полетов воздушных судов в процессе управления воздушным движением // Автоматизация. Современные технологии. – 2024. – № 12. – DOI: 10.36652/0869-4931-2024-78-12-553-557. [*Ivanova, P.I., Pechenejskiy, V.K., Chuvikovskaya, E.K.* Application of Machine Learning and Neural Network Technologies for Selecting Optimal Aircraft Flight Trajectories in Air Traffic Control // Automation. Modern Technologies. – 2024. – No. 12. – DOI: 10.36652/0869-4931-2024-78-12-553-557. (In Russian)]
65. *Саттон Р.С., Барто Э.Г.* Обучение с подкреплением // М.: ДМК-Пресс, 2020. – 552 с. [*Sutton, R.S., Barto, A.G.* Reinforcement Learning: An Introduction. Second edition, in progress. – Cambridge: MIT Press, 2014, 2015. – 352 p.]
66. *Watkins, C.J., Dayan, P.* Q-learning // Machine Learning. – 1992. – Vol. 8. – P. 279–292.
67. *Моралес М.* Грокаем глубокое обучение с подкреплением // СПб.: Питер, 2023. – 464 с. [*Morales, M.* Grokking Deep Reinforcement Learning // London: Dimensions, 2020. – 472 p.]

Статья представлена к публикации членом редколлегии
А. А. Лазаревым.

Поступила в редакцию 18.09.2025,
после доработки 27.11.2025.
Принята к публикации 16.12.2025.

Кулида Елена Львовна – канд. техн. наук,
✉ elena-kulida@yandex.ru
ORCID iD: <https://orcid.org/0009-0003-0226-9708>

Лебедев Валентин Григорьевич – д-р техн. наук,
✉ lebedev-valentin@yandex.ru
ORCID iD: <https://orcid.org/0000-0003-3206-9558>

Егоров Николай Александрович – канд. техн. наук,
✉ negorov@bk.ru
ORCID iD: <https://orcid.org/0009-0006-5141-1207>

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

© 2026 г. Кулида Е. Л., Лебедев В. Г., Егоров Н. А.



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

METHODS FOR SOLVING THE AIRCRAFT LANDING OPTIMIZATION PROBLEM

E. L. Kulida*, V. G. Lebedev**, and N. A. Egorov***

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

*✉ elena-kulida@yandex.ru, **✉ lebedev-valentin@yandex.ru, ***✉ negorov@bk.ru

Abstract. This paper considers the problem of optimizing the sequence and time of aircraft landings, which is topical for increasing the capacity of airport runways. The main approaches to solving this problem are briefly overviewed, and its mathematical statement is provided, including the key constraints and objective functions. The exact solution method and approximate ones using genetic algorithms with two different crossover operators and a heuristic algorithm are described. The time to obtain an exact solution grows exponentially with increasing the problem dimension (the number of aircraft), which makes it impractical. Approximate methods yield a suboptimal, albeit quite good, solution in real time during aircraft landing. An original simulation software complex is presented and applied to compare the efficiency of the main methods implemented. It is proposed to use deep reinforcement learning to solve the problem.

Keywords: airport runway capacity, separation criteria between aircraft landings, optimality criteria, library of methods.



ЕДИНОЕ ПОЛЕ ВЕРОЯТНОСТИ ОБНАРУЖЕНИЯ ДЛЯ ГРУППЫ СТАЦИОНАРНЫХ НАБЛЮДАТЕЛЕЙ

И. М. Рудько

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

✉ igor-rudko@mail.ru

Аннотация. Исследуется задача вычисления единого поля вероятности обнаружения для группы стационарных наблюдателей, контролирующих заданный район и работающих в активном режиме. Рассматриваются расположенные на плоскости, снабженной декартовой системой координат, стационарные наблюдатели, координаты которых известны, и рассчитывается единое поле вероятности обнаружения объекта, имеющего заданные координаты, для всех стационарных наблюдателей района. Показано, что наибольшую сложность представляют объекты, для которых необходимо учитывать рельеф местности. Если для обнаружения объектов без учета рельефа достаточно однослойной (плоской) карты вероятности обнаружения, то для таких объектов необходим расчет многослойной карты, включающей высоту движения объекта. На примерах показано, что вероятность обнаружения объекта зависит от места и высоты установки наблюдателей, высоты движения объекта наблюдения и рельефа местности. Расчет единого поля вероятности обнаружения для заданного района в виде многослойной карты позволяет существенно упростить как задачу оптимизации расположения наблюдателей, так и задачу управления подвижными объектами в конфликтной среде.

Ключевые слова: вероятность обнаружения объекта, поле вероятности обнаружения объекта, траектория движения объекта, рельеф местности.

ВВЕДЕНИЕ

При решении задач обнаружения объекта в активном режиме используются закономерности распространения радиоволн в пространстве, основными из которых являются: постоянство скорости распространения сигнала; прямолинейность распространения сигнала; направленность излучения и приема сигнала, в основе которой лежит явление интерференции радиоволн; эффект Доплера. При этом зондирующий и отраженный сигналы распространяются по прямолинейной траектории и без искажения своей формы [1]. Дальность прямой видимости для наблюдателя, стоящего на абсолютно ровной поверхности, с учетом кривизны земной поверхности по объекту на высоте 50 м, будет составлять около 39 км и увеличиваться с ростом высоты объекта. Дальность обнаружения объекта, движущегося на высоте 100 м, составит примерно 52-55 км. Для успешного обнаружения движущего-

ся объекта он должен находиться в зоне прямой видимости. По объектам, на обнаружение которых не влияет рельеф местности, можно создать так называемое сплошное поле обнаружения. С учетом рельефа местности сделать это значительно сложнее. Отметим, что в СССР так и не было создано сплошное поле обнаружения даже в европейской части страны. Если движение объекта происходит на предельно малых высотах (ПМВ) – от нескольких десятков до сотен метров [2] – то рельеф местности, лес, здания и сооружения создают так называемые углы закрытия. Сегодня 99 % перемещений движущихся объектов происходит на высотах 200–300 м, а над сильно пересеченной местностью и выше, поскольку используется измерение высоты полета либо по спутниковым системам глобального позиционирования, либо по барометрическим высотомерам. Внедрение лазерных высотомеров позволит снизить эту высоту до 30–50 м, что еще больше затруднит их обнаружение.

Обработка наблюдателем отраженных от объекта сигналов осуществляется в три этапа [1]:

- *Первичная обработка* включает операции обнаружения и измерения (оценки) параметров принимаемых сигналов. Первичная обработка проводится непосредственно наблюдателем. Совокупность оценок параметров сигналов образует отметку от объекта.

- *Вторичная обработка* производится по совокупности отметок и обеспечивает формирование траекторной информации.

- *Третичная обработка* состоит в объединении и отождествлении информации отдельных наблюдателей, входящих в систему, или информации отдельных систем.

В настоящей работе рассматривается задача обнаружения объекта с учетом рельефа местности в районе, контролируемом несколькими стационарными наблюдателями, т. е. рассматривается первичная и третичная обработка.

Зона потенциально возможного обнаружения объекта зависит от рельефа местности, т. е. обладает контрастной структурой, с чередованием зон освещенности (объект в зоне прямой видимости) и зон тени (объект вне зоны прямой видимости). Контрастная структура зоны наблюдения может выражаться любыми показателями – либо как интенсивность полезного сигнала, либо как вероятность обнаружения полезного сигнала, лишь бы они представляли собой корректную свертку всех основных влияющих факторов и были доступны для оценки. Непосредственно такую возможность характеризует **вероятностный критерий**, представляющий собой вероятность такого события, при котором за время прохождения объектом маршрута его обнаружит хотя бы один из L ($L \geq 1$) наблюдателей, расположенных в районе. Обозначим эту вероятность $P_{\text{обн}}$ и будем называть ее вероятностью обнаружения объекта. В случае независимых наблюдателей оценки вероятности наступления события обнаружения хотя бы одним наблюдателем производятся по формуле интегральной (кумулятивной, накопленной) вероятности [3]:

$$P_{\text{обн}} = 1 - \prod_{i=1}^L (1 - P_{\text{обн}i}), \quad (1)$$

где L – число наблюдателей; $P_{\text{обн}i}$ – вероятность обнаружения i -м наблюдателем. Такая задача относится к области третичной обработки. Знание координат расположения наблюдателей позволяет, используя формулу (1), рассчитать интегральную вероятность $P_{\text{обн}}(x, y)$ для всех точек области, контролируемой этими наблюдателями.

В работе [4] рассматриваются расположенные на плоскости, снабженной декартовой системой координат, стационарные средства гидроакустического наблюдения, работающие в пассивном режиме, координаты которых известны, и рассчитывается единое поле вероятности обнаружения объекта для всех стационарных гидроакустических средств по всему району $P_{\text{обн}}(x, y)$ с учетом анизотропии гидроакустического поля.

В настоящей работе рассматриваются расположенные на плоскости, снабженной декартовой системой координат XOY , стационарные наблюдатели, работающие в активном режиме, координаты которых известны, и рассчитывается единое поле вероятности обнаружения объекта для всех наблюдателей по всему району $P_{\text{обн}}(x, y)$ в зависимости от высоты установки наблюдателей, высоты объекта наблюдения и рельефа местности.

