

ISSN 2413 - 0133
Scientific journal

ИНФОРМАЦИОННЫЕ И МАТЕМАТИЧЕСКИЕ ТЕХНОЛОГИИ В НАУКЕ И УПРАВЛЕНИИ

№1(25)/2022

6+

ИНФОРМАЦИОННЫЕ И МАТЕМАТИЧЕСКИЕ ТЕХНОЛОГИИ В НАУКЕ И УПРАВЛЕНИИ

Научный журнал

№ 1 (25)



EDITORIAL BOARD

Ablameyko S.V.
 Arshinskiy L.V.
 Berestneva O.G.
 Boukhanovsky A.V.
 Bychkov I.V.
 Woern H.
 Voevodin V.V.
 Wolfengagen V.E.
 Voropai N.I.
 Gornov A.Y.
 Gribova V.V.
 Groumpos P.
 Zorina T.G.
 Kazakov A.L.
 Kalimoldaev M.N.
 Karpenko A.P.
 Komendantova N.P.
 Kureichik V.V.
 Lis R.
 Massel L.V.
 Moskvichev V.V.
 Ovtcharova J.
 Popov G.T.
 Sidorov D.N.
 Smirnov S.V.
 Stylios C.
 Taratukhin V.V.
 Khamisov O.V.
 Hodashinsky I.A.
 Chubarov L.B.
 Yusupova N.I.

РЕДАКЦИОННАЯ КОЛЛЕГИЯ ЖУРНАЛА

Абламейко С.В., академик НАН Беларуси, Минск, БГУ
 Аршинский Л.В., д.т.н., Иркутск, ИрГУПС
 Берестнева О.Г., д.т.н., Томск, ТПУ
 Бухановский А.В., д.т.н., Санкт-Петербург, НИУ ИТМО
 Бычков И.В., академик РАН, Иркутск, ИДСТУ СО РАН
 Вёрн Х., Германия, Karlsruhe Institute of Technology (KIT)
 Воеводин В.В., чл.-корр. РАН, Москва, НИВЦ МГУ
 Вольфенгаген В.Э., д.т.н., Москва, МИФИ
 Воропай Н.И., чл.-корр. РАН, Иркутск, ИСЭМ СО РАН
 Горнов А.Ю., д.т.н., Иркутск, ИДСТУ СО РАН
 Грибова В.В., д.т.н., Владивосток, ИАПУ ДВО РАН
 Грумπος Π., Греция, University of Patras
 Зорина Т.Г., д.т.н., Республика Беларусь, Институт энергетики НАН Беларуси
 Казаков А.Л., д.ф.-м.н., Иркутск, ИДСТУ СО РАН
 Калимолдаев М.Н., академик НАН РК, Республика Казахстан, ИИВТ
 Карпенко А.П., д.ф.-м.н., Москва, МГТУ им. Баумана
 Комендантова Н.П., PhD, Австрия, Лаксенбург, IASA
 Курейчик В.В., д.т.н., профессор ЮФУ, Таганрог
 Лис Р., Польша, Wroclaw University of Science and Technology
 Массель Л.В., д.т.н., Иркутск, ИСЭМ СО РАН
 Москвичев В.В., д.т.н., Красноярск, СКТБ «Наука» СО РАН
 Овчарова Ж., Германия, Karlsruhe Institute of Technology (KIT)
 Попов Г.Т., Болгария, г. София, Технический университет
 Сидоров Д.Н., д.ф.-м.н., Иркутск, ИСЭМ СО РАН
 Смирнов С.В., д.т.н., Самара, ИПУСС РАН
 Стилос Х., Греция, Technological Educational Institute of Epirus
 Таратухин В.В., Германия, ERCIS, University of Muenster
 Хамисов О.В., д.ф.-м.н., Иркутск, ИСЭМ СО РАН
 Ходашинский И.А., д.т.н., Томск, ТУСУР
 Чубаров Л.Б., д.т.н., Новосибирск, ИВТ СО РАН
 Юсупова Н.И., д.т.н., Уфа, УГАТУ

EXECUTIVE EDITORIAL

Chief Editor Massel L.V.
 Executive Editor
 Bakhvalova Z.A.
 Editor Gaskova D.A.
 Editor Massel A.G.
 Designer Pesterev D.V.

ИСПОЛНИТЕЛЬНАЯ РЕДАКЦИЯ ЖУРНАЛА

Главный редактор Массель Л.В. д.т.н., Иркутск, ИСЭМ СО РАН
 Выпускающий редактор Бахвалова З.А. к.т.н., Иркутск, ИСЭМ СО РАН
 Редактор Гаскова Д.А. к.т.н., Иркутск, ИСЭМ СО РАН
 Редактор Массель А.Г. к.т.н., Иркутск, ИСЭМ СО РАН
 Дизайнер Пестерев Д.В. Иркутск, ИСЭМ СО РАН

Адрес учредителя, издателя и редакции

Федеральное государственное бюджетное учреждение науки Институт систем энергетики им. Л.А. Мелентьева Сибирского отделения Российской академии наук (ИСЭМ СО РАН)

664033 г. Иркутск, ул. Лермонтова, 130

Тел: (3952) 42-47-00 Факс: (3952) 42-67-96

Сайт журнала - <https://www.imt-journal.ru/>

Журнал включен в Российский индекс научного цитирования (РИНЦ). Номер контракта 202-04/2016.

Журнал включен в перечень рецензируемых научных изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученых степеней кандидата и доктора наук (по состоянию на 29.03.2022) по научным специальностям 1.2.2; 2.3.1; 2.3.3; 2.3.5; 2.4.5.

Журнал зарегистрирован в Роскомнадзоре. Регистрационный номер ПИ № ФС 77 – 73539.

Отпечатано в полиграфическом участке ИСЭМ СО РАН

CC BY

*Подписка на журнал доступна со второго полугодия 2021г.

Дата выхода 11.04.2022 г. Тираж 100 экз.

© Издательство ИСЭМ СО РАН

Цена свободная. (6+)

Предисловие редактора	5
Математическое моделирование	
Донской И.Г. Определение критических условий в задаче о тепловом взрыве с помощью приближенных вариационных формулировок	7
Фоменкова А.А., Ключарев А.А., Колесникова С.И. Синтез системы управления, мониторинга и оценивания состояния анаэробного биореактора	21
Семенов Е.С., Зувев А.С., Фокин А.П. Расчёт электродинамических характеристик и электронно-волнового взаимодействия в резонаторах гириотронов на основе комплекса программ ANGEL	35
Методы искусственного интеллекта	
Карпенко А.П. Типовые структуры популяционных алгоритмов глобальной оптимизации	48
Виноградов Г.П. Отслеживание мобильных объектов средствами беспроводных сенсорных сетей	58
Цифровые двойники	
Фридман А.Я. Интеграция концептов цифровых двойников и ситуационной осведомленности в системе моделирования критических инфраструктур	70
Массель Л.В., Массель А.Г., Щукин Н.И., Цыбиков А.Р. Разработка цифрового двойника ветровой электростанции: постановка задачи и проектирование	79
Лосев А. С., Массель А.Г. Прогнозирование солнечной радиации и импутация данных для цифрового двойника солнечной электростанции	91
Моделирование и программирование	
Городняя Л.В. Место функционального программирования в организации параллельных вычислений	102
Жарков М.Л., Супруновский А.В. О моделировании железнодорожных узлов на основе теории массового обслуживания	120
Мансурова Ю.Т., Туктарова П.А., Мансуров Т.В. Статистическая и динамическая задачи минимизации инновационного риска в системе управления инновационным потенциалом предприятия	133
Памяти Н.И. Воропая	143

	Content
Editor's Preface	5
Mathematical modeling	
Donskoy I.G. Determining critical conditions in the thermal explosion problem using approximate variational formulations	7
Fomenkova A.A., Klyucharev A.A., Kolesnikova S.I. Synthesis of the system of control, monitoring and evaluation of the state of an anaerobic bioreactor	21
Semenov E.S., Zuev A.S., Fokin A.P. Calculation of electrodynamic characteristics and electron-wave interaction in gyrotron resonators based on the ANGEL software package	35
Artificial intelligence methods	
Karpenko A.P. Typical structures of population algorithms for global optimization	48
Vinogradov G.P. Tracking mobile objects using wireless sensor networks	58
Digital twins	
Fridman A.Y. Incorporation of concepts of digital twins and situational awareness in the modelling system of critical infrastructures	70
Massel L.V., Massel A.G., Shchukin N.I., Tsybikov A.R. Digital twin development of a wind farm: problem statement and construction	79
Losev A.S., Massel A.G. Time series forecasting and data imputation for the digital twin of a solar power plant	91
Modeling and programming	
Gorodnyaya L.V. On functional programming in the organization of parallel computing	102
Zharkov M.L., Suprunovsky A.V. On modeling sections of the railway network based on queuing theory	120
Mansurova Y.T., Tuktarova P.A., Mansurov T.V. Statistical and dynamic problems of minimizing innovation risk in the system of managing the innovative potential of an enterprise	133
In memory of N.I. Voropai	143

ПРЕДИСЛОВИЕ РЕДАКТОРА

Уважаемые коллеги, читатели и авторы журнала!

Благодаря нашим совместным усилиям, с 22 марта 2022 г. журнал "Информационные и математические технологии в науке и управлении" включен в перечень журналов (№ 1224), публикация статей в которых рекомендуется ВАК РФ при защите кандидатских и докторских диссертаций по специальностям:

- 1.2.2. Математическое моделирование, численные методы и комплексы программ (технические науки) (05.13.18),
- 2.3.1. Системный анализ, управление и обработка информации (технические науки) (05.13.01),
- 2.3.3. Автоматизация и управление технологическими процессами и производствами (технические науки) (05.13.06),
- 2.3.5. Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей (технические науки) (05.13.11),
- 2.4.5. Энергетические системы и комплексы (технические науки) (05.14.01)

Несмотря на то, что журнал издается с 2016 г., время от времени нам задают вопрос: что же такое «математические технологии»? Хочу еще раз напомнить, что этот термин был введен академиком Николаем Николаевичем Яненко в конце 70-х гг. прошлого века и получил распространение в 80-х гг. Привожу цитату из его статьи¹: «К настоящему времени при решении на ЭВМ задач математической физики сложилась вполне определенная технологическая цепочка: объект исследования → математическая модель → физическая модель → численный алгоритм → программа → расчет на ЭВМ. Объектом математической технологии здесь является вычислительная часть этой цепочки: математическая модель → численный алгоритм → программа → расчет на ЭВМ, которую в дальнейшем будем называть просто вычислительной цепочкой. Структура ее такова, что допускается цикличность как для всей цепочки в целом, так и между ее отдельными звеньями, т.е. является нелинейной. В силу этого основная задача математической технологии - оптимизация вычислительной цепочки».

Этот вопрос рассматривался также в работах^{2,3}

Ниже приведены названия и хронология мероприятий, в рамках которых обсуждались вопросы математических технологий:

- Научный совет по комплексной проблеме «Математическое моделирование» по руководством академика Самарского А.А.
- Комитет по технологии программирования и инструментальным средствам вычислительного эксперимента – председатель Комитета Легоньков В.И. – единственный лауреат Ленинской премии среди программистов (годы жизни 1933-1997).
- Конференция-семинар «Комплексы программ математической физики» под председательством Н.Н. Яненко (с 1972 г.), 1979 – Днепропетровск, 1980 г. – Горький (VI), 1981 – Песчанка, 1982 – Ташкент.

¹ Яненко Н.Н., Коновалов А.Н. Технологические аспекты численных методов математической физики // Acta Univ. carol. math. Phys. 1974. N 1/2. Pp. 47-53.

² Вопросы модульного анализа и параллельных вычислений в задачах математической физики / Комплексы программ математической физики: (Матер. VI Всесоюз. семинара по комплексам программ мат. физики). Новосибирск, 1980. С. 3-12.

³ Яненко Н.Н. Проблемы вычислительной механики // Очерки, статьи, воспоминания. Новосибирск: Наука, 1988. С. 72-102.

- 1986 – Шушенское, 1990 – Абакан (ак. Шокин Н.И.)
- Сибирская школа по пакетам прикладных программ (ППП) (Матросов В.М., конец 70-х – 1989 г.).
- Имитационный подход в исследованиях энергетики (Шер И.А., с 1979 по 1992 гг.).
- Информационные технологии в энергетике, сейчас Байкальская Всероссийская конференция "Информационные и математические технологии в науке и управлении" (Массель Л.В., с 1993 г. по настоящее время; в 1993-2010 гг. соруководителем конференции был профессор МФТИ Лев Николаевич Столяров (годы жизни 1937-2010), по чьей инициативе в названии конференции и появился термин «математические технологии».).

В нашем журнале история этого термина затрагивалась в статье А.К. Черкашина (ИМТ, №1, 2021. С. 17-35).

Сейчас этот термин используется, например, в таких ВУЗах, как МФТИ, ИТМО, МЭИ. Так, в МФТИ существует магистерская программа "Математические и информационные технологии", которая готовит высококвалифицированных специалистов в области информационных технологий. Основными направлениями специализации являются подготовка специалистов в области параллельных вычислений на высокопроизводительных многопроцессорных системах и решения на таких системах задач вычислительной физики (медицина, сейсмология, моделирование техногенных катастроф, высокоскоростное взаимодействие тел и др.), также, в области агентных технологий, теоретической и прикладной робототехники.

Ведущие преподаватели специализации являются представителями научной школы академика Г. С. Поспелова. Руководитель специализации профессор А. И. Эрлих под руководством Г. С. Поспелова в середине 1970-х годов одним из первых в России начал развивать и применять подходы нового по тем временам научного направления «Искусственный интеллект» в разработках и проектировании прикладных компьютерных систем. Сегодня методы и средства искусственного интеллекта, как основа новых информационных технологий, прочно заняли передовые позиции в теоретической и прикладной информатике.

Подготовка студентов ориентирована на их дальнейшую работу по развитию новых методов и средств создания сложных интеллектуальных систем различного прикладного назначения. Образовательный процесс осуществляется на кафедре Интеллектуальных систем, которую до последнего времени возглавлял академик РАН Константин Владимирович Рудаков.

В заключение хочу добавить, что сейчас задача оптимизации вычислительной цепочки «математическая модель → численный алгоритм → программа → расчет на ЭВМ» решается с применением методов искусственного интеллекта, статьи по этой тематике регулярно публикуются в нашем журнале.

Желаем успехов всем нашим авторам и надеемся на дальнейшее сотрудничество!

Главный редактор журнала "Информационные и математические технологии в науке и управлении" Л.В. Массель

Математическое моделирование

УДК 544.45; 517.9

doi: 10.38028/ESI.2022.25.1.001

Определение критических условий в задаче о тепловом взрыве с помощью приближенных вариационных формулировок

Донской Игорь Геннадьевич

Институт систем энергетики им. Л.А. Мелентьева СО РАН,
Россия, Иркутск, *donskoy.chem@mail.ru*

Аннотация. В работе предлагаются приближенные вариационные принципы для классической задачи теплового взрыва и ее модификаций (в том числе, для случаев, в которых не выполняется приближение высокой энергии активации). Вариационные задачи решаются с использованием однопараметрической тестовой функции и метода Ритца (для полиномиальной аппроксимации решения). Результаты расчетов позволяют определить зависимость критических условий от параметров задачи: интенсивности теплообмена на границе, конвективного переноса, вязкой диссипации.

Ключевые слова: тепловой взрыв, обратная вариационная задача, критические условия, численные методы

Цитирование: Донской И.Г. Определение критических условий в задаче о тепловом взрыве с помощью приближенных вариационных формулировок // Информационные и математические технологии в науке и управлении. – 2022. – № 1 (25). – С. 7-20. – DOI:10.38028/ESI.2022.25.1.001.

Введение. Типичная обратная задача вариационного исчисления – это определение вида функции Лагранжа, для которой интеграл

$$I = \int_0^x L(\xi, y, y', \dots) d\xi \quad (1)$$

имеет экстремальное значение, если функция $y(\xi)$ удовлетворяет нужным условиям (например, некоторому дифференциальному уравнению). Вариационная постановка может быть предпочтительной, например, из соображений простоты численного решения [1], или возможностей подробного анализа взаимосвязей между описываемыми процессами [2].

Уравнение стационарной теории теплового взрыва выглядит следующим образом [3]:

$$\theta'' + Fk \exp\left(\frac{\theta}{1 + Ar\theta}\right) = 0 \quad (2)$$

Здесь θ – это безразмерная температура, Fk – число Франк-Каменецкого (отношение скоростей тепловыделения при химической реакции и теплоотвода путем теплопроводности), Ar – число Аррениуса (величина $\exp(1/Ar)$ это отношение скоростей реакции при максимальной и минимальной температуре). При малых Ar решение этого уравнения существует лишь для ограниченных значений Fk : при достижении его критического значения происходит разрушение решения, которое трактуется как тепловой взрыв. Уравнению (2) и его модификациям посвящена обширная литература (см., например, [4-11]).

Функция Лагранжа для задачи о тепловом взрыве в пределе большой энергии активации ($Ar = 0$) имеет простой вид [12]:

$$L_0 = \frac{1}{2}(\theta')^2 - Fk \exp \theta \quad (3)$$

При Ar порядка 0,23 происходит вырождение теплового взрыва: решение гладко зависит от значения Fk [13, 14]. Однако даже при малых Ar условия стационарности интеграла от функции (3) не удовлетворяют в точности уравнению (2). В настоящей работе предлагаются способы построения приближенных функций Лагранжа, которые позволяют определять условия существования решения для уравнения (2) и его модификаций.

Функция Лагранжа, интеграл от которой будет иметь экстремум для решения уравнения (2), запишется следующим образом:

$$L_{Ar} = \frac{1}{2}(\theta')^2 - Fk \int_0^\theta \exp\left(\frac{y}{1+Ar y}\right) dy \quad (4)$$

Интеграл во втором слагаемом не выражается в элементарных функциях, однако может быть приближен, например, подходящим полиномом, или задан таблично. Для цилиндрической и сферической симметрии можно записать функцию Лагранжа в виде:

$$L_{Ar} = \frac{\xi^m}{2}(\theta')^2 - Fk \xi^m \int_0^\theta \exp\left(\frac{y}{1+Ar y}\right) dy,$$

где m равно 1 или 2.

Классическая задача о тепловом взрыве исследовалась с помощью вариационных методов в работах [12, 15, 16], в т.ч. для сложной кинетики химических реакций в [17, 18]. Вариационные задачи для определения критических условий теплового взрыва при течении жидкостей с разными вязкостными свойствами исследовались в работах [19, 20]. Вариационный принцип для смешанного лучистого и кондуктивного теплообмена предложен в работе [21]. Вариационные методы применялись для построения аппроксимаций и численного решения уравнений теории теплового взрыва в [22, 23]. При этом обычно принимается приближение высокой энергии активации ($Ar = 0$), а конвективный перенос не учитывается в вариационной постановке.

В настоящей работе рассматриваются приближенные функции Лагранжа для уравнения (2), учитывающие зависимость скорости реакции от параметра Ar , на основе методов возмущений [24, 25]. Предложена вариационная формулировка для уравнения теплопереноса в стационарном одномерном проточном реакторе. Результаты решения вариационных задач сравниваются с результатами численного решения дифференциальных уравнений типа (2), которое проводилось итерационным методом с полунявной разностной аппроксимацией [26]. Критические значения параметра Fk в разных вариантах находятся методом половинного деления с точностью до 10^{-6} .