1. ОСНОВНЫЕ СООТНОШЕНИЯ

Рассмотрим уравнение максимальной дальности обнаружения [5]

$$R_{\text{max}} = \sqrt[4]{\frac{E_{\text{изл}} G A \delta_{\text{ц}}}{E_{\text{пр max}} (4\pi)^2}},$$

где $E_{\text{изл}}$ – излучаемая мощность; $E_{\text{пр max}}$ – мощность принимаемого наблюдателем сигнала; G – коэффициент направленного действия излучающей антенны; $\delta_{\text{ц}}$ – эффективная поверхность рассеяния объекта (ЭПР); A – эффективная площадь приемной антенны; R_{max} – максимальная дальность обнаружения, при которой для расстояния $R \leq R_{\text{max}}$ обнаружение цели обеспечивается с вероятностью правильного обнаружения ($P_{\text{обн}}$), не меньшей допустимой для заданных вероятностей ложной тревоги $P_{\text{лт}}$ и минимально допустимой вероятности обнаружения $P_{\text{обн min}}$. Данное определение в общем виде относится к однократному (мгновенному) наблюдению при установленных зависимостях между временем наблюдения T_0 , отношением сигнал/помеха ρ , $P_{\text{обн}}$ и $P_{\text{лт}}$.

Параметры $E_{\text{изл}}$, G , A относятся к наблюдателю и, следовательно, от R не зависят. Параметр $\delta_{\text{ц}}$ также не зависит от дальности до объекта R [1]. Энергия принимаемого сигнала, как и его мощность, обратно пропорциональна четвертой степени расстояния до объекта [5]. Таким образом, мощность принимаемого наблюдателем сигнала можно представить как

$$E_{\text{пр}} = \beta / R^4, \quad (2)$$

где под β сведены все не зависящие от R параметры.



Критерием возможности обнаружения сигнала на фоне шумов различного происхождения является соотношение сигнал/помеха ρ , определяемое как [6]

$$\rho = 2 \int_{\tau} P_{\text{пр}} dt / P_{\text{ш}}, \quad (3)$$

где $P_{\text{пр}}$ – мощность полезного сигнала на входе приемного устройства наблюдателя, а $P_{\text{ш}}$ – спектральная плотность мощности шумов. Интегрирование проводится по времени посылки сигнала τ . То есть задача сводится к классическому энергетическому приемнику.

Таким образом, задача обнаружения является задачей проверки двух гипотез: H_0 (принимается только шум) и H_1 (принимается смесь сигнала и шума). Решение о наличии сигнала принимается по критерию Неймана – Пирсона, согласно которому оптимальная система обнаружения должна максимизировать вероятность правильного обнаружения $P_{\text{обн}}$ при фиксированной вероятности ложной тревоги $P_{\text{лт}}$, т. е. $P_{\text{обн}} \Rightarrow \max$ при $P_{\text{лт}} = \text{const}$ [1].

2. РАСЧЕТ ЕДИНОГО ПОЛЯ ВЕРОЯТНОСТИ ОБНАРУЖЕНИЯ ДЛЯ НЕСКОЛЬКИХ НАБЛЮДАТЕЛЕЙ

При обработке информации в активном режиме обнаружение осуществляется по результатам обработки излученного и принятого наблюдателем сигнала при наличии помех. Решение о наличии или отсутствии сигнала от объекта принимается периодически, после предварительной обработки (единичного импульса или пачки импульсов) поступившей на интервале наблюдения (накопления) длительностью T_0 реализации гауссовского случайного процесса $X(t)$ с нулевым математическим ожиданием. В отсутствие сигнала от объекта случайный процесс $X(t)$ имеет дисперсию $\sigma_{\text{ш}}^2$, при наличии сигнала от объекта – дисперсию $\sigma_{\text{с}}^2 + \sigma_{\text{ш}}^2$.

Известно [7], что плотность распределения статистики y для энергетического приемника (3), как распределение суммы квадратов нормальных случайных величин с нулевым средним и дисперсией σ^2 , описывается центральным χ^2 -распределением:

$$f(y) = \frac{1}{(2\sigma^2)^{n/2} \Gamma(n/2)} y^{n/2-1} e^{-y/2\sigma^2}, \quad y \geq 0,$$

где $\Gamma(\cdot)$ – гамма-функция; n – число степеней свободы, которое в случае цифрового анализа определяется числом N усредняемых выборочных оценок энергии $n = 2N$, а при аналоговой обработке $n = 2T_0 \Delta f$. Здесь T_0 – длительность излученного сигнала (импульса), Δf – полоса пропускания фильтра.

Числовые характеристики (два первых момента) χ^2 -распределения определяются следующими выражениями [8]:

$$m_{\chi} = n\sigma^2, \quad \sigma_{\chi}^2 = 2n\sigma^4.$$

Для гипотезы H_0 $m_0 = n\sigma_{\text{ш}}^2$, $\sigma_0^2 = 2n\sigma_{\text{ш}}^4$, а для гипотезы H_1 $m_1 = n(\sigma_{\text{ш}}^2 + \sigma_{\text{с}}^2)$, $\sigma_1^2 = 2n(\sigma_{\text{ш}}^2 + \sigma_{\text{с}}^2)^2$. Определим отношение сигнал/помеха как $\rho = \sigma_{\text{с}}^2 / \sigma_{\text{ш}}^2$, тогда

$$m_1 = n\sigma_{\text{ш}}^2(1+\rho), \quad \sigma_1^2 = 2n[\sigma_{\text{ш}}^2(1+\rho)]^2 = 2n\sigma_{\text{ш}}^4(1+\rho)^2. \quad (4)$$

При достаточно больших n χ^2 -распределение хорошо аппроксимируется нормальным распределением $N(m_{\chi}, \sigma_{\chi}^2)$.

На основании формул (2)–(4) можно построить зависимость $P_{\text{обн}}$ при фиксированной $P_{\text{лт}}$ от дальности до цели $R - P_{\text{обн}}(R)$ или зависимость $P_{\text{обн}}$ от ЭПР цели – $P_{\text{обн}}(\delta_{\text{ц}})$. Качественный характер этих зависимостей для двух фиксированных значений ЭПР $\delta_{\text{ц}}$ приведен на рис. 1, где $\delta_{\text{ц1}} > \delta_{\text{ц2}}$. Применительно к обзорным наблюдателям обычно считается, что надежное обнаружение обеспечивается при $\rho > 25$ [6].

Данная зависимость описывает этап первичной обработки. Для заданной вероятности ложной тревоги $P_{\text{лт}} = \alpha$ вероятность правильного обнаружения $P_{\text{обн}}$ возрастает с увеличением ρ , а ρ монотонно убывает с ростом расстояния R до цели.

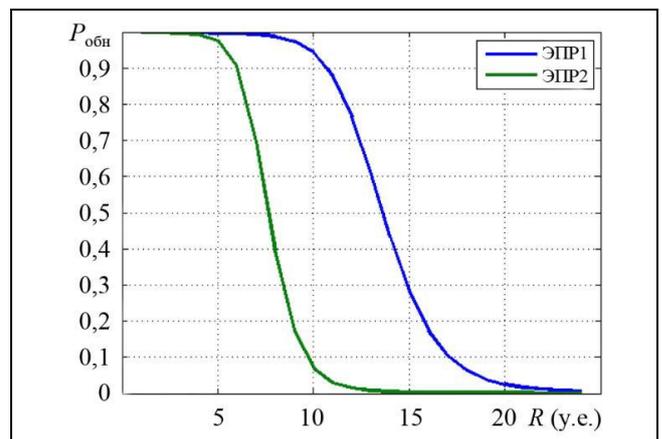


Рис. 1. Зависимость $P_{\text{обн}}$ от дальности до цели при фиксированной $P_{\text{лт}}$ для различных ЭПР

При решении задачи обнаружения необходимо учитывать влияние рельефа местности, т. е. $P_{\text{обн}}(R)$ является условной плотностью вероятности

$$\{P_{\text{обн}}(R) | P_{\text{зпв}}\}, \quad (5)$$

где $P_{\text{зпв}} = 1$ в зоне прямой видимости и $P_{\text{зпв}} = 0$ вне зоны прямой видимости (в зоне тени).

Как правило, приемная система наблюдателя работает периодически, каждый раз накапливая и обрабатывая сигнал в течение фиксированного времени T (времени разового наблюдения), за которое положение обнаруживаемого объекта и его скорость практически не изменяются. По истечении этого времени принимается решение об отсутствии или наличии объекта.

Учитывая, что каждый наблюдатель может работать со своим временем накопления, формулу вычисления интегральной вероятности (1) применить невозможно. Если бы удалось привести времена накопления (обзора) для всех наблюдателей к единому времени (базовому) $T_6 = \text{const}$, то можно было бы построить единую карту вероятности $P_{\text{обн}}(x, y)$ в координатах XOY для всех наблюдателей, находящихся в заданном районе.