1. Влияние скорости реакции при начальной температуре. Учтем зависимость скорости реакции от параметра Ar . В самом грубом приближении рассмотрим функцию Лагранжа:

$$L_A^{(0)} = \frac{1}{2}(\theta')^2 - Fk \exp\left(\frac{\theta}{1+Ar\theta}\right) \quad (5)$$

Условие стационарности приводит к уравнению:

$$\theta'' + \frac{Fk}{(1+Ar\theta)^2} \exp\left(\frac{\theta}{1+Ar\theta}\right) = 0$$

Видно, что при $Ar = 0$ полученное уравнение совпадает с уравнением (2). Для того, чтобы избавиться от знаменателя, умножим второе слагаемое в выражении (5) на $(1+Ar\theta)^2$ и рассмотрим следующее приближение для функции Лагранжа:

$$L_A^{(1)} = \frac{1}{2}(\theta')^2 - Fk(1 + Ar\theta)^2 \exp\left(\frac{\theta}{1 + Ar\theta}\right) \quad (6)$$

Условия стационарности дают на этот раз уравнение:

$$\theta'' + Fk(1 + 2Ar + 2Ar^2\theta) \exp\left(\frac{\theta}{1 + Ar\theta}\right) = 0$$

Таким образом, в следующем приближении надо поделить второе слагаемое (6) на полученный полином:

$$L_A^{(2)} = \frac{1}{2}(\theta')^2 - \frac{Fk(1 + Ar\theta)^2}{1 + 2Ar(1 + Ar\theta)} \exp\left(\frac{\theta}{1 + Ar\theta}\right)$$

Следующее уточнение даст:

$$L_A^{(3)} = \frac{1}{2}(\theta')^2 - \frac{Fk(1 + Ar\theta)^2}{1 + 2Ar(1 + Ar\theta) - \frac{Ar^2(1 + Ar\theta)^2}{1 + 2Ar(1 + Ar\theta)}} \exp\left(\frac{\theta}{1 + Ar\theta}\right) \quad (7)$$

Другой способ построения приближенных функций Лагранжа – это разложение экспоненциального множителя в ряд по Ar . Показатель экспоненты можно представить в виде суммы геометрической прогрессии:

$$\frac{\theta}{1 + Ar\theta} = \theta \left[1 - Ar\theta + (Ar\theta)^2 - (Ar\theta)^3 + \dots \right]$$

Первые члены ряда дают приближение:

$$\exp\left(\frac{\theta}{1 + Ar\theta}\right) \approx \exp\theta \left[1 - Ar\theta^2 + Ar^2\theta^3 + \frac{Ar^2}{2}\theta^4 - Ar^3\theta^4 - Ar^3\theta^5 - \frac{Ar^3}{6}\theta^6 + \dots \right]$$

Обрывая ряд на какой-либо степени Ar , можно получать приближения разного порядка. Тогда после интегрирования этого выражения можно получить приближенную функцию Лагранжа в виде:

$$L_B = \frac{1}{2}(\theta')^2 - Fk \exp\theta p(\theta), \quad (8)$$

где $p(\theta)$ – это полином от температуры:

$$p(\theta) = 1 - Ar \sum_{i=0}^2 (-1)^i \theta^{2-i} + Ar^2 \sum_{i=0}^3 (-1)^i \theta^{3-i} + \left(\frac{Ar^2}{2} - Ar^3 \right) \sum_{i=0}^4 (-1)^i \theta^{4-i} - Ar^3 \sum_{i=0}^5 (-1)^i \theta^{5-i} - \frac{Ar^3}{6} \sum_{i=0}^6 (-1)^i \theta^{6-i}$$

Аналогично могут быть получены приближенные функции Лагранжа с полиномом p , зависящим от высших степеней Ar . Эти выкладки, однако, очень громоздки, поэтому здесь не приводятся. Ниже в расчетах применяются приближения для функции Лагранжа (4) в виде (7) и (8).

Если разложить показатель степени до второго члена, можно записать:

$$\begin{aligned} L_C &= \frac{1}{2}(\theta')^2 - Fk \int_0^\theta \exp(y - Ary^2) dy = \\ &= \frac{1}{2}(\theta')^2 - \frac{Fk}{2} \sqrt{\frac{\pi}{Ar}} \exp\left(\frac{1}{4Ar}\right) \left[\operatorname{erf}\left(\sqrt{Ar}\theta - \frac{1}{2\sqrt{Ar}}\right) - \operatorname{erf}\left(-\frac{1}{2\sqrt{Ar}}\right) \right] \end{aligned} \quad (9)$$

Наконец, путем подбора можно предложить простую аппроксимацию для (4) в виде:

$$L_D = \frac{1}{2}(\theta') - Fk \left(1 + \frac{Ar}{2} \theta^2\right) \left[\exp\left(\frac{\theta}{1 + Ar\theta}\right) - 1 \right] \quad (10)$$

Все полученные здесь функции Лагранжа (7-10) являются приближенными: точное уравнение они дают только при $Ar = 0$. Можно, однако, оценить точность определения критического значения Fk с помощью предложенных приближенных вариационных принципов.

Вычисления проводятся для двухпараметрического семейства парабол. На левой границе примем граничное условие второго рода:

$$\theta'(0) = 0$$

На правой границе, в общем случае, можно записать условие третьего рода:

$$\theta'(1) + Bi\theta(1) = 0$$

Здесь Bi – число Био, равное отношению интенсивностей теплообмена с внешней средой и теплопроводности. Тогда условия на коэффициенты полинома второй степени дают уравнение для тестовой функции:

$$\theta = a \left(1 + \frac{2}{Bi} - \xi^2\right) \quad (11)$$

В частном случае, когда $Bi \gg 1$, получаем однопараметрическую тестовую функцию, которая использовалась ранее в работах [15-17]:

$$\theta = a(1 - \xi^2)$$

Для определения критического значения Fk определяется зависимость вариационного функционала от параметра a : тепловому воспламенению соответствует значение Fk , при котором эта зависимость теряет локальный минимум и становится монотонно убывающей. Для более точной локализации определяются нули производной dI/da : в дозрывных условиях существуют два нуля dI/da , которые соответствуют локальному минимуму и локальному максимуму; критическим условиям соответствует точка, в которой $d^2I/da^2 = 0$. По этому условию можно искать критическое значение Fk . На сетке значений a (с шагом 10^{-3}) при постоянном Fk для выбранной функций Лагранжа рассчитываются значения I (методом трапеций). Число Fk считается докритическим, если существует такое значение $I(a_i)$, что выполняется условие:

$$I(a_{i-1}) \geq I(a_i) \leq I(a_{i+1})$$

Соответствующее значение a_i пропорционально стационарному значению максимальной температуры $\theta(0)$. Алгоритм строится следующим образом: задаются два значения Fk , заведомо докритическое ($Fk = 0$) и заведомо закритическое (произвольное большое число, например, 10). Интервал делится пополам, для нового значения Fk^* проводится проверка на существование стационарного значения I . Если такое значение существует, то Fk^* становится нижней границей диапазона; если значения не существует, то Fk^* становится верхней границей диапазона. Итерации ведутся до тех пор, пока ширина диапазона не становится меньше заданной величины (10^{-6}).

Ограничения на точность определения критического значения Fk связаны с точностью приближения температурного распределения тестовой функцией, а также шагом по переменной ξ (поскольку значения I находятся численным интегрированием). В наших расчетах шаг по переменной ξ и a равен 10^{-3} (максимальные значения 1 и 3, соответственно). В связи с погрешностями вычислений, критическое значение Fk при $Ar = 0$ не совпадает в точности с аналитическим результатом (0,878). Это отличие, однако, невелико. На рис. 1 показана зависимость критического значения Fk от числа Ar : видно, что функция Лагранжа L_A лучше аппроксимирует поведение системы во всем диапазоне. Для сравнения приведены результаты

численного решения уравнения (2) и данные работы [27]. Как видно из табл. 1, для L_A также наблюдаются лучшая сходимость с увеличением порядка приближения (здесь Δ – отклонения от значения 0,98976..., полученного при численном решении дифференциального уравнения (2)). Приближение температурного распределения ветвью параболы для всех случаев оказывается достаточно хорошим. Приближение L_C приводит к вырождению теплового взрыва уже при Ar порядка 0,12. Остальные приближения для функции Лагранжа, наоборот, расширяют диапазон существования критических условий за пределы теоретических границ.

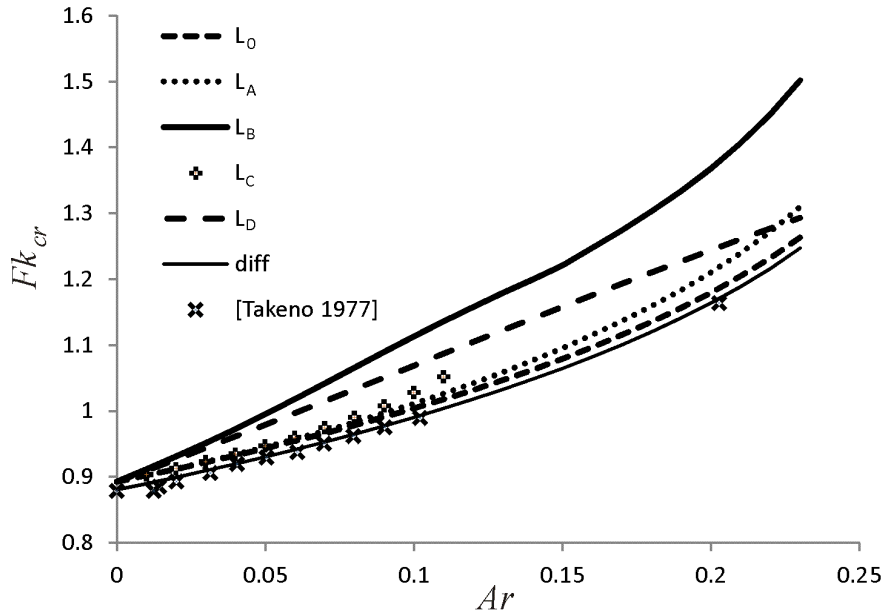


Рис. 1. Зависимость критического значения параметра Fk для задачи (1) от параметра Ar ($Bi = 0$): L_0 – (4); L_A – (7); L_B – (8); L_C – (9); L_D – (10); $diff$ – численное решение дифференциального уравнения (2).

Таблица 1. Критические условия в задаче (4) при использовании приближенных функций Лагранжа для $Ar = 0,1$.

Номер приближения	L_A (7)		L_B (8)		L_C (9)		L_D (10)	
	Fk_{cr}	$\Delta Fk, \%$	Fk_{cr}	$\Delta Fk, \%$	Fk_{cr}	$\Delta Fk, \%$	Fk_{cr}	$\Delta Fk, \%$
0	1,334	34,8	0,892	-9,8	-	-	-	-
1	0,819	-17,3	1,458	47,3	1,028	3,83	1,068	7,96
2	1,021	3,1	1,100	11,1	-	-	-	-
3	1,011	2,1	1,167	17,9	-	-	-	-

2. Влияние теплообмена с окружающей средой. При $Ar = 0$ значение I для тестовой функции (11) можно найти по формуле:

$$I = \frac{2}{3}a^2 - \frac{Fk}{2}\exp(a)\exp\left(\frac{2a}{Bi}\right)\sqrt{\frac{\pi}{a}}\operatorname{erf}\left(\sqrt{a}\right) \quad (12)$$

Критическое условие (вырождение экстремумов по параметру a) для этого выражения в пределе больших значений Bi достигается при $Fk = 0,89227...$ (отличие от точного значения менее 2%). Зависимость критического Fk от Bi воспроизводится лишь качественно: видимо, приближение параболы становится неприменимым при уменьшении Bi до величины порядка нескольких десятков. С уменьшением Bi второе слагаемое в правой части (9) становится меньше по величине, поэтому переход к тепловому взрыву происходит при меньших значе-

ниях Fk . Если считать, что критическое значение безразмерной температуры a примерно равно единице, можно оценить критическое значение Fk следующим образом:

$$Fk \approx Fk_0 \exp\left(-\frac{2}{Bi}\right) \quad (13)$$

Эта зависимость сравнивается с оценками на основе выражения (12) и результатами численного решения уравнения (2). До значений Bi порядка единицы приближение (13) дает лучшее соответствие численному решению, чем поиск стационарных условий для выражения (12). Интересно, что критическое значение Fk , полученное при определении точки перегиба для (12), примерно равно квадрату истинного значения.

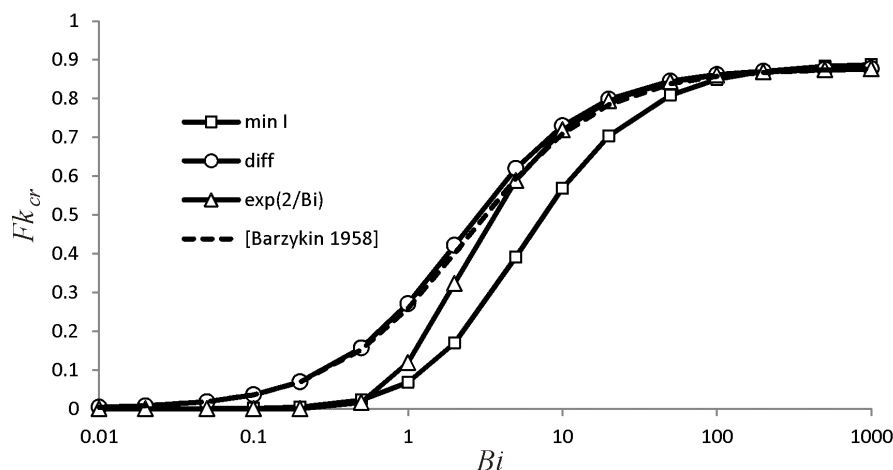


Рис. 2. Зависимости критического значения Fk от параметра Bi , полученными разными способами: квадраты – минимизация выражения (12); круги – численное решение уравнения (2); треугольники – оценка по формуле (13); пунктир – приближенная формула из работы [28].

3. Влияние конвективного переноса. Рассмотрим уравнение, описывающее докритический стационарный режим теплообмена в одномерном реакторе

$$-Pe\theta' - Bi_w\theta + \theta'' + Fk \exp\left(\frac{\theta}{1 + Ar\theta}\right) = 0 \quad (14)$$

Здесь Pe – число Пекле (отношение интенсивности конвективного переноса к кондуктивному); Bi_w – число Био для теплообмена реакционной смеси со стенкой реактора. Функция Лагранжа для уравнения (14) запишется в виде:

$$L = \exp(-Pe\xi) \left[\frac{1}{2}(\theta')^2 + \frac{Bi_w}{2}\theta^2 - Fk \int_0^\theta \exp\left(\frac{y}{1 + Ary}\right) dy \right] \quad (15)$$

Подобный вид функции Лагранжа (с экспоненциальным множителем) использовался ранее для других задач тепломассообмена в работах [29, 30] и для механических систем в [31]. Примем следующие граничные условия для уравнения (14):

$$\theta(0) = 0, \quad \theta'(1) = 0$$

Использование тестовой функции (11) оказывается слишком грубым приближением уже при небольших значениях Pe . Наличие конвективного слагаемого в (14) приводит к выполаживанию начального участка зависимости $\theta(\xi)$ и росту градиентов при приближении к $\xi = 1$. Для вычислений используется метод Рунге с полиномиальной аппроксимацией:

$$\theta(\xi) \approx \sum_{i=0}^N a_i \xi^i$$

Условия стационарности дают итерационную схему для уточнения коэффициентов полинома:

$$Da^q = s^{q-1} \quad (16)$$

Здесь q – номер итерации; элементы D и s определяются из выражений:

$$D_{ki} = ik \int_0^1 \exp(-Pe\xi) \xi^{i+k-2} d\xi + Bi_w \int_0^1 \exp(-Pe\xi) \xi^{i+k} d\xi,$$

$$s_k = Fk \int_0^1 \exp\left(\frac{\bar{\theta}}{1 + Ar\bar{\theta}} - Pe\xi\right) \xi^k d\xi$$

Здесь функция θ в уравнении для s_k считается фиксированной (в данном случае нет необходимости приближений для источника типа (7-10)). Граничные условия кроме очевидного равенства $a_0 = 0$ дают дополнительное соотношение между коэффициентами:

$$\sum_{i=1}^N ia_i = 0$$

Итерационный процесс заканчивается, когда выполняется условие:

$$\max |\theta^q - \theta^{q-1}| < \varepsilon,$$

где допустимая погрешность ε принята равной 0,001.

В работе [32] показано, что в задаче с конвективным переносом критическое значение Fk асимптотически стремится к Pe (т.е. при больших значениях Pe максимальное значение безразмерной температуры стремится к $-\ln(1 - Fk/Pe)$). На рис. 3 показаны результаты решения вариационной задачи для полинома степени $N = 6$ ($Ar = 0$): видно, что с увеличением Pe аппроксимация становится все хуже, однако при $Pe = 10$ полином шестой степени дает точность оценки критического значения Fk около 2,5%.

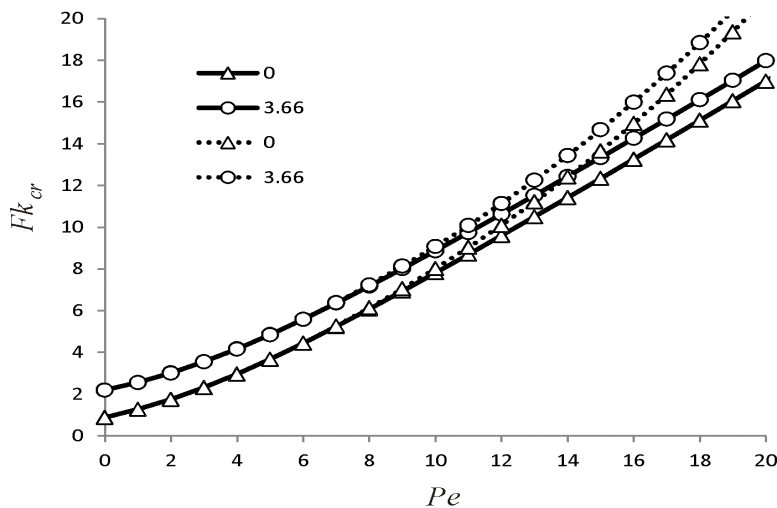


Рис. 3. Зависимость критического значения Fk от Pe : сплошные линии – численное решение уравнения (14); пунктирные линии – решение вариационной задачи с функцией Лагранжа (15) методом Ритца; треугольники – $Bi = 0$; круги – $Bi = 3,66$.

Теплопотери увеличивают критическое значение Fk : зависимости для $Bi = 0$ и $Bi = 5$ в выбранном диапазоне отличаются примерно друг от друга примерно постоянным сдвигом (при больших Bi для более точной оценки критических условий надо делать поправку на конверсию реагентов [33], в настоящей работе этот эффект не учитывается). Интересно, что с ростом Pe усиливаются различия между кривыми для разных Ar : можно сказать, что увели-

чение интенсивности конвективного переноса приводит к росту чувствительности условий зажигания к скорости реакции при начальной температуре.

Рассмотренные в предыдущих разделах постановки естественным образом распространяются на двумерный и трехмерный случай. Наличие конвективного переноса не позволяет совершить такой же простой переход к многомерным обобщениям: зависимость Pe от пространственных переменных приводит к появлению дополнительных ненужных слагаемых в уравнениях, выражающих условия стационарности в вариационной задаче. В некоторых случаях, как показано в [19, 20, 34], можно редуцировать двумерную задачу к одномерной и рассматривать только перенос в направлении, поперечному потоку:

$$\theta'' + \alpha \xi^2 + Fk \exp\left(\frac{\theta}{1 + Ar\theta}\right) = 0 \quad (17)$$

Здесь α – это интенсивность вязкой диссипации. Можно записать соответствующую уравнению (17) функцию Лагранжа для вариационной задачи в виде:

$$L = \frac{1}{2}(\theta')^2 - \alpha \theta \xi^2 - Fk \int_0^\theta \exp\left(\frac{y}{1 + Ary}\right) dy \quad (18)$$

Диссипация приводит к дополнительному тепловыделению, поэтому критическое число Fk должно уменьшаться с ростом параметра α . Автор [34] дает простую формулу для оценки критического значения Fk для уравнения (18) при $Ar = 0$:

$Fk(\alpha)/Fk(0) = \exp(-\alpha/12)$. Расчеты проводились с применением тестовой функции (11) и приближения типа (7) для источникового слагаемого. Сравнение численных результатов с этой формулой показано на рис. 5. Видно, что с ростом Ar и α эта формула становится все менее точной (хотя при $\alpha < 2$ погрешность составляет менее 5%).

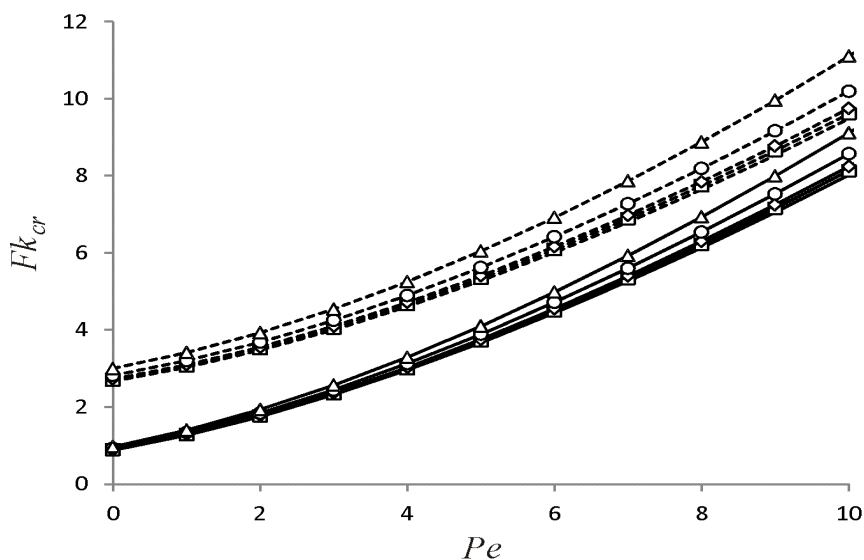


Рис. 4. Зависимость критического значения Fk от Pe при $Bi = 0$ (сплошные линии) и $Bi = 5$ (пунктирные линии); квадраты: $Ar = 0,01$; ромбы: $Ar = 0,02$; круги: $Ar = 0,05$; треугольники: $Ar = 0,1$.