В статье [4] показано, что для группы наблюдателей, имеющих различные времена разового наблюдения T_i , вероятности необнаружения за время T_i , $P_{\text{необн } i}(T_i)$, можно пересчитать в вероятности необнаружения за базовое время T_6 , $P_{\text{необн } i}(T_6)$, используя формулу

$$P_{\text{необн } i}(T_6) = [P_{\text{необн } i}(T_i)]^{T_6/T_i}. \quad (6)$$

Следовательно, вероятности необнаружения для двух наблюдателей с различными временами обзоров можно объединить, используя формулу (6), в вероятность необнаружения с единым временем (пересчитывается ко времени наблюдения одного из наблюдателей)

$$\begin{aligned} P_{\text{необн}\Sigma}(T_2) &= P_{\text{необн}2}(T_2)P_{\text{необн}1}(T_2) = \\ &= P_{\text{необн}2}(T_2)[P_{\text{необн}1}(T_1)]^{T_2/T_1}. \end{aligned}$$

Пусть наблюдатель, расположенный в районе, имеет координаты (x_l, y_l, H_l) , где H_l – высота установки наблюдателя. Тогда, используя зависимость $P_{\text{обн}}(R)$ для заданной высоты h , приведенную выше, можно, учитывая выражение (5), рассчитать вероятность необнаружения для всех точек района (x, y) , т. е. сформировать поле (матрицу) условных вероятностей необнаружения $P_{\text{необн}}(x, y | h)$.

Рассчитав по формуле (6) приведенные вероятности необнаружения для каждого наблюдателя, можно рассчитать единое (суммарное) поле для всех наблюдателей.

$$P_{\text{необн}}(x, y) = \prod_{l=1}^L P_{\text{необн } l}(x_l, y_l, H_l)^{T_M/T_l}, \quad (7)$$

где L – число наблюдателей; T_l – длительность обзора для l -го наблюдателя; $T_M = \max[T_l]$; (x_l, y_l) – координаты l -го наблюдателя; $P_{\text{необн } l}$ – вероятности необнаружения для l -го наблюдателя. Таким образом,

$$P_{\text{обн}}(x, y) = 1 - P_{\text{необн}}(x, y).$$

Кроме задания единого времени наблюдения для всех наблюдателей для расчета единого поля вероятности обнаружения, требуется выполнение еще одного очевидного условия – **равенства вероятности ложных тревог** для всех наблюдателей $P_{\text{лт}} = \alpha$.

3. РЕЗУЛЬТАТЫ МОДЕЛИРОВАНИЯ

Моделирование проводилось с использованием пакета MATLAB.

На рис. 2 представлена карта, на которой проводилось моделирование (в правой части рисунков приведена цветовая шкала высот; здесь и далее

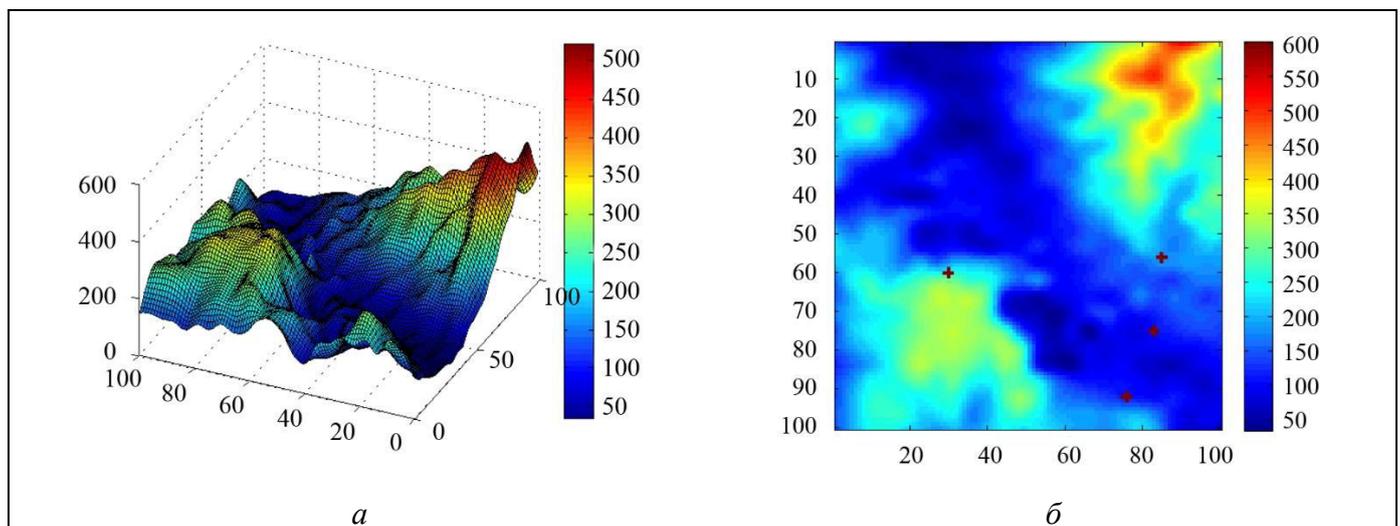


Рис. 2. Карта, на которой проводилось моделирование: a – проекция 3D, b – проекция 2D

шкалы высот и расстояний приведены в условных единицах (у.е.).

Рассматриваются два стационарных наблюдателя с координатами (x, y) и высотой установки H (соответствует высоте точки по карте + 5 у. е. над поверхностью земли):

– первый наблюдатель – $(x_1 = 30, y_1 = 60, H_1 = 263)$;

– второй наблюдатель может быть расположен в точках с тремя возможными координатами (на рис. 4, б отмечены значком «+») – $(x_2 = 85, y_2 = 56, H_2 = 193)$, $(x_2 = 83, y_2 = 75, H_2 = 104)$ и $(x_2 = 76, y_2 = 92, H_2 = 119)$.

Расстояние между первым и вторым наблюдателем во всех трех вариантах расстановки одинаково.

Рассматриваются два варианта высоты полета объекта – $h_1 = 600$ у. е. (высотный объект) и $h_2 = 200$ у. е. (ПМВ).

На рис. 3 (правая часть рисунков – цветовая

шкала вероятности обнаружения) приведен пример расчета суммарного поля вероятности обнаружения (ПВО) для двух наблюдателей при работе по высотному объекту (h_1).

Из рис. 3 видно, что оба наблюдателя имеют круговую диаграмму направленности, но так как первый наблюдатель установлен на склоне холма (см. рис. 2), то склон создает для него зону тени (см. нижнюю часть рис. 3, а). Поля вероятности обнаружения первого и второго наблюдателей не перекрываются, но учитывая, что $T_1 > T_2$, после пересчета к единому времени наблюдения по формуле (6) их единое ПВО практически образует сплошное поле вероятности обнаружения (рис. 3, з).

На рис. 4 приведен пример расчета суммарного ПВО для двух наблюдателей при работе с учетом рельефа местности (h_2). Правая часть рис. 4, а – цветовая шкала высот, правая часть рис. 4, б – цветовая шкала вероятности обнаружения.