Заключение. В статье предложены приближенные функции Лагранжа для вариационной формулировки задачи о тепловом взрыве в одномерном реакторе. С помощью простых тестовых функций сопоставлена точность разных приближений (в первую очередь, приближения высокой энергии активации химической реакции) и определены границы их применимости. Предложена вариационная формулировка задачи о тепловой устойчивости

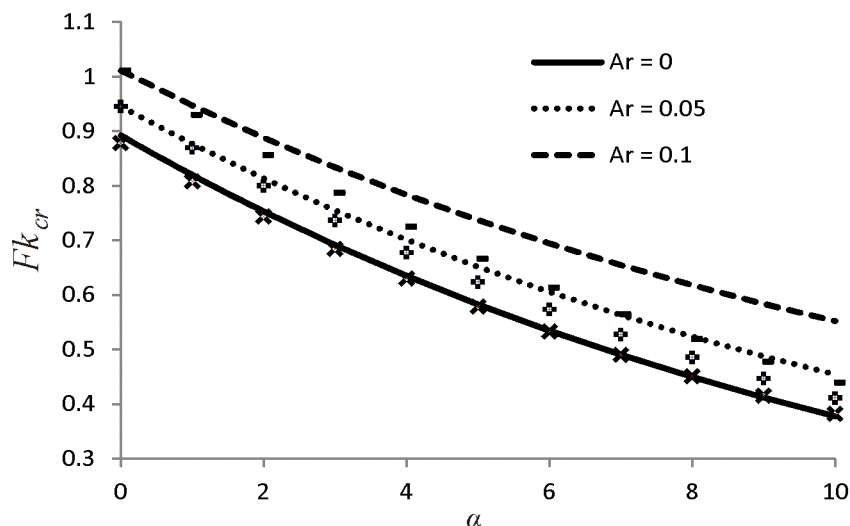


Рис. 5. Зависимость критического значения Fk в для уравнения (17) от интенсивности вязкой диссипации: сплошные линии – решение вариационной задачи (18); маркеры – оценка по формуле из работы [34].

проточного химического реактора, с помощью метода Ритца определены критические условия при наличии теплопотерь и конвективного переноса. Упрощенно рассмотрено влияние вязкой диссипации. Полученные результаты могут быть применены для оценки тепловой устойчивости реагирующих сред. Дальнейшая работа будет связана с исследованием вариационных задач, соответствующих квазистационарному распространению волн горения в реагирующих средах.

Благодарности: Работа выполнена в рамках проекта государственного задания (№ FWEU-2021-0005) программы фундаментальных исследований РФ на 2021-2030 гг.

СПИСОК ИСТОЧНИКОВ

1. Шехтер Р. Вариационный метод в инженерных расчетах / Р. Шехтер, – Мир, 1971. – 292 с.
2. Гленсдорф П., Пригожин И. Термодинамическая теория структуры, устойчивости и флуктуаций / П. Гленсдорф, И. Пригожин. – М.: Мир, 1973. – 280 с.
3. Франк-Каменецкий Д.А. Диффузия и теплопередача в химической кинетике / Д.А. Франк-Каменецкий. – М.: Наука, 1987. – 502 с.
4. Fradkin L.J., Wake G.C. The Critical Explosion Parameter in the Theory of Thermal Ignition. IMA Journal of Applied Mathematics, 1977, vol. 20, pp. 471–484, DOI: 10.1093/imamat/20.4.471.
5. Wake G.C., Sleeman M., Chen X.D., Jones J.C. Theory and applications of ignition with variable activation energy. Journal of Thermal Science, 1992, vol. 1, pp. 208–212, DOI: 10.1007/BF02663700.
6. Li S., Liao S.-J. An analytic approach to solve multiple solutions of a strongly nonlinear problem. Applied Mathematics and Computation, 2005, vol. 169, pp. 854–865, DOI: 10.1016/j.amc.2004.09.066.
7. Gordon P.V., Ko E., Shivaji R. Multiplicity and uniqueness of positive solutions for elliptic equations with nonlinear boundary conditions arising in a theory of thermal explosion. Non-linear Analysis: Real World Applications, 2014, vol. 15, pp. 51–57, DOI: 10.1016/j.nonrwa.2013.05.005.

8. Ko E., Prashanth S. Positive solutions for elliptic equations in two dimensions arising in a theory of thermal explosion. *Taiwanese Journal of Mathematics*, 2015, vol. 19, pp. 1759-1775, DOI: 10.11650/tjm.19.2015.5968.
9. Yu Y., Luo Q., Liang D. Indirect Method of the Critical Parameters of Frank-Kamenetskii Equations in Spontaneous Combustion Theory. *Procedia Engineering*, 2016, vol. 135, pp. 551-554, DOI: 10.1016/j.proeng.2016.01.099.
10. Roul P. A fast and accurate computational technique for efficient numerical solution of non-linear singular boundary value problems. *International Journal of Computer Mathematics*, 2019, Vol. 96, pp. 51-72, DOI: 10.1080/00207160.2017.1417588.
11. Yu S.-Y., Yan B. Positive Solutions for a Singular Elliptic Equation Arising in a Theory of Thermal Explosion. *Mathematics*, 2021, vol. 9, pp. 2173, DOI: 10.3390/math9172173.
12. Гришин А.М. Некоторые задачи теории воспламенения / А.М. Гришин // Прикладная механика и техническая физика, 1962. – № 5. – С. 75-79.
13. Мержанов А.Г., Зеликман Е.Г., Абрамов В.Г. Вырожденные режимы теплового взрыва / А.Г. Мержанов, Е.Г. Зеликман, В.Г. Абрамов // Докл. АН СССР, 1968. – Т. 180. – № 3. – С. 639-642.
14. Boddington T., Gray P., Robinson C. Thermal explosions and the disappearance of criticality at small activation energies: exact results for the slab. *Proceedings of the Royal Society A. Mathematical, Physical and Engineering Sciences*, 1979, vol. 368, pp. 441–461, DOI: 10.1098/rspa.1979.0140.
15. Анисимов С.И., Виткин Э.И. Некоторые вариационные задачи теории теплового взрыва / С.И. Анисимов, Э.И. Виткин // Прикладная механика и техническая физика, 1966. – № 4. – С. 150-151.
16. Зарубин В.С., Кувыркин Г.Н., Савельева И.Ю. Вариационная форма модели теплового взрыва в твердом теле с зависящей от температуры теплопроводностью / В.С.Зарубин, Г.Н. Кувыркин, И.Ю. Савельева // Теплофизика высоких температур, 2018. –Т. 56. – № 2. – С. 235-240. DOI: 10.7868/ S0040364418020102.
17. Graham-Eagle J.G., Wake G.C. Theory of thermal explosions with simultaneous parallel reactions. II. The two- and three-dimensional cases and the variational method. *Proceedings of the Royal Society A. Mathematical, Physical and Engineering Sciences*, 1985, vol. 401, pp. 195-202, DOI: 10.1098/rspa.1985.0094.
18. Донской И.Г. Численная оценка критических условий в задаче о тепловом взрыве с флуктуациями реакционной способности / И.Г. Донской // Информационные и математические технологии в науке и управлении, 2021. – № 1. – С. 54-65. DOI: 10.38028/ESI.2021.21.1.005.
19. Ajadi S.O. A note on the thermal stability of a reactive non-Newtonian flow in a cylindrical pipe. *International Communications in Heat and Mass Transfer*, 2009, vol. 36, pp. 63-68, DOI: 10.1016/j.icheatmasstransfer.2008.09.005.
20. Ajadi S.O. The influence of viscous heating and wall thermal conditions on the thermal ignition of a Poiseuille/Couette reactive flow // *Chemical physics*, 2010, v. 29, no. 8, p. 47-54.
21. Blouquin R., Joulin G. On a Variational Principle for Reaction/Radiation/Conduction Equilibria. *Combustion Science and Technology*, 1996, vol. 112, pp. 375-385, DOI: 10.1080/00102209608951968.

22. Moise A., Pritchard H.O. Newton-variational solution of the Frank–Kamenetskii thermal explosion problem. *Canadian Journal of Chemistry*, 1989, vol. 67, p. 442, DOI: 10.1139/v89-069.
23. Wazwaz A.-M. Solving the non-isothermal reaction-diffusion model equations in a spherical catalyst by the variational iteration method. *Chemical Physics Letters*, 2017, vol. 679, pp. 132-136, DOI: 10.1016/j.cplett.2017.04.077.
24. Байков В.А., Газизов Р.К., Ибрагимов Н.Х. Приближенные симметрии / В.А. Байков, Р.К. Газизов, Н.Х.Ибрагимов // Математический сборник, 1988. – Т. 136(178). – № 4(8). С. 435–450.
25. Gorgone M., Oliveri F. Approximate Noether Symmetries of Perturbed Lagrangians and Approximate Conservation Laws. *Mathematics*, 2021, vol. 9, pp. 2900, DOI: 10.3390/math9222900.
26. Donskoy I.G. Numerical estimation of thermal ignition conditions for reactive medium with Gaussian distribution of activation energy. *Journal of Physics: Conference Series*, 2021, vol. 2119, p. 012102, DOI:10.1088/1742-6596/2119/1/012102.
27. Takeno T. Ignition criterion by thermal explosion theory. *Combustion and Flame*, 1977, vol. 29, pp. 209-211, DOI: 10.1016/0010-2180(77)90108-0.
28. Барзыкин В.В., Мержанов А.Г. Краевая задача в теории теплового взрыва / В.В. Барзыкин, А.Г. Мержанов // Доклад АН СССР, 1958. –Т. 120. – № 6. – С. 1271-1273.
29. Гостинцев Ю.А. Метод приведения к обыкновенным дифференциальным уравнениям в задачах нестационарного горения пороха / Ю.А. Гостинцев // Физика горения и взрыва, 1967. – № 3. – С. 335-361.
30. Sieniutycz S. The variational principles of classical type for non-coupled non-stationary irreversible transport processes with convective motion and relaxation. *International Journal of Heat and Mass Transfer*, 1977, vol. 20, pp. 1221-1231, DOI: 10.1016/0017-9310(77)90131-4.
31. He J.-H. Variational principles for some nonlinear partial differential equations with variable coefficients. *Chaos, Solitons & Fractals*, 2004, vol. 19, pp. 847-851, DOI: 10.1016/S0960-0779(03)00265-0.
32. Дик И.Г., Толстых А.В. Воспламенение продуваемого слоя / И.Г. Дик, А.В. Толстых А.В. // Физика горения и взрыва, 1994. –№ 2. – С. 3-7.
33. Boddington T., Gray P., Wake G.C. Criteria for thermal explosions with and without reactant consumption. *Proceedings of the Royal Society A. Mathematical, Physical and Engineering Sciences*, 1977, vol. 357, pp. 403–422, DOI: 10.1098/rspa.1977.0176.
34. Adler J. Thermal explosion theory for reactive flow between parallel heated walls. *Combustion and Flame*, 1975, vol. 24, pp. 151-158, DOI: 10.1016/0010-2180(75)90142-X.

Донской Игорь Геннадьевич, к.т.н., с.н.с. лаборатории термодинамики, Институт систем энергетики им. Л.А. Мелентьева СО РАН, donskey.chem@mail.ru, Россия, г. Иркутск, ул. Лермонтова 130

Determining critical conditions in the thermal explosion problem using approximate variational formulations

Igor G. Donskoy

Melentiev Energy Systems Institute Siberian Branch of the Russian Academy of Sciences, Russia, Irkutsk, *donskoy.chem@mail.ru*

Abstract. The paper proposes approximate variational principles for the classical thermal explosion problem and its modifications (including cases where the high activation energy approximation does not hold). Variational problems are solved using a one-parameter test function and the Ritz method (for polynomial approximation of the solution). The results of the calculations make it possible to determine the dependence of the critical conditions on the parameters of the problem: the intensity of heat transfer at the inner boundary, forced convection, and viscous dissipation.

Keywords: thermal explosion, inverse variational problem, critical conditions, numerical methods

Acknowledgements: The work was carried out within the framework of the draft state task (No. FWEU-2021-0005) of the program of fundamental research of the Russian Federation for 2021-2030.

REFERENCES

1. Schechter R. Variatsionnyi metod v inzhenernykh raschetakh [Variational method in engineering calculations]. Moscow, Mir, 1971, 292 p.
2. Glansdorff P., Prigogine I. Termodinamicheskaya teoriya struktury, ustoychivosti i fluktuatsii [Thermodynamic theory of structure, stability and fluctuations]. Moscow, Mir, 1973, 280 p.
3. Frank-Kamenetskii D.A. Diffuziya i teploperedacha v khimicheskoi kinetike [Diffusion and Heat Exchange in Chemical Kinetics]. Moscow, Nauka, 1987, 502 p.
4. Fradkin L.J., Wake G.C. The Critical Explosion Parameter in the Theory of Thermal Ignition. IMA Journal of Applied Mathematics, 1977, vol. 20, pp. 471–484, DOI: 10.1093/imamat/20.4.471.
5. Wake G.C., Sleeman M., Chen X.D., Jones J.C. Theory and applications of ignition with variable activation energy. Journal of Thermal Science, 1992, vol. 1, pp. 208–212, DOI: 10.1007/BF02663700.
6. Li S., Liao S.-J. An analytic approach to solve multiple solutions of a strongly nonlinear problem. Applied Mathematics and Computation, 2005, vol. 169, pp. 854–865, DOI: 10.1016/j.amc.2004.09.066.
7. Gordon P.V., Ko E., Shivaji R. Multiplicity and uniqueness of positive solutions for elliptic equations with nonlinear boundary conditions arising in a theory of thermal explosion. Nonlinear Analysis: Real World Applications, 2014, vol. 15, pp. 51–57, DOI: 10.1016/j.nonrwa.2013.05.005.
8. Ko E., Prashanth S. Positive solutions for elliptic equations in two dimensions arising in a theory of thermal explosion. Taiwanese Journal of Mathematics, 2015, vol. 19, pp. 1759–1775, DOI: 10.11650/tjm.19.2015.5968.

9. Yu Y., Luo Q., Liang D. Indirect Method of the Critical Parameters of Frank-Kamenetskii Equations in Spontaneous Combustion Theory. *Procedia Engineering*, 2016, vol. 135, pp. 551-554, DOI: 10.1016/j.proeng.2016.01.099.
10. Roul P. A fast and accurate computational technique for efficient numerical solution of non-linear singular boundary value problems. *International Journal of Computer Mathematics*, 2019, vol. 96, pp. 51-72, DOI: 10.1080/00207160.2017.1417588.
11. Yu S.-Y., Yan B. Positive Solutions for a Singular Elliptic Equation Arising in a Theory of Thermal Explosion. *Mathematics*, 2021, vol. 9, pp. 2173, DOI: 10.3390/math9172173.
12. Grishin A.M. Nekotorye zadachi teorii vosplamenenija [Some problems of ignition theory]. *Prikladnaja mehanika i tehničeskaja fizika* [Journal of Applied Mechanics and Technical Physics], 1962, no. 5, pp. 75-79.
13. Merzhanov A.G., Zelikman E.G., Abramov V.G. Vyrozhdennye rezhimy teplovogo vzryva [Degenerate modes of thermal explosion]. *Doklady AN SSSR* [Reports of AN USSR], 1968, vol. 180, no. 3, pp. 639-642.
14. Boddington T., Gray P., Robinson C. Thermal explosions and the disappearance of criticality at small activation energies: exact results for the slab. *Proceedings of the Royal Society A. Mathematical, Physical and Engineering Sciences*, 1979, vol. 368, pp. 441–461, DOI: 10.1098/rspa.1979.0140
15. Anisimov S.I., Vitkin E.I. Nekotorye variacionnye zadachi teorii teplovogo vzryva [Some variational problems in thermal explosion theory]. *Prikladnaja mehanika i tehničeskaja fizika* [Journal of Applied Mechanics and Technical Physics], 1966, vol. 7, pp. 109–110, DOI: 10.1007/BF00917676
16. Zarubin V.S., Kuvyrkin G.N., Savel'eva I.Ju. Variacionnaja forma modeli teplovogo vzryva v tverdom tele s zavisjashhej ot temperatury teploprovodnost'ju [The Variational Form of the Mathematical Model of a Thermal Explosion in a Solid Body with Temperature-Dependent Thermal Conductivity]. *Teplofizika vysokih temperature* [High Temperature], 2018, vol. 56, pp. 223-228, DOI: 10.1134/S0018151X18010212
17. Graham-Eagle J.G., Wake G.C. Theory of thermal explosions with simultaneous parallel reactions. II. The two- and three-dimensional cases and the variational method. *Proceedings of the Royal Society A. Mathematical, Physical and Engineering Sciences*, 1985, vol. 401, pp. 195-202, DOI: 10.1098/rspa.1985.0094.
18. Donskoj I.G. Chislennaja ocenka kritičeskikh uslovij v zadache o teplovom vzryve s fluktuacijami reakcionnoj sposobnosti [Numerical estimation of critical conditions in thermal explosion problem for a medium with fluctuations of reactivity. *Informacionnye i matematicheskie tehnologii v nauke i upravlenii* [Information and Mathematical Technologies in Science and Management], 2021, no. 1, pp. 54-65. DOI: 10.38028/ESI.2021.21.1.005
19. Ajadi S.O. A note on the thermal stability of a reactive non-Newtonian flow in a cylindrical pipe. *International Communications in Heat and Mass Transfer*, 2009, vol. 36, pp. 63-68, DOI: 10.1016/j.icheatmasstransfer.2008.09.005.
20. Ajadi S.O. The influence of viscous heating and wall thermal conditions on the thermal ignition of a Poiseuille/Couette reactive flow // *Chemical physics*, 2010, v. 29, no. 8, p. 47-54.
21. Blouquin R., Joulin G. On a Variational Principle for Reaction/Radiation/Conduction Equilibria. *Combustion Science and Technology*, 1996, vol. 112, pp. 375-385, DOI: 10.1080/00102209608951968.
22. Moise A., Pritchard H.O. Newton-variational solution of the Frank–Kamenetskii thermal explosion problem. *Canadian Journal of Chemistry*, 1989, vol. 67, p. 442, DOI: 10.1139/v89-069.

23. Wazwaz A.-M. Solving the non-isothermal reaction-diffusion model equations in a spherical catalyst by the variational iteration method. *Chemical Physics Letters*, 2017, vol. 679, pp. 132-136, DOI: 10.1016/j.cplett.2017.04.077.
24. Baikov V.A., Gazizov R.K., Ibragimov N.Kh. Priblizhennyye simmetrii [Approximate symmetries]. *Matematicheskij sbornik [Mathematics of USSR – Sbornik]*, 1989, vol. 64, pp. 427–441. DOI: 10.1070/SM1989v064n02ABEH003318.
25. Gorgone M., Oliveri F. Approximate Noether Symmetries of Perturbed Lagrangians and Approximate Conservation Laws. *Mathematics*, 2021, vol. 9, pp. 2900. DOI: 10.3390/math9222900.
26. Donskoy I.G. Numerical estimation of thermal ignition conditions for reactive medium with Gaussian distribution of activation energy. *Journal of Physics: Conference Series*, 2021, vol. 2119, p. 012102. DOI:10.1088/1742-6596/2119/1/012102.
27. Takeno T.. Ignition criterion by thermal explosion theory. *Combustion and Flame*, 1977, vol. 29, pp. 209-211. DOI: 10.1016/0010-2180(77)90108-0.
28. Barzykin V.V., Merzhanov A.G. Kraevaja zadacha v teorii teplovogo vzryva [Boundary value problem in thermal explosion theory]. *Doklady AN SSSR [Reports of AN USSR]*, 1958, vol. 120, pp. 1271-1273.
29. Gostincev Yu.A. Metod privedenija k obyknovennym differencial'nym uravnenijam v zadachah nestacionarnogo gorenija poroha [Method of reduction to ordinary differential equations in problems of the nonstationary burning of solid propellants]. *Fizika gorenija i vzryva [Combustion, Explosions and Shock Waves]*, 1967, vol. 3, pp. 218–220. DOI: 10.1007/BF00791864
30. Sieniutycz S. The variational principles of classical type for non-coupled non-stationary irreversible transport processes with convective motion and relaxation. *International Journal of Heat and Mass Transfer*, 1977, vol. 20, pp. 1221-1231. DOI: 10.1016/0017-9310(77)90131-4
31. He J.-H. Variational principles for some nonlinear partial differential equations with variable coefficients. *Chaos, Solitons & Fractals*, 2004, vol. 19, pp. 847-851. DOI: 10.1016/S0960-0779(03)00265-0
32. Dik I.G., Tolstyh A.V. Vosplamenenie produvaemogo sloja [Ignition of a porous layer with a flow of heat carrier]. *Fizika gorenija i vzryva [Combustion, Explosions and Shock Wave]*, 1994, vol. 30, pp. 135–139. DOI: 10.1007/BF00786117
33. Boddington T., Gray P., Wake G.C. Criteria for thermal explosions with and without reactant consumption. *Proceedings of the Royal Society A. Mathematical, Physical and Engineering Sciences*, 1977, Vol. 357, pp. 403–422. DOI: 10.1098/rspa.1977.017
34. Adler J. Thermal explosion theory for reactive flow between parallel heated walls. *Combustion and Flame*, 1975, Vol. 24, pp. 151-158. DOI: 10.1016/0010-2180(75)90142-X

Igor G. Donskoy, *Cand. Sci. (Eng.), Senior Researcher in Laboratory of Thermodynamics, Melentiev Energy Systems Institute Siberian Branch of the Russian Academy of Sciences, donskey.chem@mail.ru, Irkutsk, Russia, Lermontov Str., 130*

Статья поступила в редакцию 18.02.2022; одобрена после рецензирования 23.03.2022; принята к публикации 25.03.2022.