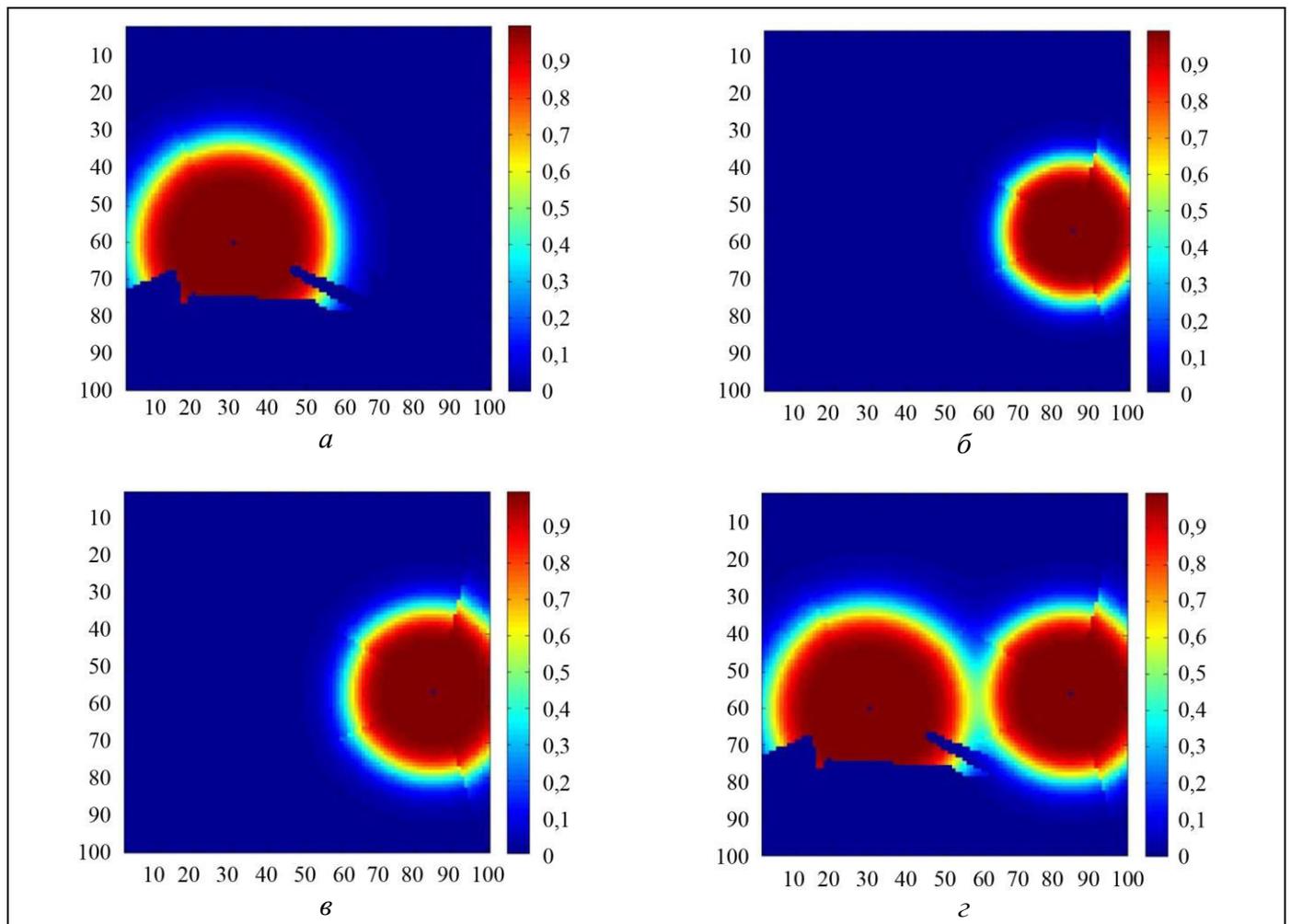


Рис. 3. Суммарное поле вероятности обнаружения для двух РЛС при работе по высотному объекту (h_1): а – ПВО для первого наблюдателя (время обзора – T_1) – $P_{обн1}$; б – ПВО для второго наблюдателя (время обзора – T_2) – $P_{обн2}$; в – ПВО для второго наблюдателя – $P_{обнз}$, пересчитанное по формуле (6) на время T_1 ; з – единое ПВО для двух наблюдателей – $P_{обн} = 1 - P_{исобн}$ по формуле (7)

На рис. 4, *a* приведено сечение географической карты по высоте движения объекта и показан рельеф, превышающий h_2 , который является районом, где объект не может появиться. Этот рельеф задает

«маску запрещенных к нахождению объекта областей», так как высота рельефа в них больше высоты движения объекта h . На картах ПВО отмечено зеленой заливкой.

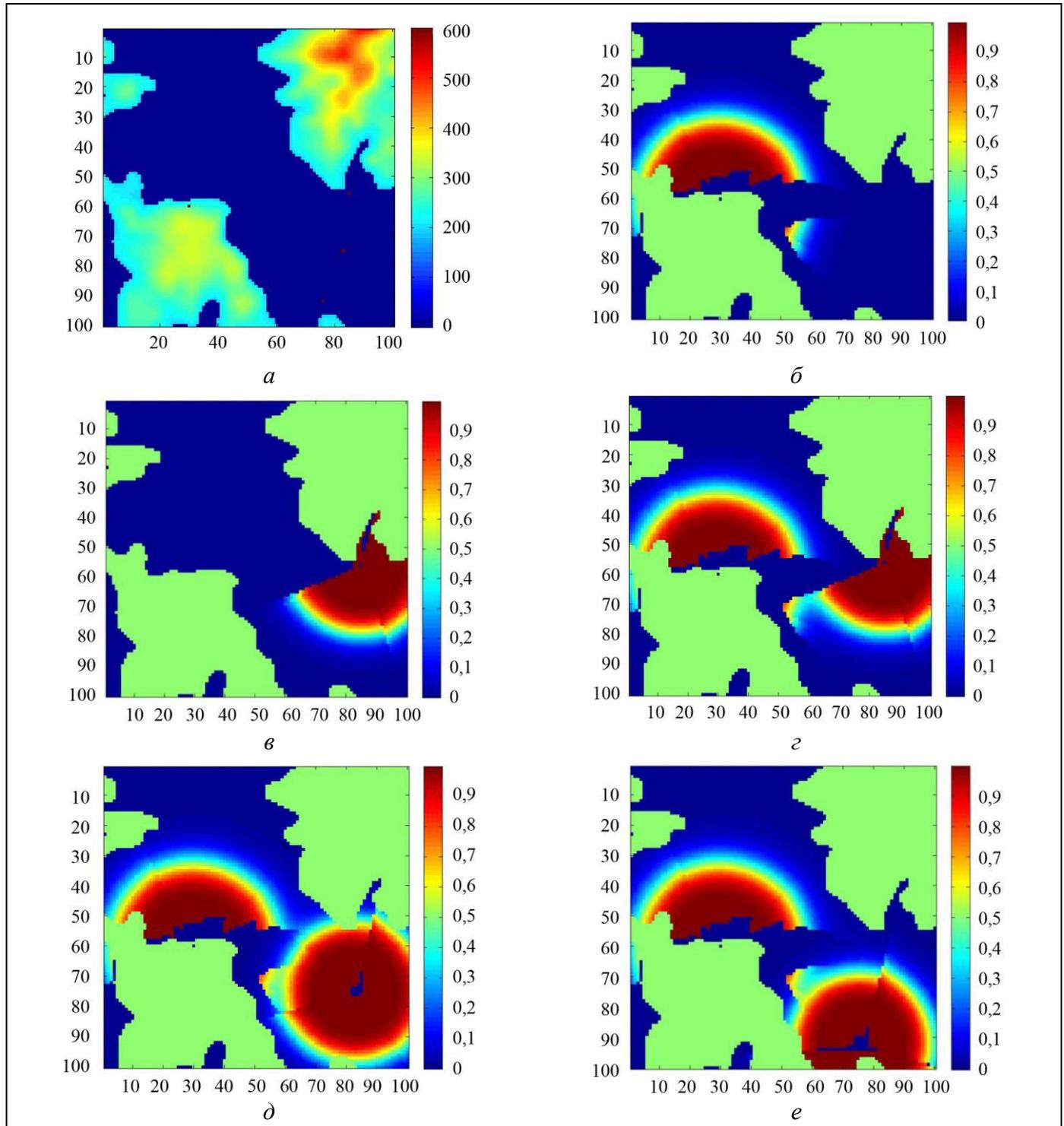


Рис. 4. Единое ПВО для двух наблюдателей при работе с учетом рельефа местности (h_2): *a* – сечение карты по высоте h_2 ; *б* – ПВО для первого наблюдателя (время обзора – T_1) – $P_{\text{обн}1}$; *в* – ПВО для второго наблюдателя – $P_{\text{обн}2}$, пересчитанное по формуле (4) на время T_1 ; *г* – единое ПВО для двух наблюдателей – $P_{\text{обн}} = 1 - P_{\text{необн}}$ по формуле (7), когда координаты второй РЛС – $(x_2 = 85, y_2 = 56, H_2 = 193)$; *д* – единое ПВО для двух наблюдателей – $P_{\text{обн}} = 1 - P_{\text{необн}}$ по формуле (7), когда координаты второго наблюдателя – $(x_2 = 83, y_2 = 75, H_2 = 104)$; *е* – единое ПВО для двух наблюдателей – $P_{\text{обн}} = 1 - P_{\text{необн}}$ по формуле (7), когда координаты второго наблюдателя – $(x_2 = 76, y_2 = 92, H_2 = 119)$

На рис 4, б хорошо видна зона тени, образованная склоном холма, на котором установлен первый наблюдатель. Появление тени объясняется тем, что $H_1 > h_2$ и склон холма образует угол закрытия, полностью перекрывающий в этом секторе зону обзора первого наблюдателя по высоте h_2 .

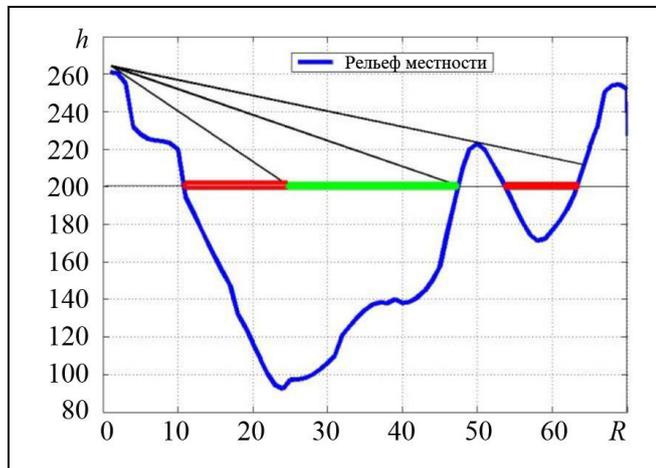


Рис. 5. Зоны видимости и тени

На рис. 5 приведен пример образования зон прямой видимости (зеленая окраска) и зон тени (красная окраска) по заданному направлению для координат карты ($x_1 = 30$, $y_1 = 60$, $H_1 = 266$) и ($x_2 = 100$, $y_2 = 40$) для $h = 200$.

Из сравнения рис. 4, а–е видно, что при обнаружении объекта с учетом рельефа взаимное расположение наблюдателей играет основную роль при формировании единого ПВО.

С учетом рельефа круговая диаграмма ПВО сильно деформируется и зависит уже от четырех параметров:

- расстояния между объектом и наблюдателем R ,
- высоты движения объекта h ,
- высоты установки антенны наблюдателя H ,
- направления на объект.