The article was submitted 02.18.2022; approved after reviewing 03.23.2022; accepted for publication 03.25.2022.

Синтез системы управления, мониторинга и оценивания состояния анаэробного биореактора

Фоменкова Анастасия Алексеевна, Ключарев Александр Анатольевич, Колесникова Светлана Ивановна

Санкт-Петербургский государственный университет аэрокосмического приборостроения, Россия, Санкт-Петербург, skolesnikova@yandex.ru

Аннотация. В статье представлена концептуальная схема мониторинга и формализации управляющего воздействия на техническое состояние сложной системы на примере анаэробного биореактора. Представлен математический аппарат и реализующее его алгоритмическое обеспечение, как основа синтеза интеллектуальной системы управления технологическими процессами на предприятиях. Приведены примеры аналитического синтеза адаптивной системы управления для нелинейного многомерного объекта и алгоритм наблюдателя неизмеряемых переменных.

Ключевые слова: система анаэробной биологической очистки, мониторинг, нелинейный синтез управления, инвариант динамического объекта, оценка состояния сложного объекта, наблюдатель состояния

Цитирование: Фоменкова А.А., Ключарев А.А., Колесникова С.И. Синтез системы управления, мониторинга и оценивания состояния анаэробного биореактора // Информационные и математические технологии в науке и управлении. – 2022. – № 1 (25). – С. 21-34. – DOI:10.38028/ESI.2021.25.1.002.

Введение. Современные системы анаэробной биологической очистки (САБО) сточных вод [1-5] позволяют очищать стоки с высокими концентрациями органических загрязнений и повышать энергоэффективность производства за счет вырабатываемого биогаза.

Актуальность разработки системы мониторинга процессов очистки сточных вод с целью обеспечения нормативных показателей не нуждается в обосновании, поскольку:

- 1) решение задач современного производства должно отвечать понятиям рационального природопользования и ресурсосбережения, относящихся к приоритетному направлению развития науки, технологий и техники РФ;
- 2) существующие системы биологической очистки обладают рядом недостатков, связанных с серьезными экологическими угрозами, не преодоленных на данный момент, несмотря на возрастающее число публикаций (см., напр., обзор в [6]);
- 3) высокая структурная и физическая сложность биологических систем очистки [4] не позволяют полностью описать все взаимосвязи процессов с учетом состояния биомассы (неустойчивого объекта);
- 4) нелинейность математического описания ограничивает возможности применения классических математических моделей и алгоритмов оценивания состояния биореакторов (часть параметров которых не подлежат непосредственному измерению);
- 5) САБО функционируют в условиях внутренних и внешних возмущений, что вкупе со сложностью математического описания создает серьезную проблему для конструирования управления с требуемыми свойствами.

Под **сложным динамическим объектом** будем понимать объект [7], характеризующийся свойствами: 1) отсутствие (недостаточность) априорной информации для построения адекватной модели или невозможность (структурно сложный объект) полного аналитического

го описания; 2) нелинейность имеющихся моделей описания; 3) динамическая нестационарность сопровождающего его поведение процесса.

Состоянию $X(t)$ в некоторый момент t в фазовом пространстве соответствует точка $X = (X_1, \dots, X_n)$ – изображающая точка, движение которой в фазовом пространстве по определенной траектории (фазовой) сопоставлено изменению $X(t)$ [7 - 10].

Макросостояние объекта (системы) есть некоторая подобласть фазового пространства динамического объекта, характерные физически интерпретируемые (экспертно) свойства которого могут быть формализованы предельным ($t \rightarrow \infty$) уравнением $\psi(X) = 0$, где $\psi(X)$ – целевая макропеременная [9].

Мониторинг – система наблюдения, оценивания и прогноза основных показателей процесса анаэробной биологической очистки, результаты которой являются основанием для выбора соответствующих управленческих решений по обеспечению режима функционирования объекта САБО требуемого качества [8].

При изменении параметров существующий динамический режим САБО может потерять устойчивость, но некоторое время еще оставаться в прежней области фазового пространства. При этом, несмотря на то, что статистические свойства наблюдаемого объекта значимо не меняются, со временем устанавливается другой динамический режим, и важно как можно раньше и точнее обнаружить «предвестники» момента изменения состояния системы.

Система управления – система, состоящая из управляющей подсистемы (в частности, регулятора) и объекта управления [8]. Здесь в состав управляющей подсистемы входят подсистемы мониторинга состояния объекта и блок формирования управляющих воздействий.

1. Постановка задачи. Разработку математического и алгоритмического обеспечения мониторинга и формирование целевого воздействия на системы анаэробной биологической очистки сточных вод, направленного на сохранение работоспособности и снижение аварийности в процессе эксплуатации, свяжем с решением следующих задач:

- 1) формализовать математическую модель САБО, исходя из особенностей процессов, как объект мониторинга и управления;
- 2) построить наблюдатель неизмеряемых переменных состояния САБО;
- 3) построить систему управления САБО на принципах нелинейной адаптации и метода аналитического проектирования агрегированных регуляторов (АКАР) [9-11].

2. Формализация системы мониторинга и управления.

2.1. Формализация САБО как объекта мониторинга. Для синтеза системы мониторинга удобно рассматривать САБО как объект класса динамических систем, допускающих представление:

$$Z = \{X, Y\} = \{X(t), Y(t), t_0 \leq t \leq T\}, t \geq t_0, \quad (1)$$

где $X(t) \in R^n$ – вектор переменных состояния системы; $Y(t) = G(X(t), \xi(t)) \in R^m, m < n$ – вектор измеряемых переменных, $G()$ – известная функция своих аргументов; $\xi(t)$ – неизвестное ограниченное возмущение как функция времени.

Вообще говоря, полагается, что процесс $Z(t)$ характеризует состояние САБО и принимает значения в произвольном измеримом (фазовом) пространстве $(Z, \mathfrak{F}_Z)$ со счетно-порожденной σ -алгеброй подмножеств пространства Z (общее предположение, охватывающие практически все возможные ситуации).

Относительно динамики поведения процесса $Z(t)$ на $[t_0, \infty)$ выдвинуто $I > 0$ альтернативных гипотез $(\Omega_1, \dots, \Omega_I)$, составляющих полную группу событий и интерпретируемых как классы состояний САБО. Другими словами, временной фрагмент реализации случайного процесса $Z(t)$ может находиться только в одной из заданных экспертом, или выделенных по обучающей выборке по заданным ограничениям областей (классов) $\Omega_i, i = \overline{1, I}$.

Предполагается, что с вероятностью 1 за конечный промежуток времени происходит конечное число изменений состояний САБО и для определенности процесс (1) допускает представление:

$$Z(t) = \sum_{i=1}^I Z_i(t) \chi_i(t), \quad (2)$$

где $Z_i(t)$ – вообще говоря, случайный элемент (при фиксированном t), принимающий значения из $(Z, \mathfrak{F}_Z)$, $\chi_i(t)$ – индикаторная функция со значением 1, если САБО в момент t находится в состоянии $\Omega_i, i = \overline{1, I}$, и ноль – иначе. Наблюдение $Y(t)$ осуществляется в соответствии с дискретным планом $t \in \{t_0, t_1, \dots, t_n, \dots\}$, $t_j = t_0 + j\Delta$, $j = 0, 1, 2, \dots$, отвечающим системе реального времени (масштабы времени наблюдающей (управляющей) системы и контролируемого ею объекта синхронизированы) (рис. 1).



Рис. 1. Структурная схема взаимодействия двух систем: мониторинга САБО и управления

Схема на рис.1 есть основа для синтеза соответствующих алгоритмов реализации каждого положения и их информационного взаимодействия для организации приемлемого управления сложной системой САБО.

2.2. Формализация задачи управления макросостояниями САБО. Ставится задача синтеза системы управления в пространстве состояний для динамического объекта вида (3), функционирующего в условиях неопределенностей, на характер которых влияют и неполнота описания, и наличие внутренних и внешних возмущений (рис. 2):

$$\begin{aligned} \dot{X} &= F(X, \xi) + U, \quad \hat{X} = g(X), \\ U &= U(\psi, \hat{X}), \quad \psi = \psi(\hat{X}), \quad \dim(U) = \dim(\psi) < n, \end{aligned} \quad (3)$$

где $\hat{X}$ – вектор оценки состояния, $g()$ – функция от измеряемых переменных (наблюдатель состояния), $\psi(X)$ – целевая макропеременная, условие $\psi(X(t)) = 0, t \rightarrow \infty$ определяет так называемое целевое макросостояние (многообразие или инвариант системы в случае его асимптотической устойчивости) для объекта управления $\dot{X} = F(X, \xi) + U$; $F()$ – известная функция; U – управление, конструкция которого должна обеспечивать асимптотическую

устойчивость системы управления (3) в целом с заданным временем переходных процессов в ней (класс АКАР-управлений, см. раздел 2.3).

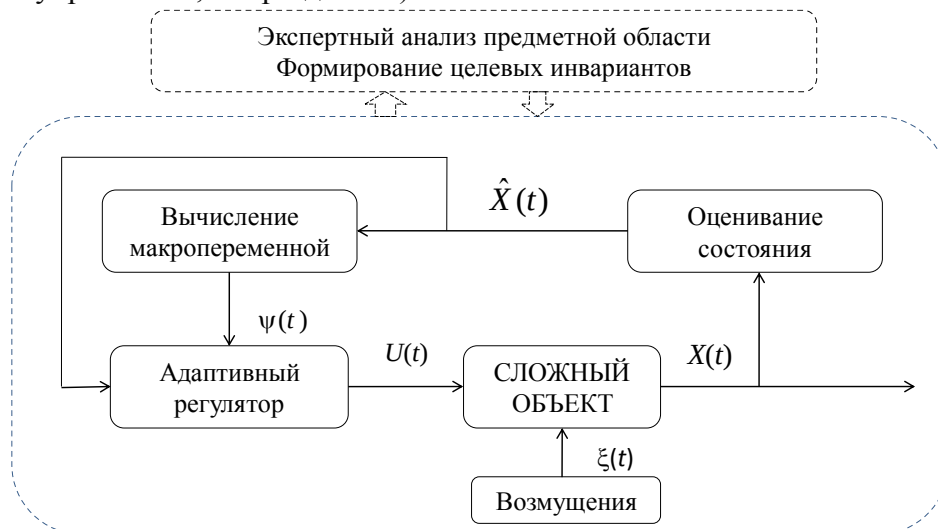


Рис. 2. Структурная схема системы управления на основе алгоритма нелинейной адаптации

2.3. Математическая модель САБО как объекта управления. Согласно базовой модели ADM-1 [12] для описания работы анаэробного биореактора достаточно рассмотрения двухстадийной модели процесса (рис. 3), исходя из которой концентрация органического загрязнения сточной воды пересчитывается в эквивалентную концентрацию глюкозы S . На вход в биореактор подается сточная вода с эквивалентной концентрацией глюкозы S_{in} .

Двухстадийный процесс анаэробного брожения, протекающий по схеме: $S + B_1 \rightarrow P$, $P + B_2 \rightarrow G$, схематично представлен на рис. 3 с вектором состояний $X \in R^7$, $X = (S, B_1, P, B_2, G, \theta, Q)^T$.

Динамика процесса очистки следующая. В биореактор-смеситель объемом V в момент времени t подается субстрат с расходом $Q(t)$ и концентрацией органических загрязнений $S_{in}(t)$.

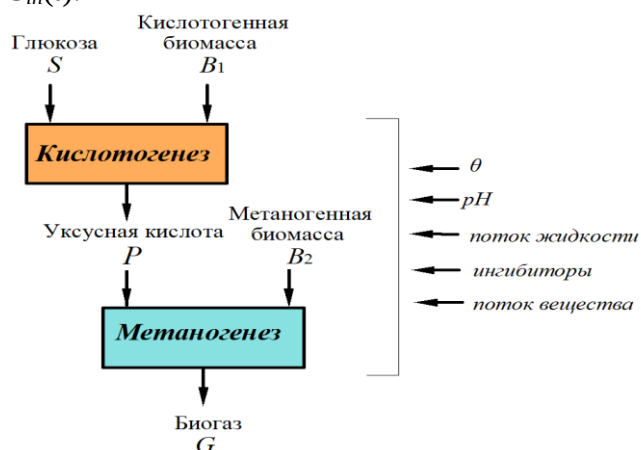


Рис. 3. Схема двухстадийного процесса анаэробного разложения органического загрязнения сточной воды в анаэробном биореакторе

На выходе биореактора формируется поток очищенной воды с концентрацией загрязнений $S_{out}(t) = S(t) + P(t)$, состоящих из остатков исходного субстрата с концентрацией $S(t)$ и промежуточных продуктов анаэробного брожения с концентрацией $P(t)$. В процессе очистки исходное органическое загрязнение проходит стадии кислотогенеза и метаногенеза. В результате жизнедеятельности кислотогенной биомассы с концентрацией B_1 образуются промежуточные продукты очистки, которые перерабатываются в биогаз метаногенной биомассой с концентрацией B_2 .

Двухстадийный процесс анаэробной биологической очистки сточных вод в биореакторе-смесителе можно описать системой дифференциальных уравнений:

$$\begin{aligned}\dot{S}(t) &= \frac{Q(t)}{V} (S_{in}(t) - S(t)) - k_1(\theta) B_1(t) - k_2(\theta) \frac{S(t) B_1(t)}{k_3(\theta) + S(t)}, \\ \dot{B}_1(t) &= -\frac{Q(t)}{V} B_1(t) + k_4(\theta) \frac{S(t) B_1(t)}{k_3(\theta) + S(t)}, \\ \dot{P}(t) &= -\frac{Q(t)}{V} P(t) + k_5(\theta) B_1(t) + k_6(\theta) \frac{S(t) B_1(t)}{k_3(\theta) + S(t)} - k_7(\theta) B_2(t) - k_8(\theta) \frac{P(t) B_2(t)}{k_9(\theta) + P(t)}, \\ \dot{B}_2(t) &= -\frac{Q(t)}{V} B_2(t) + k_{10}(\theta) \frac{P(t) B_2(t)}{k_9(\theta) + P(t)}, \\ \dot{G}(t) &= -G(t) + k_{11}(\theta) \frac{P(t) B_2(t)}{k_9(\theta) + P(t)} \frac{k_{12}(\theta)}{k_{12}(\theta) + P(t)},\end{aligned}\quad (4)$$

дополнив которую уравнениями

$$\begin{aligned}\dot{\theta}(t) &= z_1(t) + u_1(t), \\ \dot{Q}(t) &= z_2(t) + u_2(t),\end{aligned}\quad (5)$$

получим объект управления. Здесь $z_1(t), z_2(t)$ и $u_1(t), u_2(t)$ - неизвестные ограниченные функции времени (возмущения) и функции управления, соответственно; $\theta(t)$ - рабочая температура в биореакторе, $k_i(\theta)$ - кинетические параметры процесса с законом изменения $k_i(\theta) = k_{i35^\circ C} \exp(c(\theta - 35^\circ C))$, $k_{i35^\circ C}$ - значение параметра при нормальной температуре жизнедеятельности биомассы $35^\circ C$, c - заданная постоянная.

Определяя конкретное целевое состояние САБО как формализованное уравнение $\psi(\hat{X}(t)) = 0$, $t \rightarrow \infty$, получим задачу управления состоянием в виде (3)-(5).

Далее для краткости записи будем ссылаться на систему (3)-(5), используя обозначение $\Theta = (X; u + \xi)$, $X \in R^7, u, \xi \in R^2$.

2.4. Формализация макросостояний САБО. Традиционно формулируемое техническое состояние динамической системы через совокупность характерных значений структурных параметров (признаков в определенный момент времени при определенных условиях внешней среды), увязанных с данными технической документации, в структуре мониторинга (рис. 1) и управления в частности предполагает следующую формализацию (как пример).

1. Основной целью функционирования биореактора в системе очистки сточных вод является снижение концентрации органического загрязнения сточной воды до (или ниже) заданного нормативного значения $S_{норм}$. При этом желательно максимально полно преобразовать загрязнения в биогаз. Если G^* - расчетное количество биогаза, которое возможно получить из заданного сырья, $\psi_1^*(t)$ и $\psi_2^*(t)$ - целевые инварианты для системы анаэробной биологической очистки, то формальная модель цели управления примет вид:

$$\begin{aligned}\psi_1^*(t) &= S_{out}(t) + P_{out}(t) - S_{норм} \rightarrow 0; \\ \psi_2^*(t) &= G_{out}(t) - G^* \rightarrow 0, t \rightarrow \infty.\end{aligned}\quad (6)$$

2. Увеличение количества газа не всегда приводит к повышению энергоэффективности $E(t)$ системы из-за увеличения расходов на обогрев биореактора, $E(t) = \frac{E_G(t) - E_{heating}(t)}{E_{heating}(t)}$,

где $E_G(t)$ - количество энергии, которую потенциально можно получить из биогаза, $E_{heating}(t)$ - энергия, расходуемая для обогрева биореактора и в общем случае определяется из уравнения

теплового баланса системы. Тогда целевой инвариант для системы очистки сточных вод примет вид:

$$\begin{aligned}\psi_1^*(t) &= S_{out}(t) + P_{out}(t) - S_{norm} \rightarrow 0; \\ \psi_2^*(t) &= E(t) \rightarrow \max, t \rightarrow \infty.\end{aligned}\quad (7)$$

3. Для биогазовых установок описание целевого макросостояния целесообразно выбрать в виде:

$$\begin{aligned}\psi_1^*(t) &= P(t) - P^* \rightarrow 0; (P(t) \rightarrow \min, P(t) \leq P^*) \\ \psi_2^*(t) &= E(t) \rightarrow \max; (\psi_2^*(t) = E(t) \rightarrow 1-).\end{aligned}\quad (8)$$

Соотношения (6)-(8) представляют собой примеры формализации макросостояний САБО (инвариантов в случае асимптотической устойчивости целевых систем).

3. Решение задачи управления достижением целевого макросостояния.

3.1. Суть алгоритма нелинейной адаптации (НАД). Алгоритм энергосберегающего (осторожного) управления на принципах нелинейной адаптации (НАД) [13] есть по сути расширение классического метода аналитического проектирования агрегированных регуляторов (АКАР)) [9].

Под **адаптацией** будем понимать процесс изменения усредненных характеристик динамического объекта с неопределенностью в описании среды функционирования, а под алгоритмом адаптации способность асимптотически переводить состояние системы в окрестность заданного инвариантного множества в такой среде.

Как известно [9], синтез АКАР-управлений в общем виде допускает иерархичность процедуры конструирования управления. Так, в задачах стабилизации иерархичность синтеза управления связана с необходимостью осуществлять управление переменными, прямой доступ к которым невозможен. При этом используется совокупность частных критериев (целевых функционалов) $\{\Phi^s\}_{s=1,e}$ с определенным правилом подчинения и согласованности, каждый из критериев отражает некоторое локальное требование к качеству соответствующей подсистемы управления.

Постановка непрерывной задачи классического АКАР-управления. Рассматривается нелинейный векторный объект с описанием в виде системы ОДУ $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}), \mathbf{x}, \mathbf{f} \in R^n, \mathbf{u} \in R^m, m \leq n$. Требуется определить закон управления $\mathbf{u}(t) = \mathbf{u}(\mathbf{x}(t))$ в пространстве состояний для вывода объекта управления из некоторого начального состояния $\mathbf{x}(0)$ в окрестность многообразия $\psi(t) = \psi(\mathbf{x}(t)) = 0, t \rightarrow \infty$ и обеспечить асимптотически устойчивое движение объекта в его окрестности.

Имеют место предположения:

1) глобальный минимум функционала качества, отражающего вариационный принцип при достижении цели управления $\psi(\mathbf{x}) = 0, t \rightarrow \infty$:

$$\Phi = \int_0^\infty \sum_{i=1}^m (\psi_i^2(t) + w_i^2 \dot{\psi}_i^2(t)) dt \rightarrow \min, \psi \xrightarrow{t \rightarrow \infty} 0, \psi \in R^m$$

где $\mathbf{w} \in R^m$ - параметр синтезируемой системы управления, регулирующий длительность движения изображающей точки объекта управления до окрестности $\psi(\mathbf{x}) = 0, t \rightarrow \infty$;

2) ограниченность правой части и реализуемость поведения системы $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u})$ при всех допустимых управлениях;

3) существование устойчивой целевой системы для исходной модели, отвечающей одновременно динамическим и/или инженерным требованиям, не противоречащим физиче-

ским возможностям системы [14];

Ранее в [11] был конструктивно доказан ряд утверждений об условиях существования такого управления для скалярного объекта.

Основные положения алгоритма адаптации объекта, функционирующего под воздействием ограниченного возмущения.

1. Расширение фазового пространства за счет перевода внешнего возмущения ξ во внутреннюю дополнительную фазовую переменную z с законом изменения [13, 15]

$$\dot{z}(t) = \eta \psi^*(t), \quad \eta = \text{const} > 0$$

2. Вывод структуры синтезируемого регулятора реализуется методологией классического синтеза АКАР для полученной на шаге 1 замкнутой системы. При этом на конечном e -ом уровне иерархии синтеза используется макропеременная вида $\psi^{(e)} = \psi^* + kz$, $k = \text{const} > 0$, где ψ^* - целевая макропеременная из постановки задачи управления, и формулируется соответствующая вариационная задача $(\Phi^e, \psi^{(e)})$ с функциональным дифференциальным уравнением линейного вида, доставляющим устойчивое решение уравнению Эйлера-Лагранжа для вариационной задачи $(\Phi^e, \psi^{(e)})$

$$w_e \dot{\psi}^{(e)}(t) = 0, t \rightarrow 0^*$$

Функционал Φ^e является критерием качества искомого управления при перемещении изображающей точки системы от многообразия $\psi^{(e-1)}(t) = 0, t \rightarrow \infty$ к многообразию $\psi^{(e)}(t) = 0, t \rightarrow \infty$. Верхний индекс e (для объекта САБО $e=1,2$ в зависимости от вида целевой макропеременной ψ) означает конечный номер иерархии в процессе синтеза классического АКАР-регулятора.

Возникающие постоянные в процессе синтеза w_s есть параметры системы управления, пропорциональные скорости достижения изображающих точек s -й подсистемы окрестности $\psi_s(t) = 0, t \rightarrow \infty$ согласно частному критерию $\Phi^s, s = \overline{1, e}$.

3.2. Конструирование системы управления для САБО на основе алгоритма нелинейной адаптации. Рассмотрим скалярный случай конструирования управления с целью достижения заданного целевого множества состояний (для краткости изложения без ограничения общности).

Алгоритм интегральной нелинейной адаптации для задачи $\psi^*(t) = G(t) - G^*, t \rightarrow \infty$ с учетом возмущения по скалярному управлению. Здесь G^* - целевое (желаемое, допустимое) значение выхода биогаза. Здесь полагаем $Q = \text{const}$, и в описании (5) имеется одно уравнение

$$\dot{\theta}(t) = z(t) + u(t) \tag{9}$$

Шаг 1. Расширение фазового пространства (превращение исходного описания в замкнутую систему для синтеза управления и компенсации возмущения):

$$\begin{aligned} \dot{S}(t) &= -\frac{Q(t)}{V} (S_{in}(t) - S(t)) - k_1(\theta) B_i(t) - k_2(\theta) \frac{S(t) B_1(t)}{k_3(\theta) + S(t)}, \\ \dot{B}_1(t) &= -\frac{Q(t)}{V} B_1 + k_1(\theta) \frac{S(t) B_1(t)}{k_3(\theta) + S(t)}, \\ \dot{P}(t) &= -\frac{Q(t)}{V} P(t) + k_5(\theta) B_1(t) + k_6(\theta) \frac{S(t) B_1(t)}{k_3(\theta) + S(t)} - k_7(\theta) B_2(t) - k_8(\theta) \frac{P(t) B_2(t)}{k_9(\theta) + P(t)}, \\ \dot{B}_2(t) &= -\frac{Q(t)}{V} B_2(t) + k_{10}(\theta) \frac{P(t) B_2(t)}{k_9(\theta) + P(t)}, \\ \dot{G}(t) &= -G(t) + k_{11}(\theta) \frac{P(t) B_2(t)}{k_9 + P(t) k_{12}(\theta) + P(t)}, \\ \dot{\theta}(t) &= z(t) + u(t), \end{aligned} \tag{10}$$

$$\ddot{x}(t) = \eta \psi^*(t), \eta = \text{const} > 0.$$

Обозначим $\Theta_0 = (X; u + \xi)$, $X = (S, B_1, P, B_2, G, \theta)^T \in R^6$, $u, \xi \in R^1$ – исходное описание, а расширенную систему (10) – исходную для алгоритма синтеза регулятора через $\Theta_1 = (X_1; u + z)$, $X_1 = (S, B_1, P, B_2, G, \theta, z)^T$, $X_1 \in R^7$, $u, z \in R^1$.

Замечание 1. Возмущение $z(t)$ есть ограниченная и непрерывная функция времени. Для моделирования используются результаты: а) непрерывные функции могут быть представлены как решения дифференциальных уравнений [16]; б) возмущения $\xi(t)$ могут быть представлены как суперпозиция кусочно-непрерывных функций $z(t)$ – волновое представление [15].

Шаг 2. Вводим вспомогательную макропеременную ψ_1 (для передачи функции управления от переменной $\theta(t)$ к внутренней переменной управления φ , где $\varphi = \varphi(G, z)$):

$$\psi_1 = \theta - \varphi. \quad (11)$$

Шаг 3. Выписываем уравнение для устойчивых экстремалей вариационной задачи

$$(\Phi_1 \psi_1), \Phi_1 = \int_0^\infty (\psi_1^2(t) + T_1^2 \dot{\psi}_1^2(t)) dt \rightarrow \min: \\ T_1 \dot{\psi}_1 + \psi_1 = 0, T_1 = \text{const} > 0 \quad (12)$$

Подставляем (11) в (12):

$$T_1 \dot{\psi}_1 + \psi_1 = T_1 \dot{\theta} - T_1 \dot{\varphi} + \psi_1 = 0, \dot{\varphi}(G, z) = \frac{\partial \varphi}{\partial G} \dot{G} + \frac{\partial \varphi}{\partial z} \dot{z} = \frac{\partial \varphi}{\partial G} \dot{f}_G + \frac{\partial \varphi}{\partial z} \eta \psi^*,$$

где f_G – правая часть описания в системе (10).

Из (12) следуют формулы для *структуры управления*:

$$u = \dot{\varphi} - T_1^{-1} \psi_1 - z = \frac{\partial \varphi}{\partial G} f_G + \frac{\partial \varphi}{\partial z} \eta \psi^* - T_1^{-1} \psi_1 - z \quad (13)$$

Выражение (13) определяют структуру регулятора для (9) с точностью до неизвестной функции внутреннего управления φ и ее частных производных по переменным G, z .

Шаг 4. Осуществляем декомпозицию базовой модели на многообразии $\psi_1 = 0$, получаем систему уравнений, исходную для дальнейшего синтеза внутреннего управления φ . Заменяем в (8) $\theta = \varphi$. Получим описание

$$\Theta_2 = (X_2; \varphi), X_2 = (S, B_1, P, B_2, G, z)^T, X_2 \in R^6, \varphi \in R^1 \quad (14)$$

Система (14) – исходная для алгоритма синтеза регулятора.

Шаг 5. Вводим вспомогательную макропеременную второго этапа синтеза регулятора $\psi_2 = \psi^* + \mu z$, $\mu = \text{const} > 0$ и выписываем уравнение для устойчивых экстремалей вариационной задачи $(\Phi_2 \psi_2), \Phi_2 = \int_0^\infty (\psi_2^2(t) + T_2^2 \dot{\psi}_2^2(t)) dt \rightarrow \min:$

$$T_2 \dot{\psi}_2 + \psi_2 = T_2 (\dot{\psi}^* + \mu \dot{z}) + \psi_2 = 0, T_2 = \text{const} > 0 \quad (15)$$

Подставляем производные целевых макропеременных $\psi^*(t) = G(t) - G^*$ в (15), с учетом уравнений (14), получим явный вид функции φ и ее частных производных $\frac{\partial \varphi}{\partial G}, \frac{\partial \varphi}{\partial z}$.

Итоговая система управления для достижения требуемого состояния $\psi^*(t) = G(t) - G^*$, $t \rightarrow \infty$ согласно алгоритму нелинейной адаптации имеет вид совокупности уравнений (4), (9), (13) (рис. 4).

Как следует из рис. 4, вопрос типа и уровня возмущений с практически приемлемым качеством системы управления требует отдельного рассмотрения и связан с аккуратностью подбора параметров регулятора (здесь это величины T_1, T_2, η , влияющие на скорость достижения установившегося режима и качество переходного процесса).

Утверждение 1. Управление $u(t)$, если существует, обеспечивает асимптотическую устойчивость объекту управления САБО (4), (9) в окрестности $\psi^*(t) = 0, t \rightarrow \infty$.

Доказательство утверждения 1 конструктивное [9-11] и непосредственно следует из самого метода АКАР [9] и его обобщения [13] - алгоритма интегральной адаптации для данного описания (4), (9) (см. шаги 1-5) при выполнении следующих условий корректного применения АКАР-конструирования регулятора:

а) существует инвариант для объекта управления $\psi^*(t) = \psi^*(X(t)) = 0, t \rightarrow \infty$ с аттрактивным [10] свойством, где $\psi^*(X(t))$ непрерывно дифференцируемая функция;

б) управление должно обеспечить существование и ограниченность решений исходной системы;

в) исходная система в окрестности $\psi^*(t) = 0, t \rightarrow \infty$ стабилизируема.

Тогда согласно следствию (15) из уравнения Эйлера-Лагранжа (см. приложение) для функционала Φ_2 получим $\psi_2 = \psi^* + \mu z = 0, t \rightarrow \infty$, или предельное уравнение $\psi^* = -\mu z, t \rightarrow \infty$, согласно которому в декомпозированной системе на многообразии $\psi_2 = 0$ уравнение, отвечающее за моделирование возмущения, примет вид $\dot{z}(t) = -\mu\eta z(t), \mu, \eta = \text{const} > 0$, что гарантирует $\psi^* = -\mu z, t \rightarrow \infty$ и, следовательно, асимптотически устойчивое достижение цели управления $\psi^*(t) = 0, t \rightarrow \infty$.

Замечание 2. В [17] доказано, что свойство аттрактивности инвариантного многообразия является достаточным для асимптотической стабилизации заданного положения равновесия системы (см. также теорему 1.1 в [10]). Отдельно решается задача доказательства инвариантности заданного целевого многообразия, реализующего желаемое технологическое состояние объекта управления.

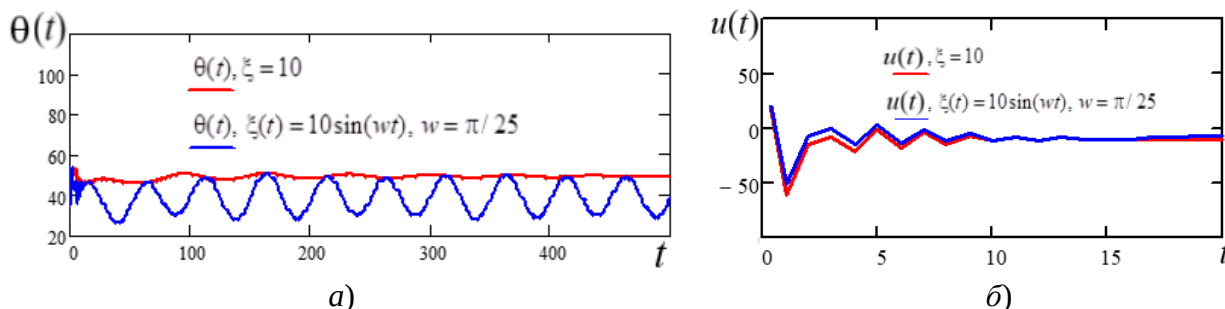


Рис. 4. Графики а) переменной управления θ и б) функции управления $u(t)$ под воздействием разного типа возмущений (постоянные и гармонические)

Замечание 3. Аналитически сформулированная цель управления есть основание для синтеза системы синергетического управления (цена вопроса). Вопрос инвариантности заданного целевого многообразия, реализующего желаемое технологическое (и не только, см. [9]) состояние объекта управления, как правило, исследуется отдельно (в том числе на основе имитационного моделирования, экспертного заключения).

3.3. Конструирование наблюдателя состояния для системы управления на основе алгоритма нелинейной адаптации. Поскольку в описании (4) не все переменные измеряются, то требуется алгоритм оценивания неизмеряемых переменных и представление их как функций от измеряемых. Сложность объекта (4) заключается еще и в возможных параметрических флуктуациях в силу нестабильных биохимических процессов, весьма чувствительных к внешним воздействиям.

Цель данного раздела – показать принципиальную возможность встраивания в систему управления САБО (с частным описанием (4), (9)) наблюдателя состояния и, следовательно, реализуемости схемы на рис. 2. Поставим задачу построения наблюдателей $\hat{B}_1(t), \hat{B}_2(t)$

для переменных $B_1(t), B_2(t)$ как функций от переменных состояния S, P, G, θ для объекта управления (4), (5) с обозначением $\Theta_0 = (X; u + \xi)$, $X = (S, B_1, P, B_2, G, \theta)^T \in R^6, u, \xi \in R^1$.

Согласно алгоритму адаптации НАД, для синтеза управления и компенсации возмущений следует осуществить перевод в расширенное фазовое пространство $\Theta_0 \rightarrow \Theta_1$, где $\Theta_1 = (X_1; u + z)$, $X_1 = (S, B_1, P, B_2, G, \theta, z)^T$, $X_1 \in R^7, u, z \in R^1$; $\dot{z}(t) = \eta \psi^*(t)$, $\eta = \text{const} > 0$.

Алгоритм построения наблюдателя состояния для применения метода нелинейной адаптации в задаче управления (4), (9). Для компактного изложения алгоритма на основе АКАР-наблюдателей [9] нам удобно унифицировать обозначения:

$$Y = (Y_1, Y_2, Y_3, Y_4, Y_5, Y_6, Y_7)^T = (S, P, G, z, \theta, B_1, B_2)^T.$$

1. Задаются некоторые промежуточные функции, зависящие от наблюдаемой и измеряемых переменных (здесь предполагаем линейную зависимость):

$$\phi_1(Y) = \sum_{l=1}^5 \frac{Y_l}{N_l} + Y_6, \quad \phi_2(Y) = \sum_{l=1}^5 \frac{Y_l}{N_l} + Y_7. \quad (16)$$

где N_l - подбираемые параметры модели наблюдателя переменных Y_6, Y_7 .

2. Вводится совокупность макропеременных $\psi_i = \phi_i - \hat{\phi}_i$, $i=1, 2$ с требованием удовлетворять решению соответствующей вариационной задачи $(\Phi, \psi), \Phi = \int_0^\infty \sum_{l=1}^2 (\psi_l^2(t) + L_l^2 \dot{\psi}_l^2(t)) dt \rightarrow \min$, $\psi \in R^2$ и, соответственно, уравнениям:

$$\dot{\psi}_i - L_i \psi_i = 0, \quad i=1, 2. \quad (17)$$

Здесь $\hat{\phi}_i = \hat{\phi}_i(t)$ есть наблюдатель для Y_7, Y_8 в виде искомым функций времени, выбор постоянных L_i , $i=1, 2$ должен обеспечивать асимптотическую устойчивость решениям (17).

Решение вариационной задачи (Φ, ψ) с условием $\psi = \phi - \hat{\phi} \xrightarrow{t \rightarrow \infty} 0$, $\psi, \phi, \hat{\phi} \in R^2$, обеспеченное уравнением (17), приводит к предельному равенству $\phi = \hat{\phi}$, $t \rightarrow \infty$.

3. Определяется соотношение из уравнения (17) с учетом вида (4) и (9):

$$\sum_{l=1}^5 \frac{\partial \phi_1}{\partial Y_l} \dot{Y}_l + \dot{Y}_6 - \frac{d\hat{\phi}_1}{dt} - L_1(\phi_1 - \hat{\phi}_1) = 0, \quad \sum_{l=1}^5 \frac{\partial \phi_2}{\partial Y_l} \dot{Y}_l + \dot{Y}_7 - \frac{d\hat{\phi}_2}{dt} - L_2(\phi_2 - \hat{\phi}_2) = 0. \quad (18)$$

4. Исключение из уравнений (18) переменных, для которых строится наблюдатель состояния. С этой целью вводятся вспомогательные далее определяемые функции $\Gamma_{1,l}, \Gamma_{2,l}, \gamma_i$, $l=1, 5, i=1, 2$, не зависящие от неизмеряемых (наблюдаемых) переменных, такие, чтобы выполнялось равенство (представляем (18) в виде (19)):

$$\sum_{l=1}^5 \frac{\partial \phi_1}{\partial Y_l} \dot{Y}_l + \dot{Y}_6 - L_1 \phi_1 = \sum_{l=1}^5 \Gamma_{1,l} \dot{Y}_l + \gamma_1, \quad \sum_{l=1}^5 \frac{\partial \phi_2}{\partial Y_l} \dot{Y}_l + \dot{Y}_7 - L_2 \phi_2 = \sum_{l=1}^5 \Gamma_{2,l} \dot{Y}_l + \gamma_2. \quad (19)$$

Сопоставляя представления (18) и (19) (в предположении, что величины $\Gamma_{i,l}, \gamma_i$, $l=1, 5, i=1, 2$ определены), получим собственно оценочные функции времени $\hat{\phi}_i$ наблюдателя ϕ_i в виде линейных уравнений (относительно функций $\hat{\phi}_i$, $i=1, 2$)

$$\frac{d\hat{\phi}_1}{dt} = L_1 \hat{\phi}_1 - \left(\sum_{l=1}^5 \Gamma_{1,l} \dot{Y}_l + \gamma_1 \right), \quad \frac{d\hat{\phi}_2}{dt} = L_2 \hat{\phi}_2 - \left(\sum_{l=1}^5 \Gamma_{2,l} \dot{Y}_l + \gamma_2 \right). \quad (20)$$

С учетом описаний (4), (9), (18)-(20) получим

$$\sum_{l=1}^4 \frac{f_l}{N_l} + \frac{1}{N_5} (f_5 + u + z) + f_5 - L_1 \phi_1 = \sum_{l=1}^4 \Gamma_{1,l} f_l + \Gamma_{1,5} (f_5 + u + z) + \gamma_1, \quad (21)$$

$$\sum_{l=1}^4 \frac{f_l}{N_l} + \frac{1}{N_5} (f_5 + u + z) + f_6 - L_2 \phi_2 = \sum_{l=1}^4 \Gamma_{2,l} f_l + \Gamma_{2,5} (f_5 + u + z) + \gamma_2.$$

Выберем функции в (21) $\Gamma_{i,l}$, γ_i , $l = \overline{1,5}$, $i = 1, 2$ так, чтобы исключить наблюдаемые переменные, понятно, что это можно осуществить неоднозначно. Например, собирая из (21) все члены с управлением в виде:

$$\gamma_1 = (u + \hat{z})(N_5^{-1} - \Gamma_{1,5}), \quad \gamma_2 = (u + \hat{z})(N_5^{-1} - \Gamma_{2,5}). \quad (22)$$

и подставляя γ_i , $i = 1, 2$ в (21), получим

$$\sum_{l=1}^4 \frac{f_l}{N_l} + \frac{1}{N_5} f_5 + f_6 - L_1 \phi_1 = \sum_{l=1}^4 \Gamma_{1,l} f_l + \Gamma_{1,5} f_5,$$

$$\sum_{l=1}^4 \frac{f_l}{N_l} + \frac{1}{N_5} f_5 + f_7 - L_2 \phi_2 = \sum_{l=1}^4 \Gamma_{2,l} f_l + \Gamma_{2,5} f_5.$$

С учетом представления (16)

$$\sum_{l=1}^4 \frac{f_l}{N_l} + \frac{1}{N_5} f_5 + f_6 - L_1 \left(\sum_{l=1}^5 \frac{Y_l}{N_l} + Y_6 \right) = \sum_{l=1}^4 \Gamma_{1,l} f_l + \Gamma_{1,5} f_5, \quad (23)$$

$$\sum_{l=1}^4 \frac{f_l}{N_l} + \frac{1}{N_5} f_5 + f_7 - L_2 \left(\sum_{l=1}^5 \frac{Y_l}{N_l} + Y_7 \right) = \sum_{l=1}^4 \Gamma_{2,l} f_l + \Gamma_{2,5} f_5.$$

5. Из (23) методом неопределенных коэффициентов получают соотношения, связывающие величины $\Gamma_{i,l}$, $l = \overline{1,5}$, $i = 1, 2$ и описание объекта.