Тем не менее, для фиксированной высоты движения объекта h также можно рассчитать вероятность обнаружения для всех точек района $P_{\text{обн}}(x, y)$, т. е. построить карту вероятности обнаружения для $h = \text{const}$. Такой набор карт вероятности можно описывать трехмерной матрицей в координатах (x, y, h) для ряда фиксированных высот h_j ($1 \leq j \leq J$) при фиксированных координатах и высотах каждого наблюдателя H_l ($1 \leq l \leq L$) – получается так называемая многослойная карта.

Очевидно, что описанный выше расчет единого ПВО по формулам (5)–(7) легко обобщается на произвольное число наблюдателей.

4. ПРИМЕНЕНИЕ ЕДИНОГО ПОЛЯ ВЕРоятНОСТИ ОБНАРУЖЕНИЯ

Существует достаточно большой круг задач, для которых расчет единого ПВО нескольких наблюдателей представляет интерес. Эти задачи можно условно разделить на две группы.

Первая группа. При решении задачи обнаружения высотного объекта вычисление единого поля вероятности $P_{\text{обн}}(x, y)$ для всего района позволяет объединить информацию от отдельных наблюдателей, имеющих отличающиеся друг от друга технические характеристики (см. рис. 3). Карта при этом делается одномерной.

Вычисление единого поля вероятности $P_{\text{обн}}(x, y, h)$ для всего района позволяет наложить это поле на географическую карту. Тогда оператор может визуально (как известно, глаз человека очень хорошо справляется с такой задачей) определять слабые места, т. е. районы, в которых $P_{\text{обн}}$ мала, и, например, отправлять туда дополнительные средства или изменять расположение наблюдателей, чтобы прикрыть слабые места. Расчет единого ПВО в виде многослойной карты позволяет упростить решение задач оптимизации. Например, при решении задачи обнаружения объекта – упростить задачу оптимального размещения фиксированного числа РЛС в районе по заданному критерию. Такая задача к настоящему времени еще не рассматривалась.

Ко **второй группе** относятся задачи пересечения района, контролируемого несколькими стационарными наблюдателями, или уклонения от обнаружения этими наблюдателями. Данная задача относится к классу задач об управлении, получивших в литературе название задач управления подвижными объектами в конфликтной среде. Таким образом, задача сводится к построению маршрута, определяемого по вероятностному критерию. Построение маршрута на однослойной карте с препятствиями является хорошо изученной задачей как в теоретическом, так и в практическом плане. Для ее решения, как правило, используются классические алгоритмы поиска на графах, такие как, например, алгоритм Дейкстры. Они широко и успешно применяются при наличии одного слоя карты. При этом их сложно распространить на многослойные карты [9].

В случае использования вероятностного критерия выбор маршрута заключается в минимизации накопленной вероятности по формуле (1) для однослойной карты [10–12], но при этом не используется единое ПВО. В работе [13] рассмотрена за-

дача выбора траектории движения маневрирующего объекта и закона изменения его скорости при движении объекта в трехмерной анизотропной среде распространения сигнала, когда его пытаются обнаружить несколько наблюдателей, расположенных в заданном районе (вычисление $P_{\text{обн}}$ для каждого наблюдателя проводится отдельно). В этих задачах предварительный расчет единого ПВО существенно облегчил бы динамическое программирование.

Знание многослойной карты единого ПВО района позволяет, например, при движении объекта с огибанием рельефа на заданной высоте легко свести многослойную карту к однослойной, на которой задача построения маршрута решается значительно проще. Кроме того, на однослойной карте оператор может визуальным образом выбрать маршрут, как показано на рис. 6 (использована карта рис. 4, з).

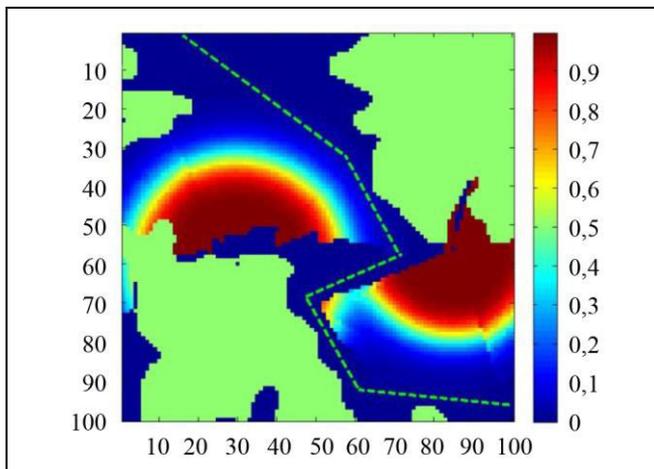


Рис. 6. Маршрут, построенный оператором

ЗАКЛЮЧЕНИЕ

Наличие единого поля вероятности обнаружения для всех наблюдателей, контролирующих заданный район, предоставляет следующие возможности:

- может быть обобщено на случай обнаружения движущегося объекта с учетом рельефа местности, путем добавления еще двух координат: высоты расположения наблюдателя и высоты движения объекта;

- может быть использовано для визуализации, т. е. представления оператору;

- наличие единого поля вероятности для всего района, позволяет использовать его как основу для комплексирования вероятностей обнаружения от других источников информации. Например, от

наблюдателей, расположенных на движущихся объектах [14]. Пассивные локационные системы также могут использоваться для обнаружения и сопровождения движущихся объектов [14]. Кроме локационных методов обнаружения, в настоящее время исследуются нелокационные методы обнаружения подвижных объектов [15], к которым, в частности, относятся: тепловое обнаружение, радиочастотное сканирование, обнаружение с помощью оптических камер, обнаружение акустических сигналов, а также подходы, использующие классические методы машинного обучения с извлечением признаков в качестве предварительной обработки (искусственный интеллект). Включение ПВО от таких источников информации в единое поле вероятности обнаружения возможно при условии наличия моделей для расчетов вероятностей обнаружения объекта этими средствами.

ЛИТЕРАТУРА

1. Радиолокационные системы / под ред. В.П. Бердышева. – Красноярск: Сиб. федер. ун-т, 2011. – 402 с. [*Radiolokatsionnyye sistemy / pod red Berdysheva V.P. Krasnoyarsk: Sib. Federal'nyy univer., 2011. – 402 s. (In Russian)*]
2. Мануйленко В.Г., Удин Е.Г. Теоретические основы крылатых управляемых ракет. – СПб: Университет ИТМО, 2020. – 201 с. [*Manuilenko, V.G., Udin, E.G. Teoreticheskiye osnovy krylatykh upravlyayemykh raket. – SPb: Universitet ITMO, 2020. – 201 s. (In Russian)*]
3. Абчук В.А., Суздаль В.Г. Поиск объектов. – М.: Советское радио, 1977. – 336 с. [*Abchuk, V.A., Suzdal, V.G. Poisk ob'ektov. – Moscow: Sovetskoe radio, 1977. – 336 s. (In Russian)*]
4. Рудько И.М. Единое поле вероятностей обнаружения для разнородных средств наблюдения // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2020. – № 53. – С. 93–101. [*Rudko, I.M. Unified Field of Detection Probabilities for Heterogeneous Means of Observations // Tomsk State University Journal of Control and Computer Science. – 2020. – No. 53. – P. 93–101. (In Russian)*]
5. Теоретические основы радиолокации / под ред. Я.Д. Ширмана. – М.: Советское радио, 1970. – 560 с. [*Teoreticheskiye osnovy radiolokatsii / pod red. Ya.D. Shirmana. – Moscow: Sovetskoye radio, 1970. – 560 s. (In Russian)*]
6. Залогин Н.Н., Калинин В.И., Скнарья А.В. Активная локация с использованием широкополосных хаотических сигналов // РЕНСИТ. – 2011. – Т. 3. – № 1. – С. 3–17. [*Zalogin, N.N., Kalinin, V.I., Sknarya, A.V. The Active Location with the Use of Ultrawide-Band Chaotic Signals // RENSIT. – 2011. – Vol. 3, no. 1. – P. 3–17. (In Russian)*]
7. Левин Б.Р. Теоретические основы статистической радиотехники. Т. 2. – М.: Советское радио, 1968. – 504 с. [*Levin, B.R. Theoretical Foundations of Statistical Radio Engineering. Vol. 2. – Moscow: Sovetskoye radio, 1968. – 504 p. (In Russian)*]
8. Крамер Г. Математические методы статистики. – М.: Мир, 1975. – 648 с. [*Cramer, H. Mathematical Methods of Statistics. – Princeton: Princeton University Press, 1946. – 589 p.*]