Метод неопределенных коэффициентов приводит к неединственным соотношениям. Традиционное его применение сводится к приравнению коэффициентов при одинаковых степенях в полиномах, либо при одинаковых наименованиях функций. Согласно идеологии машинного обучения сам способ определения коэффициентов, может быть параметром наблюдателя (рис. 5).

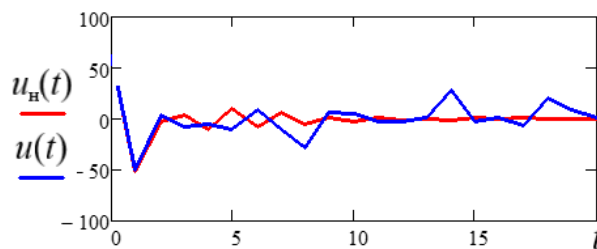


Рис. 5. Сглаживающий эффект управления u_n с наблюдателем состояния

Заключение. Результаты настоящего исследования, а именно: алгоритмы достижения заданных свойств технического объекта, позволяющие реализовать желаемые целевые состояния САБО, и алгоритмы оценивания и управления процессами в биореакторе для достижения и поддержания работоспособного состояния могут быть использованы в автоматизированных системах управления технологическим процессом очистки, при проектировании и выборе технических средств реализации системы мониторинга и поддержки принятия решений при эксплуатации анаэробных систем очистки производственных сточных вод.

Благодарности. Работа поддержана РФФИ (грант № 20-08-00747).

СПИСОК ИСТОЧНИКОВ

1. Kalyuzhnyi S.V. Batch anaerobic digestion of glucose and its mathematical modeling. II. Description, verification and application of model. *Biores. Technol*, 1997, vol. 59, pp. 249-258.
2. Boe K., Batstone D.J., Steyer J.-P., Angelidaki I. State indicators for monitoring the anaerobic digestion process. *Water Res*, 2010, vol. 44, pp. 5973–5980. DOI: 10.1016/j.watres.2010.07.043
3. Hill A., Tait S., Baillie C., Virdis B., McCabe B. Microbial electrochemical sensors for volatile fatty acid measurement in high strength wastewaters: A review. *Biosensors and Bioelectronics*, 2020, vol. 165, pp. 112409-112420. DOI: 10.1016/j.bios.2020.112409
4. Lyberatos G., Skiadas I.V. Modelling of anaerobic digestion - a review. *Global Nest: the Int. J*, 1999, vol. 1, no. 2, pp. 63-76.
5. Madsen M., Holm-Nielsen J.B., Esbensen K.H. Monitoring of anaerobic digestion processes: A review perspective. *Renewable and Sustainable Energy Reviews*, 2011, no. 15, pp. 3141– 3155.
6. Aamir Ishaq Shah, Mehraj U. Din Dar, Rouf Ahmad Bhat, J.P. Singh, Kuldip Singh, Shakeel Ahmad Bhat. Prospectives and challenges of wastewater treatment technologies to combat contaminants of emerging concerns. *Ecological Engineering*, 2020, vol. 152, 105882, URL: <https://doi.org/10.1016/j.ecoleng.2020.105882>
7. Растригин Л. А. Адаптация сложных систем / Л. А. Растригин – Рига: Зинатне, 1981. – 375 с.
8. Интеллектуальные технологии мониторинга и управления структурной динамикой сложных технических объектов: монография / М.Ю. Охтилев, Б.В. Соколов, Р.М. Юсупов. – М.: Наука, 2006. – 408 с.
9. Колесников А.А. Синергетика и проблемы теории управления: сборник научных трудов / А.А. Колесников. – М.: Физматлит, 2004. – 504 с.
10. Тюкин И.Ю., Терехов В.А. Адаптация в нелинейных динамических системах. СПб.: ЛКИ, 2008. – 384 с.
11. Kolesnikova S.I. Synthesis of the Control System for a Second Order Non-Linear Object with an Incomplete Description. *Automation and Remote Control*, 2018, vol. 79, no. 9, pp. 1556–1566. DOI: 10.1134/S0005117918090023
12. Batstone D., Keller J., Angelidaki I., Kalyuzhnyi S., Pavlostathis S., Rozzi A., Sanders W., Siegrist H., Vavilin V.A. The IWA anaerobic digestion model No 1. *Wat. Sci. Technol*, 2002, vol.45, no. 10, pp.65-73.
13. Колесников А.А. Метод интегральной адаптации нелинейных систем на инвариантных многообразиях / А.А. Колесников // Труды 3-ей мультikonференции по проблемам управления. – СПб., 2010. – С. 29-34.
14. Krasovskiy A.A. Problems of Control Physical Theory. *Automatic and Telemechanic*, 1990, no. 1, pp. 3-28.
15. Джонсон С. Теория регуляторов, приспособляющихся к возмущениям / Фильтрация и стохастическое управление в динамических системах / С. Джонсон; под ред. К.Т. Леондеса. – М: Мир, 1980. – 406 с.
16. Галиуллин А.С. Обратные задачи динамики и задачи управления движениями материальных систем / А.С. Галиуллин // Дифференц. уравнения, 1972. – Т. 8. – № 9. – С. 1535–1541.
17. Astolfi A., Ortega R. Immersion and invariance: A new tool for stabilization and adaptive control of nonlinear systems // *IEEE Trans. on Automatic Control*, 2003, vol. 48, no. 4. – pp. 590–605.

Фоменкова Анастасия Алексеевна, аспирант, Санкт-Петербургский государственный университет аэрокосмического приборостроения, a.a.fomenkova@mail.ru, Россия, Санкт-Петербург, ул. Большая Морская, д. 67, лит. А

Ключарев Александр Анатольевич, к.т.н., доцент, Санкт-Петербургский государственный университет аэрокосмического приборостроения, ak@aanet.ru, Россия, Санкт-Петербург, ул. Большая Морская, д. 67, лит. А

Колесникова Светлана Ивановна, д.т.н., профессор, Санкт-Петербургский государственный университет аэрокосмического приборостроения, skolesnikova@yandex.ru, Россия, Санкт-Петербург, ул. Большая Морская, д. 67, лит. А

UDC 519.7, 519.718

doi:10.38028/ESI.2022.25.1.002

Synthesis of the system of control, monitoring and evaluation of the state of an anaerobic bioreactor

Anastasia A. Fomenkova, Alexander A. Klyucharev, Svetlana I. Kolesnikova

St.Petersburg State University of Aerospace Instrumentation,
Russia, St. Petersburg, skolesnikova@yandex.ru

Abstract. The article presents a conceptual scheme for monitoring and formalizing the control action on the technical state of a complex system using an anaerobic bioreactor as an example. The mathematical apparatus and the algorithmic support that implements it are presented as the basis for the synthesis of an intelligent process control system at enterprises. Examples of analytical synthesis of an adaptive control system for a non-linear multidimensional object and an algorithm for the observer of an unmeasured variable are given.

Keywords: anaerobic biological treatment system, monitoring, non-linear control synthesis, dynamic object invariant, assessment of the state of a complex object, state observer

Acknowledgements: The work was supported by the Russian Foundation for Basic Research (grant no. 20-08-00747).

REFERENCES

1. Kalyuzhnyi S.V. Batch anaerobic digestion of glucose and its mathematical modeling. II. Description, verification and application of model. *Biores. Technol*, 1997, vol. 59, pp. 249-258.
2. Boe K., Batstone D.J., Steyer J.-P., Angelidaki I. State indicators for monitoring the anaerobic digestion process. *Water Res*, 2010, vol. 44, pp. 5973–5980. DOI: 10.1016/j.watres.2010.07.043
3. Hill A., Tait S., Baillie C., Virdis B., McCabe B. Microbial electrochemical sensors for volatile fatty acid measurement in high strength wastewaters: A review. *Biosensors and Bioelectronics*, 2020, vol. 165, pp. 112409-112420. DOI: 10.1016/j.bios.2020.112409
4. Lyberatos G., Skiadas I.V. Modelling of anaerobic digestion - a review. *Global Nest: the Int. J*, 1999, vol. 1, no. 2, pp. 63-76.
5. Madsen M., Holm-Nielsen J.B., Esbensen K.H. Monitoring of anaerobic digestion processes: A review perspective. *Renewable and Sustainable Energy Reviews*, 2011, no. 15, pp. 3141– 3155.
6. Aamir Ishaq Shah, Mehraj U. Din Dar, Rouf Ahmad Bhat, J.P. Singh, Kuldip Singh, Shakeel Ahmad Bhat. Prospectives and challenges of wastewater treatment technologies to combat contaminants of emerging concerns. *Ecological Engineering*, 2020, vol. 152, 105882, URL: <https://doi.org/10.1016/j.ecoleng.2020.105882>

7. Rastrigin L.A. Adaptatsiya slozhnykh sistem [Adaptation of complex systems]. Riga: Zinatne, 1981, 375 p.
8. Okhtilev M.Yu., Sokolov B.V., Yusupov R.M. Intel'ktual'nyye tekhnologii monitoringa i upravleniya strukturnoy dinamiko slozhnykh tekhnicheskikh ob"yektov: monografiya [Intelligent Technologies for Monitoring and Controlling the Structural Dynamics of Complex Technical Objects: monograph]. M.: Science, 2006, 408 p.
9. Kolesnikov A.A. Novye nelinejnye metody upravleniya poletom [New nonlinear methods of flight control]. M.: Fizmatlit, 2013, 193 p.
10. Tyukin I.Yu., Terekhov V.A. Adaptatsiya v nelinejnykh dinamicheskikh sistemah [Adaptation in nonlinear dynamical systems]. SPb.: LKI, 2008, 384 p. (in Russian).
11. Kolesnikova S.I. Synthesis of the Control System for a Second Order Non-Linear Object with an Incomplete Description. Automation and Remote Control, 2018, vol. 79, no. 9, pp. 1556–1566. DOI: 10.1134/S0005117918090023
12. Batstone D., Keller J., Angelidaki I., Kalyuzhnyi S., Pavlostathis S., Rozzi A., Sanders W., Siegrist H., Vavilin V.A. The IWA anaerobic digestion model. No 1. Wat. Sci. Technol, 2002, vol.45, no. 10, pp.65-73.
13. Kolesnikov A.A. Metod integral'noy adaptatsii nelineynykh sistem na invariantnykh mnogoobraziyakh [The method of integral adaptation of nonlinear systems on invariant manifolds] // Trudy 3-ey mul'tikonferentsii po problemam upravleniya = Proceedings of the 3rd multiconference on control problems. SPb, 2010, pp. 29-34
14. Krasovskiy A.A. [Problems of Control Physical Theory] // Avtomatika i telemekhanika =Automatic and Telemechanic. 1990. no. 1. Pp. 3-28.
15. Dzhonson S. Teoriya regulyatorov, prispособlivayushchikhsya k vozmushcheniyam [The theory of regulators to adapt to disturbances]. V kn. Fil'tratsiya i stokhasticheskoe upravlenie dinamicheskikh sistemakh [In the book. Filtering and stochastic control of dynamical systems], ed. by Leondes K.T. Moscow, Mir, 1980, 406 p. (in Russian).
16. Galiullin A.S. Obratnyye zadachi dinamiki i zadachi upravleniya dvizheniyami material'nykh sistem [Inverse problems of dynamics and problems of motion control of material systems] // Differents. Uravneniya = Differ. equations. 1972, vol. 8, no. 9, pp. 1535–1541.
17. Astolfi A., Ortega R. Immersion and invariance: A new tool for stabilization and adaptive control of nonlinear systems // IEEE Trans. on Automatic Control, 2003, vol. 48, no. 4, pp. 590–605.

Anastasia A. Fomenkova, postgraduate student, St.Petersburg State University of Aerospace Instrumentation, a.a.fomenkova@mail.ru, 67, Bolshaya Morskaya Str., St. Petersburg, Russia

Alexander A. Klyucharev, associate Professor, St.Petersburg State University of Aerospace Instrumentation, ak@aanet.ru, 67, Bolshaya Morskaya Str., St. Petersburg, Russia

Svetlana I. Kolesnikova, professor, St.Petersburg State University of Aerospace Instrumentation, skolesnikova@yandex.ru, 67, Bolshaya Morskaya Str., St. Petersburg, Russia,

Статья поступила в редакцию 05.03.2022; одобрена после рецензирования 25.03.2022; принята к публикации 29.03.2022.

The article was submitted 03.05.2022; approved after reviewing 03.25.2022; accepted for publication 03.29.2022.

Расчёт электродинамических характеристик и электронно-волнового взаимодействия в резонаторах гиротронов на основе комплекса программ ANGEL

Семенов Евгений Сергеевич, Зуев Андрей Сергеевич, Фокин Андрей Павлович

Институт прикладной физики РАН,

Россия, Нижний Новгород, semes@ipfran.ru

Аннотация. В статье приведены самосогласованные модели, используемые на различных этапах разработки мощных источников электромагнитного излучения – гиротронов. Представлены методы их решения, реализованные в комплексе программ ANGEL, который позволяет определить характеристики электродинамической системы, рассчитать стартовые токи и эффективность электронно-волнового взаимодействия.

Ключевые слова: математическое моделирование, неоднородное уравнение струны, задача Штурма–Лиувилля, электронно-волновое взаимодействие, стартовый ток, гиротрон

Цитирование: Семенов Е.С., Зуев А.С., Фокин А.П. Расчёт электродинамических характеристик и электронно-волнового взаимодействия в резонаторах гиротронов на основе комплекса программ ANGEL // Информационные и математические технологии в науке и управлении. – 2022. – № 1 (25). – С. 35-47. – DOI:10.38028/ESI.2022.25.1.003.

Введение. Перспективным прибором в миллиметровом и субмиллиметровом диапазоне длин волн является прибор вакуумной СВЧ-электроники – гиротрон [1, 2]. Разработка гиротрона вне зависимости от задач, которые перед ним стоят, начинается с предварительного определения его рабочих параметров. На первом шаге выбирается желаемая частота излучения, которая определяется, как правило, требованиями приложений при учёте разного рода ограничений, например, магнитной системы, источника питания и др. Так, требуемое магнитное поле для первоначальных оценок может быть найдено из условия циклотронного синхронизма

$$\omega = n\omega_H + \kappa_{\parallel}v_{\parallel},$$

где ω – частота выходного излучения, $n = 1, 2, 3, \dots$ – номер синхронной с волной гармоники гирочастоты, $\omega_H = eB/\gamma_0 m_0$ – релятивистская гирочастота электрона, e и m_0 – заряд и масса покоя электрона, $\gamma_0 = 1 + U_b^2/511^2$ – Лоренц-фактор, U_b – ускоряющее напряжение, $\kappa_{\parallel}$ – продольное (в направлении постоянного магнитного поля) волновое число, $v_{\parallel}$ – продольная скорость электронов. Требуемое магнитное поле зависит от желаемой частоты излучения, номера рабочей циклотронной гармоники и ускоряющего напряжения. Гиротрон обычно работает вблизи критической частоты ω_{co} одной из собственных мод кругового волновода TE_{mp} . Выбор рабочей моды определяет радиус регулярной части резонатора $R_{reg} \approx c v_{m,p}/\omega_{co}$, где c – скорость света, $v_{m,p}$ – p -ый корень производной функции Бесселя J_m . Гиротрон работает на одной из квазикритических мод резонатора. Далее в зависимости от мощности выходного излучения и других требований к источнику следует оптимизация основных его параметров: ускоряющего напряжения, силы тока электронного пучка, электронно-оптической системы, профиля резонатора и др.

Полноценный расчёт электронно-волнового взаимодействия с учётом всех нюансов – крайне трудоёмкая задача. Для синтеза резонаторов применяют широкий диапазон моделей, которые отличаются степенью учёта различных факторов. Некоторые из этих моделей рассмотрены в настоящей работе. На первом шаге, как правило, оптимизируется профиль резонатора гиротрона. С целью определения электродинамических параметров решается так называемая «холодная задача» – электродинамическая задача без учёта электронного пучка (см. раздел 1). Такой подход позволяет для заданной геометрии резонатора найти частоту и добротность выбранной рабочей моды и определить продольную структуру ВЧ-поля.

Рабочая мода выбирается, в том числе, с учётом проблемы конкуренции мод. Это особенно актуально в случае гиротронов, работающих на высоких циклотронных гармониках, из-за высокой конкуренции со стороны мод, синхронных с более низкими циклотронными гармониками. Для первоначального анализа условий конкуренции мод строится спектр соседних мод, расположенных вблизи рабочей моды. При этом для каждой рассматриваемой моды определяется структурный фактор связи поля с пучком (см. п. 3). Далее для анализа условий самовозбуждения рабочей моды и оценки вероятности возбуждения паразитных мод рассчитываются стартовые токи выбранной моды и соседних паразитных мод (см. раздел 3).

Системы уравнений, которые используются при моделировании электронно-волнового взаимодействия в гиротроне, известны в том или ином виде достаточно давно [2–10]. В данной работе эффективность электронно-волнового взаимодействия и мощность выходного излучения находятся в рамках стационарной самосогласованной модели, учитывающей нефиксированность продольной структуры ВЧ-поля в резонаторе гиротрона, неоднородность статического магнитного поля, разброс осцилляторных скоростей электронов и разброс ведущих центров электронных траекторий (см. раздел 4). Для анализа сценариев включения гиротрона и конкуренции мод используются нестационарные модели [3, 9]. Описанные в данной работе модели позволяют выполнить предварительную оптимизацию параметров гиротрона, используя минимальные вычислительные и временные ресурсы. Окончательную проверку работоспособности синтезированного в рамках упомянутых выше моделей резонатора вместе с электронным пучком осуществляют при помощи PIC-кодов (программные пакеты CST, KARAT и др.), которые позволяют смоделировать работу прибора с учётом максимального количества факторов, исходя из «первых принципов», однако требуют мощных вычислительных комплексов и занимают до нескольких суток расчета одной реализации.

На сегодняшний день имеется множество программ, позволяющих провести расчёт электронно-волнового взаимодействия в резонаторе гиротрона по усредненным моделям. Такими примерами являются пакеты программ MAGY [4], GYROSIM [6], EURIDICE [7], КЕДР [11] и др. В данной работе описаны модели расчёта, реализованные в комплексе программ ANGEL, который используется в ИПФ РАН и ЗАО «НПП Гиком» при расчёте гиротронов.

1. Собственные моды резонатора. Канонический резонатор гиротрона имеет азимутальную симметрию и состоит из входной секции (закритического сужения), регулярной части (чаще всего цилиндр постоянного радиуса) и выходной секции с выводом излучения в сторону коллектора. Пример профиля резонатора $R_r(z)$ показан на рис. 1. В резонаторе гиротрона формируются различные типы колебаний (моды TE_{mp}), которые характеризуются азимутальным (m), радиальным (p) и продольным (q) индексами.

Индексы $|m| > 0$, $p > 1$, $|q| > 1$ обозначают число вариаций поля по азимутальной, радиальной и продольной цилиндрическим координатам соответственно. При учёте

взаимодействия поля с электронным пучком индексам m и q приписывается знак: для мод, в которых СВЧ поле распространяется в сторону, противоположную продольному движению электронов, используется индекс $q < 0$; для мод, в которых СВЧ поле имеет азимутальное вращение в противоположную от электронов сторону, используется индекс $m < 0$. Азимутально-симметричные моды имеют индекс $m = 0$. В типичном для гиротрона случае «одногогорбой» продольной вариации ($q = 1$) индекс q часто опускается.

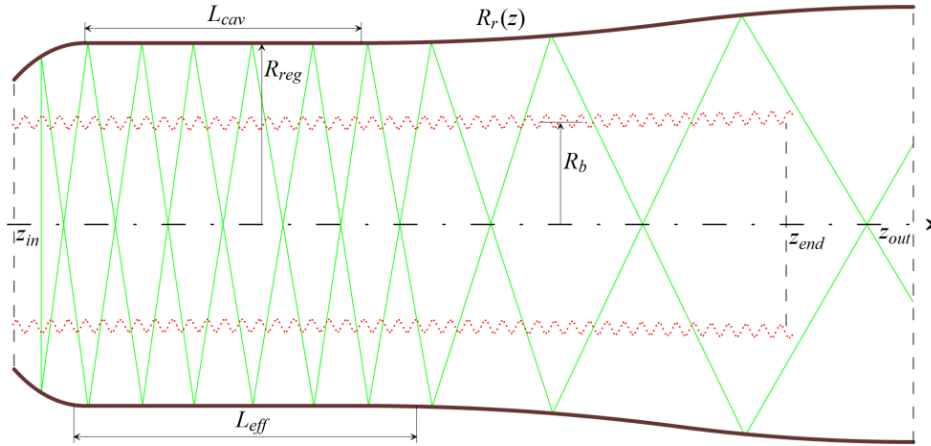


Рис. 1. Резонатор с профилем $R_r(z)$. Пунктиром показан трубчатый электронный пучок со средним радиусом R_b , тонкими линиями – схема распространения лучей электромагнитного поля

Далее полагается, что резонатор образован отрезками круглых (азимутально-симметричных) слабо нерегулярных волноводов, радиус которых близок к критическому радиусу рабочих колебаний моды TE_{mp} . В этом случае продольное распределение $f_{||}(z)$ СВЧ поля описывается уравнением неоднородной струны [12–14]

$$\frac{d^2 f_{||}}{dz^2} + \kappa_{||}^2 \cdot f_{||} = 0, \quad z \in [z_{in}, z_{out}] \quad (1)$$

с граничными условиями излучения на обоих концах [14]:

$$\frac{df_{||}}{dz} - i \kappa_{||} f_{||} = 0, \quad z = z_{in}; \quad \frac{df_{||}}{dz} + i \kappa_{||} f_{||} = 0, \quad z = z_{out}. \quad (2)$$

Условие излучения следует обеспечить на достаточно большом расстоянии от квазирегулярной части волновода, где выполняется условие ВКБ-приближения:

$$\left| \frac{d\kappa_{||}}{dz} \right| \ll |\kappa_{||}|^2.$$

Продольное волновое число $\kappa_{||}$ с учётом омических потерь запишем в виде

$$\kappa_{||}^2(z) = \kappa^2 - \kappa_{\perp}^2(z) \cdot \Omega(z),$$

где $\kappa_{\perp}(z) = v_{m,p}/R_r(z)$ – поперечное волновое число в данном поперечном сечении z ; комплексный множитель $\Omega(z)$ позволяет учесть потери энергии на нагрев резонатора. Волновое число $\kappa = \omega/c$ определяется комплексной круговой частотой

$$\omega = \omega' + i \omega'' = 2\pi f_0 \cdot \left(1 + \frac{i}{2Q}\right),$$

где f_0 – частота осцилляций СВЧ поля в «холодном» резонаторе (без учёта взаимодействия с электронным пучком), $Q = \omega'/2 \omega''$ – добротность резонатора.

Омические потери в стенках резонатора можно учесть, введя множитель $\Omega(z) = 1 +$

$\frac{i}{Q_{ohm}} \cdot \frac{R_{reg}}{R_r(z)}$ при $\kappa_{\perp}$. Здесь $Q_{ohm} = \left(1 - \frac{m^2}{v_{m,p}^2}\right) \cdot \frac{R_{reg}}{\delta_{skin}}$ – омическая добротность, $\delta_{skin} = k_{skin} \cdot \delta_{idl}$ – толщина скин-слоя, $\delta_{idl} = \sqrt{\frac{c}{\pi Z_0 \sigma \mu_r f}}$ – толщина скин-слоя гладкого металла (без

шероховатостей), $k_{skin} \approx 1.5 \dots 2$ – коэффициент потерь, учитывающий влияние микронеровностей на поверхности резонатора [15, 16].

Заметим, что задача на собственные значения (1)–(2) имеет дискретный спектр $\{\omega_q\}$ (см. рис. 2). На регулярной части типичных резонаторов собственная функция $|f_{||}(z)|$ имеет q локальных максимумов, что облегчает визуальный контроль найденных решений (см. рис. 3). Модель неоднородной струны адекватно описывает поле в резонаторе лишь при соблюдении условия Френеля [17]:

$$k_F \equiv \frac{L_q^2}{\lambda R_{reg}} \gg 1,$$

где $L_q = L/q$ — эффективная длина регулярной части резонатора с учётом количества продольных вариаций поля q .

Поиск пары (ω', ω'') , удовлетворяющей граничному условию при $z = z_{out}$, можно вести по методике, описанной в разделе 2, уточняя это решение методом Ньютона. Для интегрирования ОДУ (1) используется метод Нистрема, работающий точнее и быстрее стандартно применяемого метода Рунге—Кутты 4-го порядка.

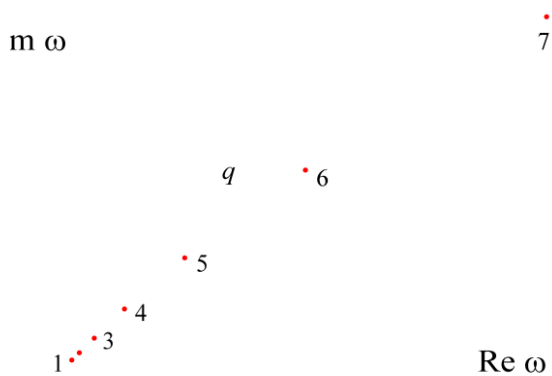


Рис. 2. Типичный спектр решений ω_q

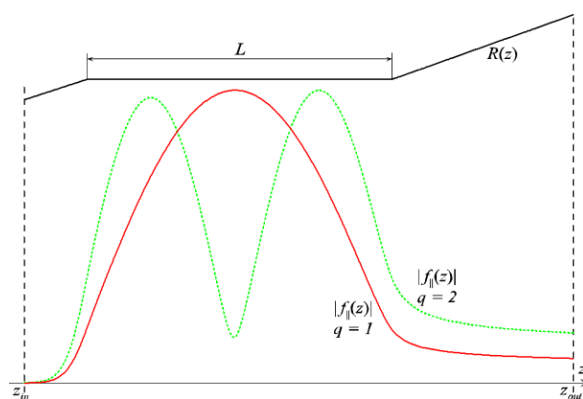


Рис. 3. Модуль продольного распределения поля для первой и второй вариаций (сплошная и пунктирная линии)

2. Метод решения задачи Штурма-Лиувилля. Задача сводится к отысканию такой комплексной частоты ω , при которой обращается в нуль функция

$$\hat{f} \equiv f'_z(z_{out}) + i \kappa_{||}(z_{out}) f_{||} = 0. \quad (3)$$

Традиционно это делается методом градиентного или покоординатного спуска после локализации минимума $|\hat{f}(\omega)|^2$ путём полного перебора с достаточно крупным шагом в области параметров, разумно выбранной из физических соображений.

Однако на эту задачу можно взглянуть, как на задачу отыскания нуля комплексной функции комплексного параметра, и применить принцип аргумента [18]: «Если $\hat{f}(\omega)$ – мероморфная функция в односвязной ограниченной области Ω и Γ – замкнутый контур, принадлежащий вместе со своей внутренностью области Ω и не проходящий ни через нули, ни через полюсы функции $\hat{f}$, то

$$\frac{1}{2\pi i} \int_{\Gamma} \frac{d\hat{f}}{\hat{f}} \equiv \frac{1}{2\pi i} \int_{\Gamma} \frac{\hat{f}' \omega}{\hat{f}} d\omega = \frac{1}{2\pi} \Delta_{\Gamma} \arg \hat{f} = N - P,$$

где N и P – число нулей и полюсов функции $\hat{f}$, лежащих внутри контура Γ с учётом их кратности (т. е. нуль или полюс порядка M считается M раз); $\Delta_{\Gamma} \arg \hat{f}$ – изменение аргумента функции $\hat{f}(\omega)$ при обходе точкой ω контура Γ в положительном направлении».

Выбрав начальное приближение $\omega_0 = \omega'_0 + i\omega''_0$ и некоторую его окрестность $\delta_0 = \delta'_0 + i\delta''_0$, $\delta'_0 > 0$, $\delta''_0 > 0$, получаем прямоугольный контур Γ_0 , натянутый на точки $(\omega'_0 \pm \delta'_0, \omega''_0 \pm \delta''_0)$. Поскольку априори известно, что полюсов у функции (3) нет, обход контура позволяет вычислить количество попавших в него нулей функции $\hat{f}$:

$$N(\omega_0, \delta_0) = \frac{1}{2\pi} |\Delta_\Gamma \arg \hat{f}|.$$

Если при обходе этого контура $N(\omega_0, \delta_0)$ окажется нулевым, следует удваивать δ_0 до тех пор, пока N не станет положительным (см. рис. 4). После этого контур рассекается пополам и по действительной, и по мнимой осям, и аналогичным образом исследуется каждый из четырёх получившихся контуров. Далее для контура с $N > 0$ проводится та же процедура разделения на четыре части; таким образом, окрестность искомого нуля функции стягивается до требуемых размеров по каждой из осей. В результате получаем координату нуля $\hat{f}(\omega^*) = 0$ с заданной точностью δ^* .

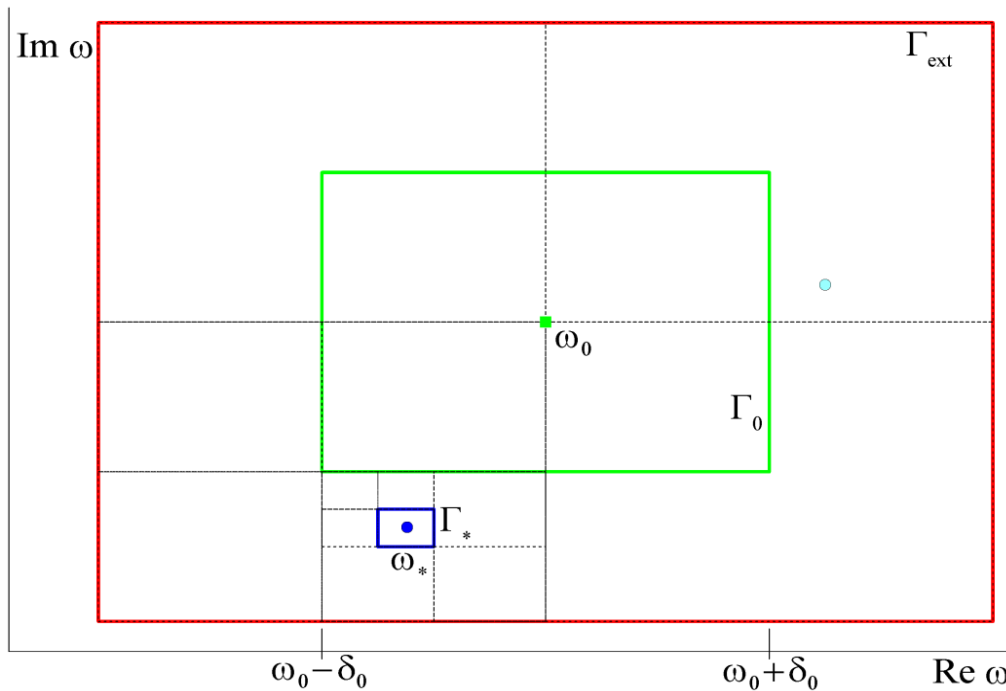


Рис. 4. Расширение и стягивание контура вокруг нуля функции

Финальный этап отгонки ω можно значительно ускорить, используя многомерный метод Ньютона, который быстро сходится, когда стартует с начальной точки ω^* , близкой к решению. Исходное комплексное уравнение $\hat{f}(\omega) = 0$ заменяется на систему нелинейных уравнений:

$$u(\eta, \xi) = 0, \quad v(\eta, \xi) = 0,$$

где $\hat{f} \equiv u + iv$, $\omega \equiv \eta + i\xi$.

Каждая итерация метода Ньютона производит переход от текущей точки (η_j, ξ_j) к следующей (η_{j+1}, ξ_{j+1}) , которая является точным решением линеаризованной системы:

$$\begin{pmatrix} \eta_{j+1} \\ \xi_{j+1} \end{pmatrix} = \begin{pmatrix} \eta_j \\ \xi_j \end{pmatrix} - G^{-1} \cdot \begin{pmatrix} u_j \\ v_j \end{pmatrix},$$

где

$$G^{-1} = \begin{pmatrix} \frac{\partial v}{\partial \xi} d^{-1} & -\frac{\partial u}{\partial \xi} d^{-1} \\ -\frac{\partial v}{\partial \eta} d^{-1} & \frac{\partial u}{\partial \eta} d^{-1} \end{pmatrix},$$

— обратная матрица Гессiana

$$G = \begin{pmatrix} \frac{\partial u}{\partial \eta} & \frac{\partial u}{\partial \xi} \\ \frac{\partial v}{\partial \eta} & \frac{\partial v}{\partial \xi} \end{pmatrix}, \quad d = \det G \equiv \frac{\partial u}{\partial \eta} \cdot \frac{\partial v}{\partial \xi} - \frac{\partial u}{\partial \xi} \cdot \frac{\partial v}{\partial \eta}$$

Функции u_j , v_j , а также все их производные, аппроксимированные конечными разностями, берутся в точке (η_j, ξ_j) .

3. Стартовые токи. Рабочую моду выбирают таким образом, чтобы в окрестностях её частоты находилось бы минимальное количество других мод, которые могут составить конкуренцию рабочей. Возможность возбуждения каждой из этих мод в первом приближении пропорциональна фактору её связи с пучком

$$G_{m,p,n} = \frac{J_{m-n}^2(v_{m,p} \cdot \frac{R_b}{R_{reg}})}{(v_{m,p}^2 - m^2) \cdot J_m^2(v_{m,p})},$$

где R_b – средний радиус ведущих центров электронных орбит (радиус встраля).

Дополнительным инструментом анализа конкуренции мод является построение стартовых токов $I_{st}(B_0)$ [2, 19]. Генерация какой-либо моды СВЧ поля в «мягком» режиме возможна, когда ток пучка I_b превосходит стартовый ток этой моды I_{st} при заданном магнитном поле B_0 . Конкуренцию, как правило, выигрывает первоначально возбуждавшаяся мода с меньшим стартовым током. Моду и радиус встраля выбирают такими, чтобы в полосу резонанса кроме рабочей моды попали лишь такие моды, стартовый ток которых превышает ток электронного пучка.

В рассматриваемой нами модели используются стационарные одномодовые самосоогласованные уравнения поля и движения электронов в релятивистском случае, в которых учтён разброс осцилляторных скоростей и разброс ведущих центров электронных траекторий. Электроны в пучке представлены фракциями: N_θ начальных фаз влёта (относительно поля), N_g групп с различными осцилляторными скоростями $v_\perp$ (но с постоянной начальной энергией пучка), N_R фракций с радиусам ведущих центров электронных орбит (для моделирования несоосности пучка и оси магнитного поля, широких пучков, многолучевых систем и т.п.), N_γ групп с различными энергиями γ_0 . Статическое магнитное поле может иметь неоднородности $B(z)$.

Совместно решаются $N_g \times N_R \times N_\gamma$ троек уравнений движения электронов

$$\begin{aligned} & \frac{d\Psi_x}{dz} + i \cdot \left(\frac{\kappa}{n} \gamma_0 - \frac{\omega_{Ho}}{c} \right) \cdot \frac{n\Psi_x - \Psi_z}{p_\parallel} + i\kappa \frac{p_\perp^2 \Psi_x + \Psi_y + 2p_\parallel^2 \Psi_z}{2n\gamma_0 p_\parallel} = \\ & = \kappa_\perp \cdot \frac{\Im_{m,p,n}}{\Re_{m,p}} \cdot \left\{ \left(i\gamma_0 F + \frac{1}{\kappa} \frac{dF}{dz} \right) \cdot \frac{p_\perp^{2n-2}}{p_\parallel} \cdot \left(\frac{J_{n-1}}{2p_\perp^{n-1}} \right) - i \frac{F p_\perp^{2n}}{2p_\parallel^2} \cdot \frac{\kappa_\perp}{\kappa} \cdot \left(\frac{J_n}{p_\perp^n} \right) \right\} + \left(n + \frac{p_\perp^2}{p_\parallel^2} \right) \cdot \frac{\Psi_x}{2B} \cdot \frac{dB}{dz}, \end{aligned} \quad (4)$$

$$\begin{aligned} & \frac{d\Psi_y}{dz} + i \cdot \left(\frac{\kappa}{n} \gamma_0 - \frac{\omega_{Ho}}{c} \right) \cdot \frac{n\Psi_y + p_\perp^2 \Psi_z}{p_\parallel} - i\kappa \frac{p_\perp^2 \Psi_x + \Psi_y + 2p_\parallel^2 \Psi_z}{2n\gamma_0 p_\parallel} p_\perp^2 = \\ & = \kappa_\perp \cdot \frac{\Im_{m,p,n}}{\Re_{m,p}} \cdot \left\{ \left(i\gamma_0 F + \frac{1}{\kappa} \frac{dF}{dz} \right) \cdot \frac{p_\perp^{2n+2}}{p_\parallel} \cdot \left(\frac{J_{n+1}}{2p_\perp^{n+1}} \right) + i \frac{F p_\perp^{2n+2}}{2p_\parallel^2} \cdot \frac{\kappa_\perp}{\kappa} \cdot \left(\frac{J_n}{p_\perp^n} \right) \right\} + \\ & + \left(n + 2 + \frac{p_\perp^2}{p_\parallel^2} \right) \cdot \frac{\Psi_y}{2B} \cdot \frac{dB}{dz}, \end{aligned} \quad (5)$$

$$\begin{aligned} & \frac{d\Psi_z}{dz} + i \cdot \left(\frac{\kappa}{n} \gamma_0 - \frac{\omega_{Ho}}{c} \right) \cdot \frac{n\Psi_z}{p_\parallel} = \\ & = -\frac{\kappa_\perp}{\kappa} \cdot \frac{dF}{dz} \cdot \frac{p_\perp^{2n}}{2p_\parallel^3} \cdot \frac{\Im_{m,p,n}}{\Re_{m,p}} \cdot \left(\frac{J'_n}{p_\perp^{n-1}} \right) + \left\{ \left(n + 3 \frac{p_\perp^2}{p_\parallel^2} \right) \cdot \Psi_z - \frac{p_\perp^2 \Psi_x + \Psi_y}{p_\parallel^2} \right\} \cdot \frac{1}{2B} \cdot \frac{dB}{dz} \end{aligned} \quad (6)$$

и уравнение для продольной структуры поля, согласованного с электронным пучком:

$$\frac{d^2 F}{dz^2} + \kappa_\parallel^2 F = I \cdot \kappa \cdot \kappa_\perp \cdot \left[\frac{\Im_{m,p,n}}{\Re_{m,p}} \cdot \left(\frac{J'_n}{p_\perp^{n-1}} \right) \cdot (n\Psi_x - \Psi_z) \right] \quad (7)$$

с начальными условиями

$$F(z_{in}) = 1, \quad \Psi_x(z_{in}) = 0, \quad \Psi_y(z_{in}) = 0, \quad \Psi_z(z_{in}) = 0 \quad (8)$$

и граничными условиями излучения

$$\frac{dF}{dz} = i\kappa_{\parallel}F, \quad z = z_{in}, \quad \frac{dF}{dz} = -i\kappa_{\parallel}F, \quad z = z_{out}. \quad (9)$$

Всюду в п. 3 и п. 4 используются следующие обозначения: $p_{\perp} = \gamma \frac{v_{\perp}}{c}$, $p_{\parallel} = \gamma \frac{v_{\parallel}}{c}$ – нормированные поперечный и продольный импульсы электронов в винтовом пучке, $\omega_{Ho} = eB/m_0$ – нерелятивистская гирочастота, $B(z)$ – внешнее неоднородное магнитное поле; $J_l \equiv J_l(\xi)$, $J'_l \equiv dJ_l(\xi)/d\xi$, $\xi = p_{\perp} \cdot \omega_{co}/\omega_{Ho}$, $\omega_{co}(z) = c\kappa_{\perp}$, $\gamma_0 = \gamma(z_{in})$. Скобки $[\dots]$ означают усреднение по всем $N_g \times N_R \times N_Y$ фракциям пучка с учётом соответствующих функций распределения. Функции $\mathfrak{J}_{m,p,n}$ и $\mathfrak{R}_{m,p}$ для TE_{mp} мод полого резонатора определены выражениями

$$\mathfrak{J}_{m,p,n}(z) = J_{m-n}(\kappa_{\perp} \cdot R_b),$$

$$\mathfrak{R}_{m,p}^2 = (v_{m,p}^2 - m^2) \cdot J_m^2(v_{m,p}).$$

Модель (4)–(9) применима, в том числе, и в случае конических или сложно профилированных резонаторов, а также может быть использована для расчёта гиротронов с коаксиальными резонаторами путем соответствующей модификации функций $\mathfrak{J}_{m,p,n}$ и $\mathfrak{R}_{m,p}$.