9. Моторин Д.Е., Попов С.Г. Алгоритм многокритериального поиска траекторий движения робота на многослойной карте // Информационно-управляющие системы. – 2018. – № 3. – С. 45–53. – DOI: 10.15217/issn1684-8853.2018.3.45 [Motorin, D.E., Popov, S.G. Multi-Criteria Path Planning Algorithm for a Robot on a Multilayer Map // Information and Control Systems. – 2018. – No. 3. – P. 45–53. – DOI: 10.15217/issn1684-8853.2018.3.45 (In Russian)]
10. Dogan, A. Probabilistic Path Planning for UAV // Proceedings of 2nd AIAA “Unmanned Unlimited” Conf. and Workshop & Exhibit. – San Diego, CA, USA, 2003. – P. 1–7.
11. Абрамянц Т.Г., Галяев А.А., Маслов Е.П. и др. Уклонение подвижного объекта в конфликтной среде от обнаружения системой разнородных наблюдателей // Проблемы управления. – 2015. – № 2. – С. 31–37. [Abramyants, T.G., Galyaev, A.A., Maslov, E.P. et al. Evasion of a Moving Object from Detection by a System of Heterogenous Observers in the Threat Environment // Automation and Remote Control. – 2017. – Vol. 72, no. 5. – P. 345–354.]
12. Добровидов А.В., Кулида Е.Л., Рудько И.М. Выбор траектории движения объекта в конфликтной среде // Проблемы управления. – 2011. – № 3. – С. 64–75. [Dobrovodov, A.V., Kulida, E.L., Rudko, I.M. Control of Object Movement in Threat Environment // Control Sciences. – 2011. – № 3. – P. 64–75. (In Russian)]
13. Добровидов А.В., Кулида Е.Л., Рудько И.М. Оптимизация траектории движения объекта по вероятностному критерию в режиме пассивной гидролокации в анизотропной среде // Проблемы управления. – 2014. – № 4. – С. 31–37. [Dobrovodov, A.V., Kulida, E.L., Rudko, I.M. Optimization of the Object Movement Path on the Probabilistic Criterion in the Mode of Passive Sonar in an Anisotropic Medium // Control Sciences. – 2014. – № 4. – P. 31–37. (In Russian)]
14. Khawaja, W., Ezuma, M., Semkin, V., et al. A Survey on Detection, Classification, and Tracking of UAVs Using Radar and Communications Systems // IEEE Communications Surveys and Tutorials. – 2025. – Vol. 28. – P. 3272–3310. – DOI: <https://doi.org/10.1109/COMST.2025.3554613>
15. Zitar, R.A., Al-Betar, M., Ryalat, M., Kassaymeh, S. A Review of UAV Visual Detection and Tracking Methods // Proc. 9th Annual Conf. on Computational Science & Computational Intelligence (CSCI'22). – Las Vegas, NV, USA, 2022. – Paper no. fhal-04108638f.

Статья представлена к публикации членом редколлегии
Б. В. Павловым.

Поступила в редакцию 30.06.2025,
после доработки 08.12.2025.
Принята к публикации 16.12.2025.

Рудько Игорь Михайлович – канд. техн. наук, Институт проблем управления им. В.А. Трапезникова РАН, г. Москва
✉ igor-rudko@mail.ru
ORCID iD: <https://orcid.org/0000-0001-6146-9177>

© 2025 г. И. М. Рудько



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

A UNIFIED DETECTION PROBABILITY FIELD FOR A GROUP OF STATIONARY OBSERVERS

I. M. Rudko

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

✉ igor-rudko@mail.ru

Abstract. This paper addresses the problem of calculating a unified detection probability field for a group of stationary observers monitoring a given area and operating in active mode. Stationary observers are located on a plane equipped with the Cartesian coordinate system, and their coordinates are known. For an object with given coordinates, a unified detection probability field is calculated for all stationary observers in the area. As shown, the greatest complexity arises for the objects necessitating the consideration of the terrain relief. For such objects, it is required to calculate a multilayer map with the altitude of the object’s movement, in contrast to a single-layer (flat) detection probability map, which is sufficient for detecting objects without taking the terrain relief into account. Examples are provided to demonstrate that the probability of object detection depends on the location and altitude of the observers, the altitude of the object’s movement, and the terrain relief. With a unified detection probability field calculated for a given area in the form of a multilayer map, it is much easier to optimize the location of observers (on the one hand) and control moving objects in a conflict environment (on the other).

Keywords: probability of object detection, detection probability field for an object, trajectory of an object, terrain relief.



XXXIII МЕЖДУНАРОДНАЯ КОНФЕРЕНЦИЯ «ПРОБЛЕМЫ УПРАВЛЕНИЯ БЕЗОПАСНОСТЬЮ СЛОЖНЫХ СИСТЕМ»

В декабре 2025 г. в Институте проблем управления им. В.А.Трапезникова РАН состоялась XXXIII Международная научная конференция «Проблемы управления безопасностью сложных систем».

В работе конференции приняли участие 86 авторов из 28 организаций, представивших 63 доклада. Работа конференции велась по следующим секциям:

1. Общетеоретические и методологические вопросы обеспечения безопасности.
2. Проблемы обеспечения экономической и социально-политической безопасности.
3. Проблемы обеспечения информационной безопасности.
4. Кибербезопасность. Особенности обеспечения безопасности в социальных сетях.
5. Экологическая и техногенная безопасность.
6. Методы моделирования и принятия решений при управлении безопасностью сложных систем.
7. Автоматизированные системы и средства обеспечения безопасности сложных систем.

Открывший конференцию председательствующий на пленарном заседании *И.В. Чернов* представил слово *Д.И. Правикову*, который посвятил свое выступление памяти выдающегося российского ученого, доктора технических наук, профессора, автора более ста научных и научно-популярных работ как специального, так и междисциплинарного характера *Сергея Павловича Расторгуева* (1958–2017 гг.). Сфера фундаментальных и прикладных научных интересов С.П. Расторгуева крайне широка и лежит в различных областях информатики, программирования, криптографии, антивирусной защиты, аватаризации (переноса основ биологической жизни и сознания в компьютерную систему), философии и

социологии, теории обучения и педагогики. Сергей Павлович – один из основоположников российской научной школы информационного противоборства, автор многих научных трудов по теории и практике информационных войн, ставших классикой данного в недалеком прошлом принципиально нового научного направления, причем сам термин «информационная война» стал пониматься именно в его трактовке после публикации в 1999 г. одноименной монографии С.П. Расторгуева.

Основа философских взглядов Сергея Павловича сформировалась в период, когда нейронные сети стали актуальным и интенсивно развивающимся научно-техническим направлением. Именно отсюда, как подчеркнуто в выступлении, возник его постулат об одном из основополагающих свойств сетей: обучение осуществляется не только через создание, но и через разрушение связей и уничтожение элементов. Данный и, на первый взгляд, неочевидный тезис положен в основу описания основных свойств информационного оружия – при (само)обучении сложных систем возможно (само)ограничение, подавление системы.

Научное, творческое наследие С.П. Расторгуева настолько глубоко и разнопланово, что, как подчеркнул выступающий, научному сообществу еще только предстоит его постичь и осмыслить. Опираясь интуитивно понятными базовыми смыслами («жизнь», «смерть», «информация», «знание»), С.П. Расторгуев в своих выводах и философских построениях выходил далеко за пределы обыденности и был нацелен на критическое осмысление окружающей действительности. Полученные С.П. Расторгуевым научные результаты остаются исключительно актуальными и в настоящее время.

Программную часть конференции открыл развернутый обзорный доклад *Т.С. Ахромеевой*,



Г.Г. Малинецкого, С.А. Торопыгиной «Новые подходы к системному анализу больших проектов». Как утверждают авторы в первой части доклада, в современных условиях основанные на системном анализе традиционные междисциплинарные подходы к решению задач анализа, структуризации и управления развитием крупномасштабных систем в различных предметных областях переживают глубокий кризис, распавшись на множество слабо связанных друг с другом методологических направлений.

Основной причиной сложившейся ситуации, отмечается в докладе, является то, что, как показывает практика, традиционные подходы не справляются с главными стоящими сегодня задачами проектирования развития больших систем – анализом, прогнозом динамики их развития, управлением рисками. Выход из сложившегося положения авторы видят в создании новых инструментов для решения актуальных задач государственного развития, опирающихся на опыт применения методов и моделей теории самоорганизации – синергетики при анализе сложных проблем и разработке больших проектов. Данная методология может быть эффективно использована для описания свойств и характеристик конкурирующих крупномасштабных проектов, а также соперничества стран, блоков, цивилизаций и этносов в экономике, силовом или ином противостоянии.

В частности, рассматривая преимущества синергетики и анализируя методологию решения рассматриваемых задач, авторы отмечают, что предлагаемый подход позволяет бороться с проблемами «проклятия размерности» путем выделения параметров порядка – ведущих переменных и степеней свободы, которые с течением времени начинают определять динамику других характеристик исследуемой сложной системы. В других случаях, где ключевыми являются так называемые сверхбыстрые процессы, по сравнению с которыми все остальные «замирают», появляется возможность упрощения, связанного с использованием параметров порядка, которые в данном случае, напротив, описывают быстрые переменные.

Вторая часть доклада посвящена обзору публикаций, отражающих использование различных фрагментов рассматриваемого подхода к решению проблем технологического развития военно-промышленного комплекса геополитических противников России на двадцатилетнем временном

горизонте геополитическими противниками России. В заключении авторы доклада отмечают, что высокий уровень напряженности международной обстановки убедительно показывает необходимость развития методов и технологий системного анализа с целью их использования в сфере управления взаимодействием (противоборством) конкурирующих субъектов.

Отличительной чертой настоящей конференции является достаточно большое число интересных и разноплановых по тематике докладов, посвященных поиску путей решения широкого круга проблем обеспечения безопасного и устойчивого поступательного социально-экономического развития России в нынешних крайне непростых условиях.

Доклад В.В. Цыганова «Механизмы безопасного устойчивого развития в многополярном мире» посвящен проблемам экономического развития страны в условиях обостряющихся противоречий и негативного влияния процессов глобализации. Как отмечается в докладе, в настоящее время выдвинутая на конференции ООН по окружающей среде и развитию (Рио-де-Жанейро, 1992 г.) Концепция устойчивого развития, в центр внимания которой помещены интересы не только нынешнего, но и будущих поколений, сегодня используется глобальным центром капитала в своих интересах в качестве инструмента влияния и манипулирования экологическими нормами путем обложения импорта из развивающихся стран экологическим налогом. Подобными методами глобалисты навязывают развивающимся странам завышенные требования к экологичности их производств и выпускаемой продукции, которые не могут быть выполнены без использования дорогостоящих западных технологий. В связи с этим возникает необходимость модификации механизмов устойчивого развития таким образом, чтобы их применение было безопасным для развивающихся стран. В первую очередь это касается механизмов безопасного устойчивого развития ключевых экономических субъектов этих стран, обеспечивающих их самостоятельность, а также возможность адаптации и самоорганизации при региональных или глобальных изменениях.

В докладе рассмотрен разработанный автором робастный механизм решения данной задачи, включающий процедуры формирования показателей безопасности, алгоритмы расчета норм этих

показателей для определения их категорий, а также свертки данных категорий с целью комплексной оценки безопасности устойчивого развития, на основе которой осуществляется стимулирование лиц, принимающих решения. Возможности практического использования разработанного робастного механизма проиллюстрированы на подробном примере обеспечения безопасного устойчивого развития железнодорожного транспорта России с учетом экологических требований в области охраны окружающей среды (задача управления капиталным ремонтом тепловозного парка в рамках соответствующей программы ОАО «РЖД»).

Доклад *А.Е. Абрамова, М.И. Черновой, Ф.С. Левина* «Проблемы развития российского финансового рынка и его стратегического планирования» посвящен комплексной оценке промежуточных результатов реализации федерального проекта «Развитие финансового рынка» и анализу существующих проблем в сфере финансового регулирования и стратегического планирования, препятствующих его успеху. В докладе утверждается, что необходимым условием успешной реализации стратегии развития российского финансового рынка с опорой на внутренние сбережения и инвестиции является ее встраивание в систему стратегического планирования развития российской экономики в целом. Это определяет необходимость более активного участия финансового регулятора в развитии рынка, которое в большей мере должно сводиться к тому, чтобы через исполнение своих классических функций обеспечивать создание условий для роста инновационной активности бизнеса.

Как отмечают авторы, в соответствии с Федеральным законом № 172-ФЗ «О стратегическом планировании в Российской Федерации» Банк России является участником системы государственного долгосрочного планирования. Однако в настоящее время серьезной проблемой является то, что в Федеральном законе № 86-ФЗ «О Центральном банке Российской Федерации (Банке России)» функции мегарегулятора прописаны весьма фрагментарно и сводятся лишь к развитию финансового рынка и обеспечению его стабильности. В результате ряд актуальных задач финансового регулирования и надзора (преодоление «провалов рынка», поддержка конкуренции, защита прав инвесторов и потребителей финансовых услуг, обеспе-

чение устойчивости финансовых организаций и др.) практически не пересекается с целями Банка России, определенными в упомянутом выше федеральном законе.

В докладе подчеркивается: несмотря на достигнутые за прошедшие годы (по ряду направлений значимые) успехи в интеграции политики развития финансового рынка в систему стратегического планирования, что во многом стало возможным благодаря преодолению определенного разрыва между программными документами по развитию финансового рынка и программами развития российской экономики в целом, остается еще ряд требующих решения проблем. На основе приведенных в докладе результатов комплексного и детального анализа основных целеполагающих и нормативно-регулирующих документов, а также ключевых показателей развития финансового рынка России на период до 2030 г. в качестве основной проблемы авторы выделяют недостаточную обоснованность количественных ориентиров, отражающих динамику развития финансового рынка, а также фрагментарность мер, направленных на стимулирование внутренних долгосрочных сбережений. Для решения данной проблемы, как утверждается в докладе, в долгосрочном плане необходимо на законодательном уровне более полно определить ключевые функции финансового регулятора, а также оптимизировать структуры документов стратегического финансового планирования (в том числе устранить дублирование и унифицировать систему прогнозных показателей по перспективам развития финансового рынка на 2030 г., содержащихся в Стратегии развития финансового рынка РФ до 2030 г. и федеральном проекте «Развитие финансового рынка»).

В докладе *А.Д. Козлова, Н.Л. Ноги* «Риски на финансовом рынке и их оценка» представлена разработанная авторами и основанная на совместном использовании эконометрических методов и методов нечеткой логики методика оценки риска при осуществлении операций на финансовом рынке. Структурно данная методика представляет собой последовательность нескольких взаимосвязанных между собой этапов.

На первом этапе проводится обследование финансового рынка и определяется набор параметров, характеризующих как систематические, так и несистематические риски. Сформированный на основе проведенного анализа набор подразделяет-



ся на группы финансовых, внутренних и внешних экономических, а также при необходимости иных параметров. На следующем этапе из полученного исходного набора с использованием экспертных процедур выделяются подмножества в наибольшей степени влияющих на риски финансовых потерь параметров и формируется нечеткая база знаний. Данные параметры представляются в форме лингвистических переменных, нормируемых в диапазоне от 0 до 1. Далее формируется таблица продукционных правил, где каждой строке ставится в соответствие определенное значение уровня риска. На третьем этапе по данной таблице строится линейная модель множественной регрессии и определяются стандартизованные коэффициенты уравнения для их ранжирования. На завершающем четвертом этапе определяются переменные, от которых в наибольшей степени зависит риск финансовых потерь, и проверяется их взаимозависимость. Качество полученной модели оценивается путем вычисления коэффициентов множественной детерминации. Проверка статистической значимости коэффициентов регрессии и уравнения регрессии в целом осуществляется на основе использования F -критерия Фишера и t -критерия Стьюдента. В заключительной части доклада приведен пример, иллюстрирующий практические возможности использования предложенной методики инвесторами для расчета прогнозных оценок возможных финансовых потерь в условиях неопределенности и риска.

Исследованию широкого комплекса методологических и прикладных проблем управления социально-экономическим развитием России, ее регионов и отдельных экономических субъектов посвящена достаточно большая группа представленных на конференции докладов: *Н.И. Комков, В.В. Сулягин, Н.Н. Володина* «Возможное согласование эффективности управления экономическим и социальным развитием»; *В.А. Ириков, Д.Р. Гончар* «Прорывное социально-экономическое развитие страны в 2025–2030 гг.: новые направления, особенности, возможности и примеры кратного роста»; *З.К. Авдеева, С.В. Коврига* «Граф стратегического планирования в системе обеспечения национальной безопасности РФ»; *В.В. Шумов* «Моделирование конфликтов с использованием методов военной кибернетики и учения о безопасности»; *В.В. Ничепорчук* «Интеллектуальные сервисы территориального управле-

ния безопасностью»; *А.В. Рожнов* «Обоснование развития информационно-аналитической системы при реализации гибридных моделей технологии анализа среды функционирования в задачах прогнозного моделирования»; *В.И. Меденников* «Цифровые инструменты, обеспечивающие более экологичный и безопасный путь развития человечества»; *О.Б.о. Байрамов* «Комбинированное применение моделей штрафов и экологического страхования для управления рисками деградации водных бассейнов»; *Т.Х. Усманова, О.В. Демьянова* «Формирование социально-экономической безопасности МТК “Север-Юг”»; *Л.Е. Мистров* «Метод распределения разнородного ресурса для обеспечения конфликтной устойчивости взаимодействия организационно-технических систем»; *Д.Р. Гончар* «Народосбережение как новый целевой показатель более успешного социально-экономического развития страны в 2025–2030 гг.»; *В.О. Сиротюк, Л.В. Богатырева* «Комплексная безопасность субъектов системы управления интеллектуальной собственностью»; *Н.Н. Лантер* «Особенности арктических концепций зарубежных стран в 2025–2030 гг.».

Теоретическим и практическим проблемам разработки методологии сценарного анализа и технологий имитационного моделирования процессов управления развитием сложных социально-экономических систем на государственном, отраслевом, региональном и объектовом уровнях был посвящен целый ряд интересных докладов, среди которых можно отметить работы *Д.А. Кононова, И.В. Чернова* «Уязвимость сложной системы: иерархия понятий»; *Г.В. Гореловой* «Устойчивое и безопасное развитие сложных систем, Юг России, когнитивное моделирование»; *В.Л. Шульца, И.В. Чернова, А.Б. Шелкова* «Анализ методов и технологий верификации моделей сценарного управления»; *М.Ю. Дмитриевой, И.Д. Бутусова, Ю.А. Гоголадзе* «Поддержка управления сложными социально-экономическими системами с применением методов ситуационного сценарного анализа»; *Н.В. Команича* «Использование сценарного анализа в СППР обеспечения региональной безопасности: краткий обзор»; *М.Ю. Дмитриевой, Л.В. Богатыревой* «Сценарная модель исследования угроз безопасного развития городской инфраструктуры»; *Е.Д. Ермолаева* «Перспектива использования сценарного анализа для обеспечения информационной безопасности ИИ».

Доклад *Д.И. Правикова, В.А. Буркина* «Обоснование показателей комплексной оценки защищенности критической информационной инфраструктуры предприятий нефтегазового комплекса» посвящен разработке методов обеспечения безопасности производственной и информационно-технологической инфраструктуры предприятий, осуществляющих добычу углеводородных ископаемых. Как отмечается в представленной работе, объекты нефтегазового комплекса представляют собой сложные социотехнические системы, в которых подсистемы автоматизированного управления технологическими процессами, информационные ресурсы, персонал и регламентирующие документы образуют единую взаимозависимую структуру. Управление их безопасностью требует перехода от исключительно нормативной проверки к количественно верифицируемой интегральной оценке защищенности, позволяющей сопоставлять объекты, приоритизировать организационно-технические мероприятия и обосновывать управленческие решения на основе измеряемых показателей.

В докладе представлена разработанная авторами методика расчета комплексного показателя защищенности критической информационной инфраструктуры предприятий нефтегазового комплекса, представляющего собой взвешенную сумму десяти нормированных частных метрик. Данные метрики отражают широкий диапазон показателей безопасности (охват критичных сегментов информационной структуры централизованным сбором и корреляцией событий безопасности, охват узлов минимально необходимым набором защитных мер, оценки своевременности устранения критических уязвимостей защищаемых объектов, а также внепланового времени простоя технологических систем и оперативности восстановления после инцидента, уровень подготовленности персонала к действиям в нештатных и критических ситуациях и т. д.). Отбор метрик осуществляется с учетом отраслевой специфики, а каждая метрика имеет прямую связь с действиями, предусмотренными соответствующими регламентными, нормативными и иными организационными документами и требованиями, что обеспечивает сопоставимость показателей, возможность аудита результатов и повышение эффективности оперативного управления риском в контуре промышленной и информационной безопасности. В целом предло-

женный подход представляет собой практико-ориентированный инструмент управления защищенностью критической информационной инфраструктуры, связывающий целевые уровни безопасности с конкретными действиями и ресурсами, а также обеспечивающий воспроизводимость оценок и прозрачность принимаемых решений.

Решению разнообразных проблем управления информационной безопасностью посвящена традиционно большая группа представленных на конференции докладов: *Н.Г. Кереселидзе* «Модель информационной безопасности с учетом санкций»; *С.А. Шустов, Р.В. Мецераков* «Повышение безопасности видеоданных за счет использования помехоустойчивого метода видеостеганографии на основе глубоких нейронных сетей»; *Л.Н. Логинова, А.Д. Дроздов* «Подходы и инструменты сбора информации из открытых источников для мониторинга и выявления угроз информационной безопасности»; *Р.Э. Асратян, С.С. Владимирова, Е.А. Курако, В.Л. Орлов* «REST-сервисы, основанные на языке C# и обеспечивающие защиту информации в средах Windows–Linux»; *Д.С. Игнатов* «Алгоритм онлайн мониторинга репутации по поисковым запросам»; *А.Д. Домашкин, Л.Н. Логинова* «Особенности архитектуры и специфика обеспечения информационной безопасности высоконагруженных информационных систем»; *А.А. Широкий* «Управление рисками компьютерной сети с древовидной топологией»; *И.Ф. Михалевич, А.М. Франшико Нелсон* «Моделирование угроз безопасности объектов критической информационной инфраструктуры Республики Ангола»; *В.В. Ведущев, Р.В. Батищев* «Использование принципов бережливого управления для обеспечения информационной безопасности промышленного предприятия»; *И.Ф. Михалевич, Д.И. Пчелинцев* «Использование регулярных выражений для управления информационной безопасностью интеллектуальных транспортных систем»; *Л.А. Баранов, Н.Д. Иванова, И.Ф. Михалевич* «Формализация риска информационной безопасности интеллектуальных системах водного транспорта как нечеткой лингвистической оценки на основе теории принятия решений»; *Ю.А. Клименко, А.П. Преображенский, И.А. Тихонов* «Моделирование системы безопасности умного дома».

Ряд интересных и разнообразных по прикладной тематике докладов посвящен вопросам обеспечения безопасности функционирования про-



мышленных и транспортных систем и объектов, среди которых можно отметить работы *В.В. Лещенко* «Кибернетика безопасности энергетических систем с ядерным реактором»; *Е.В. Кловача, И.А. Кручининой, В.А. Ткаченко* «Оценка результативности при проведении аудита систем управления промышленной безопасностью»; *С.К. Сомова* «Безопасность и надежность сложных технических систем»; *М.В. Мамченко* «Подход к оценке изменения критических характеристик робототехнических систем во времени»; *О.Б. Скворцова, В.И. Сташенко* «Надежность проводящих элементов энергооборудования с импульсной модуляцией»; *В.К. Мусаева* «Математическое моделирование безопасности наземного защитного сооружения при внешнем волновом сейсмическом воздействии»; *Р.Ш. Хабибулина* «Имитационная модель оперативных фаз реагирования на пожары объектов топливно-энергетического комплекса с робастной оптимизацией времени, риска и ресурсов»; *Д.А. Вольфа, Р.Р. Галина* «Общетеоретические и методологические вопросы обеспечения безопасности при разработке полетного контроллера для БПЛА: управление в процессе постановки экспериментов»; *Д.М. Мельника* «Формирование структуры системы управления безопасностью полетов беспилотных авиационных систем на основе интероперабельности»; *А.Г. Давыдовского* «Оценка безопасности системы «беспилотный авиационный комплекс – персонал – среда» на основе сценарного анализа»; *Е.А. Тимме* «Математическая модель влияния факторов космического полета на качество деятельности космонавта и формирования эргономического риска»; *Л.А. Баранова, С.Е. Иконникова, А.Е. Ермаковой* «Безопасность транспортной инфраструктуры и систем управления транспортными средствами»; *О.А. Тельминова* «Модель угроз для системы управления колесной платформой с многоагентной многоуровневой нейросетевой реализацией»; *А.И. Исаковой, А.С. Мецераковой* «Особенности организации перевозочного процесса на строящейся высокоскоростной магистрали «Москва – Санкт-Петербург»»; *А.С. Иконникова* «Разработка микросервиса, предназначенного для анализа и прогнозирования износа элементов контактной сети железных дорог»; *В.А. Зорина* «Применение технологий искусственного интеллекта для обеспечения информационной безопасности медицинских систем и устройств».

Подробно ознакомиться с представленными работами можно в опубликованных в электронном виде материалах¹ на официальном сайте конференции: https://iccss2025.ipu.ru/conf_proceedings.

Проведение очередной XXXIV-й конференции «Проблемы управления безопасностью сложных систем» планируется в ноябре-декабре 2026 г. в Институте проблем управления им. В.А. Трапезникова РАН. О дате и времени проведения конференции будет сообщено в информационном письме оргкомитета, которое будет опубликовано на официальном сайте XXXIV-й конференции (<https://iccss2026.ipu.ru/>), а также разослано участникам прошлых конференций, заинтересованным лицам и профильным организациям. Телефон оргкомитета (495) 198-17-20, доб. 1407, e-mail: iccss@ipu.ru. Технический секретарь конференции – *Альфия Фариссовна Ибрагимова*.

Ученый секретарь Оргкомитета конференции
А. Б. Шелков

Координатор мероприятий Оргкомитета конференции
Л. В. Богатырева

Шелков Алексей Борисович – канд. техн. наук,
✉ abshelkov@gmail.com,
ORCID iD: <https://orcid.org/0000-0003-1408-5212>

Богатырева Лариса Владимировна – канд. ист. наук,
✉ lbogat@mail.ru,
ORCID iD: <https://orcid.org/0000-0003-2744-0404>

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

© 2026 г. Шелков А. Б., Богатырева Л. В.



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

¹ Проблемы управления безопасностью сложных систем: материалы XXXIII Международной конференции, 17 декабря 2025 г., Москва / под общ. ред. А.О. Калашникова, И.В. Чернова; Институт проблем управления им. В.А. Трапезникова РАН Минобрнауки РФ [и др.]. – Электрон. текст. данн. – Москва: ИПУ РАН. – 2025.

33RD INTERNATIONAL CONFERENCE ON PROBLEMS OF COMPLEX SYSTEMS SECURITY CONTROL

A. B. Shelkov* and L. V. Bogatyreva**

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

*✉ abshelkov@gmail.com, **✉ lbogat@mail.ru

Abstract. The conference took place in December 2025. Scientific results presented by the conference participants are briefly described below. The conference included the following sections: general theoretical and methodological issues of security support; problems of economic and sociopolitical security support; problems of information security support; cybersecurity and security aspects in social networks; ecological and technogenic security; modeling and decision-making for complex systems security control; and automated systems and means of complex systems security support. Special attention was paid to the theoretical and applied problems of improving the effectiveness of Russia's economic, information, ecological and technogenic security management processes. In total, 86 authors from 28 organizations presented 63 papers at the conference.

Keywords: conference, complex systems, security control.