Уравнения (4)–(7) с условиями (8)–(9) получены линеаризацией (в предположении малой амплитуды поля) уравнений (10)–(12), приведённых в разделе 4. При этом использовались следующие обозначения:

$$\Psi_x = \frac{\langle p_c^{n-1} \cdot \delta p_c \rangle_{\theta_0}}{p_{\parallel}}, \quad \Psi_y = \frac{\langle p_c^{n+1} \cdot \delta p_c^* \rangle_{\theta_0}}{p_{\parallel}}, \quad \Psi_z = \frac{\langle p_c^n \cdot \delta p_{\parallel} \rangle_{\theta_0}}{p_{\parallel}^2}.$$

Здесь скобки $\langle \dots \rangle_{\theta_0}$ означают усреднение импульсов электронов по фазе влёта θ_0 . Поперечная $p_{c,tot}$ и продольная $p_{\parallel,tot}$ компоненты импульса разделены на две части: основные части, зависящие лишь от неоднородного статического магнитного поля (p_c , $p_{\parallel}$), и малые поправки (δp_c , $\delta p_{\parallel}$), получаемые частицами при взаимодействии с СВЧ полем:

$$p_{c,tot}(z) = p_c(z) + \delta p_c(z), \quad |\delta p_c| \ll |p_c| \forall z, \quad \delta p_c(z_{in}) = 0,$$

$$p_{\parallel,tot}(z) = p_{\parallel}(z) + \delta p_{\parallel}(z), \quad |\delta p_{\parallel}| \ll |p_{\parallel}| \forall z, \quad \delta p_{\parallel}(z_{in}) = 0.$$

Распределение поперечной $p_{\perp}(z) \equiv |p_c(z)|$ и продольной компонент импульса вдоль оси находится из системы ОДУ для каждой скоростной и энергетической фракции:

$$\begin{cases} \frac{dp_{\perp}}{dz} = \frac{p_{\perp}}{2B} \frac{dB}{dz}, \\ \frac{dp_{\parallel}}{dz} = -\frac{p_{\perp}^2}{p_{\parallel} \cdot 2B} \frac{dB}{dz} \end{cases}$$

с начальными импульсами $p_{\perp,in} = p_0 g_0 / \sqrt{1 + g_0^2}$, $p_{\parallel}(z_{in}) = \sqrt{\gamma_0^2 - 1 - p_{\perp,in}^2}$, $p_0 = \sqrt{\gamma_0^2 - 1}$,

где $g_0 = v_{\perp}/v_{\parallel}$ – питч-фактор данной электронной фракции.

Собственными значениями задачи (4)–(9) являются частота ω (определяющая волновой вектор $\kappa = \omega/c$) и стартовый ток I_{st} , из которого выражается нормированный ток

$$I = \frac{8e_0}{c m_0} \cdot 10^{-7} \cdot I_{st}.$$

Поиск пары (ω, I_{st}) , удовлетворяющей граничному условию (9) при $z = z_{out}$, можно вести по методике, аналогичной описанной в п. 2, уточняя это решение методом Ньютона. Система уравнений движения и поля интегрируются методом Рунге-Кутты 4-го порядка.

4. Расчёт КПД гиротрона. В данном разделе приводятся стационарные одномодовые самосогласованные уравнения поля и движения электронов. Совместно решаются $N_{\theta} \times N_g \times N_R \times N_Y$ пар уравнений движения электронов [5, 10]

$$\frac{dp_c}{dz} + i \frac{p_c}{p_{\parallel}} \left(\frac{\kappa}{n} \gamma - \frac{\omega_{Ho}}{c} \right) = \kappa_{\perp} \frac{\mathfrak{J}_{m,p,n}}{\mathfrak{R}_{m,p}} \cdot \left\{ \left(\frac{i\gamma F}{p_{\parallel}} + \frac{1}{\kappa} \frac{dF}{dz} \right) \cdot (p_c^*)^{n-1} \cdot \left(\frac{J_{n-1}}{2p_{\perp}^{n-1}} \right) + \right. \quad (10)$$

$$+ \left(\frac{i\gamma F^*}{p_{\parallel}} - \frac{1}{\kappa} \frac{dF^*}{dz} \right) \cdot p_c^{n+1} \cdot \left(\frac{J_{n+1}}{2p_{\perp}^{n+1}} \right) - ip_c \cdot \operatorname{Re} \left(\frac{1}{\kappa} \frac{dF}{dz} \cdot (p_c^*)^n \right) \cdot \left(\frac{J_n}{p_{\perp}^n} \cdot \frac{\kappa_{\perp}}{\kappa} \right) \left\} + \frac{p_c}{2B} \frac{dB}{dz},$$

$$\frac{dp_{\parallel}}{dz} = -\operatorname{Re} \left(\frac{1}{\kappa} \frac{dF}{dz} \cdot (p_c^*)^n \right) \cdot \left(\frac{J_n'}{p_{\perp}^{n-1}} \right) \cdot \frac{\kappa_{\perp}}{p_{\parallel}} \cdot \frac{\Im_{m,p,n}}{\Re_{m,p}} - \frac{p_{\perp}^2}{p_{\parallel} \cdot 2B} \frac{dB}{dz} \quad (11)$$

и уравнение для продольной структуры поля, согласованного с электронным пучком:

$$\frac{d^2 F}{dz^2} + \kappa_{\parallel}^2 F = \frac{l}{\Re_{m,p}} \cdot \kappa \cdot \kappa_{\perp} \cdot \left[\Im_{m,p,n} \cdot \frac{p_c^n}{p_{\parallel}} \cdot \left(\frac{J_n'}{p_{\perp}^{n-1}} \right) \right]. \quad (12)$$

с начальными условиями

$$F(z_{in}) = F_{in}, \quad p_c(z_{in}) = p_{\perp,in} \cdot e^{i\theta_0/n}, \quad p_{\parallel}(z_{in}) = p_{\parallel,in} \quad (13)$$

и граничными условиями излучения

$$\frac{dF}{dz} = i\kappa_{\parallel} F, \quad z = z_{in}, \quad \frac{dF}{dz} = -i\kappa_{\parallel} F, \quad z = z_{out}. \quad (14)$$

Все обозначения и начальные значения импульсов см. в разделе 3.

Поиск пары (ω, F_{in}) , удовлетворяющей граничному условию (14) при $z = z_{out}$, можно вести по методике, аналогичной описанной в разделе 2, уточняя это решение методом Ньютона. Система уравнений движения и поля интегрируются методом Рунге-Кутты 4-го порядка. При этом волновой КПД гиротрона определяется по формуле

$$\eta_{rf} = -\operatorname{Im} \left\{ F^* \frac{dF}{dz} \right\} \frac{1}{\kappa l (\gamma_0 - 1)}, \quad z \in [z_{in}, z_{out}].$$

В качестве примера рассмотрен технологический гиротрон, работающий на второй гармонике циклотронной частоты на моде $TE_{0,2}$ с частотой излучения 28 ГГц. На рис. 5 показан скриншот графического окна программы ANGEL, демонстрирующий профиль резонатора, продольную структуру поля и волновой КПД в зависимости от продольной координаты z .

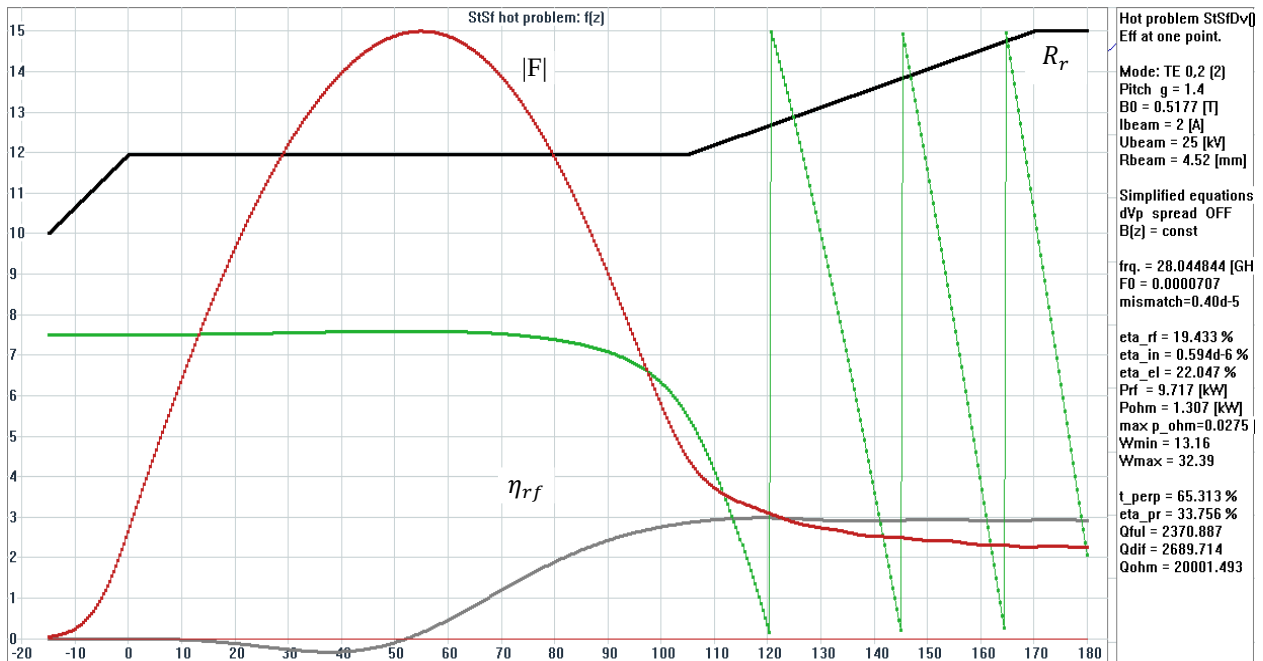


Рис. 5. Продольная структура поля (амплитуда и фаза) в резонаторе, согласованная с пучком; зависимость КПД от продольной координаты

Результаты расчётов стартового тока и волнового КПД технологического гиротрона в рабочей полосе приведены на рис. 6, где линия стартового тока отсекает типичную для всех гиротронов зону жёсткого возбуждения рабочей моды.

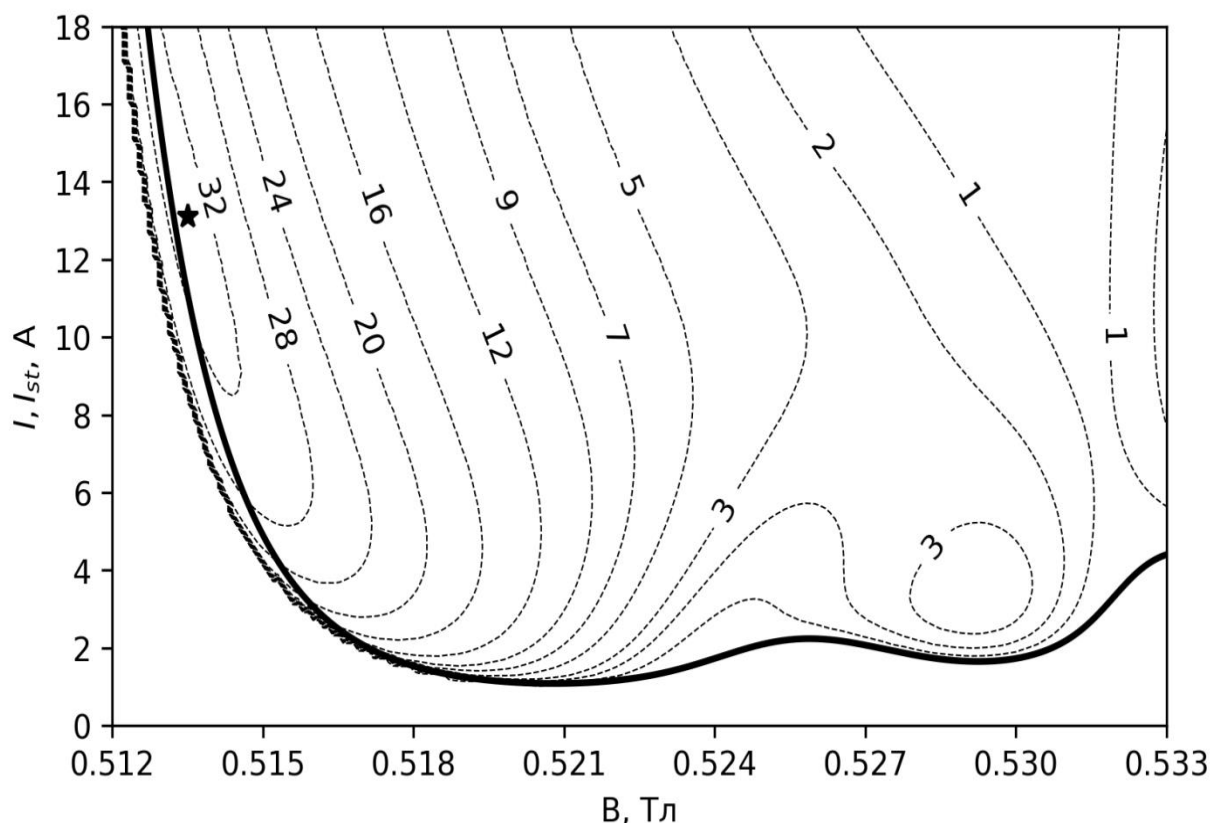


Рис. 6. Изолинии КПД (в %) от магнитного поля и тока пучка. Сплошной линией показан стартовый ток, полученный по математической модели из раздела 3

Закключение. В работе представлены эффективные математические модели, используемые на различных этапах разработки гиротронов, реализованные в комплексе программ ANGEL. Приведённые модели легко обобщаются на случай расчёта электронно-волнового взаимодействия в гиротронах с коаксиальными резонаторами. Результаты расчётов по этим моделям прошли неоднократную проверку как в сравнении с расчётами в других программах, так и с экспериментальными данными гиротронов. В настоящее время комплекс программ ANGEL активно используется в ИПФ РАН и ЗАО «НПП Гиком» при разработке гиротронов суб-ТГц и ТГц диапазона частот.

Благодарности. Авторы благодарны А.С. Сергееву за неоценимые практические советы, М.Ю. Глявину, В.Е. Запелову, М.Д. Проявину и А.С. Седову за ценные замечания и внимание к работе. Работа поддержана грантом Президента РФ №МК-3413.2021.1.2.

СПИСОК ИСТОЧНИКОВ

1. Flyagin V.A., Gaponov A.V., Petelin I., Yulpatov V.K. The gyrotron. IEEE Transactions on Microwave Theory and Techniques. 1977, vol. 25, no. 6, pp. 514–521, DOI: 10.1109/TMTT.1977.1129149.
2. Nusinovich G.S. Introduction to the Physics of Gyrotrons. Baltimore: The Johns Hopkins University, 2004, 336 p.
3. Гинзбург Н.С., Завольский Н.А., Нусинович Г.С., Сергеев А.С. Установление автоколебаний в электронных СВЧ генераторах с дифракционным выводом излучения / Н.С. Гинзбург, Н.А. Завольский, Г.С. Нусинович // Изв. вузов. Радиофизика, 1986. – Т. 29. – № 1. – С. 106–114.

4. Botton M., Antonsen T.M., Levush B., Nguyen K.T., Vlasov A.N. MAGY: a time-dependent code for simulation of slow and fast microwave sources. *IEEE Trans, Plasma Sci*, 1998, vol. 26, no. 3, pp. 882–892, DOI: 10.1109/27.700860.
5. Moiseev M.A., Zapevalov V.E., Zavolsky N.A. Efficiency enhancement of the relativistic gyrotron. *Int. J. Infrared Millimeter Waves*. 2001, vol. 22, no. 6, pp. 813–833, DOI: 10.1023/A:1014954012067
6. Sabchevski S., Idehara T., Saito T., Ogawa I., Mitsudo S., Tatematsu Y. Physical Models and Computer Codes of the GYROSIM (GYROtron SIMulation) Software Package. FIR Center Report FIR FU-99, 2010, URL: http://fir.u-fukui.ac.jp/FIR_FU99S.pdf
7. Avramides K.A., Pagonakis I.Gr., Iatrou C.T., Vomvroidis J.L. EURIDICE: a code-package for gyrotron interaction simulations and cavity design. *EPJ Web of Conferences*, 2012, vol. 32, p. 04016, DOI: 10.1051/epjconf/20123204016
8. Wang P., Chen X., Xiao H., Dumbrajs O., Qi X., Li L. GYROCOMPU: toolbox designed for the analysis of gyrotron resonators. *IEEE Trans. Plasma Sci*. 2020, vol. 48, no. 9, pp. 3007–3016, DOI: 10.1109/TPS.2020.3013299
9. Semenov E., Zapevalov V., Zuev A. Methods for Simulation the Nonlinear Dynamics of Gyrotrons. *Communications in Computer and Information Science*. 2021, vol. 1413, pp. 49-62, DOI: 10.1007/978-3-030-78759-2_4
10. Завольский Н.А., Запёвалов В.Е., Моисеев М.А. Численное моделирование процессов электронно-волнового взаимодействия в резонаторах мощных гиротронов с частотой 300 ГГц / Н.А. Завольский, В.Е. Запёвалов, М.А. Моисеев М.А. // *Изв. вузов. Радиофизика*, 2021. – Т. 64. – № 3. – С. 192–205.
11. Колосов С.В., Зайцева И.Е. Компьютерная программа Гуго-К для разработки и проектирования гирорезонансных приборов / С.В. Колосов, И.Е. Зайцева // *СВЧ электроника*, 2017. – Т. 2. – С. 44-46.
12. Вайнштейн Л.А. Электромагнитные волны / Л.А. Вайнштейн. – М. Радио и связь, 1988. – 440 с.
13. Каценеленбаум В.З. Теория нерегулярных волноводов с медленно меняющимися параметрами / В.З. Каценеленбаум. – М.: изд-во АН СССР, 1961. – 216 с.
14. Власов С.Н., Жислин Г.М., Орлова И.М., Петелин М.И., Рогачева Г.Г. Открытые резонаторы в виде волноводов переменного сечения / С.Н. Власов, Г.М. Жислин, И.М. Орлова [и др.]// *Известия вузов. Радиофизика*, 1969. – Т. 12. – № 8. – С. 1236–1244.
15. Глявин М.Ю., Морозкин М.В., Цветков А.И., Лубяко Л.В., Голубятников Г.Ю., Куфтин А.Н., Запёвалов В.Е., Холопцев В.В., Еремеев А.Г., Седов А.С., Малыгин В.И., Чирков А.В., Фокин А.П., Соколов Е.В., Денисов Г.Г. Автоматизированный микроволновый комплекс на основе работающего в непрерывном режиме гиротрона с рабочей частотой 263 ГГц и выходной мощностью 1 кВт / М.Ю. Глявин, М.В. Морозкин, А.И. Цветков [и др.] // *Изв. вузов. Радиофизика*, 2015. – Т. 58. – № 9. – С. 709–719.
16. Запёвалов В.Е., Зуев А.С., Паршин В.В., Семенов Е.С., Серов Е.А. Снижение омических потерь в резонаторах терагерцовых гиротронов малой мощности / В.Е. Запёвалов, А.С. Зуев, В.В. Паршин [и др.] // *Изв. вузов. Радиофизика*, 2021. – Т. 64. – № 4. – С. 265–275.
17. Nusinovich G.S., Dumbrajs O. Field Formation in the Interaction Space of Gyrotrons. *J. Infrared, Millimeter and Terahertz Waves*. 2016, vol. 37, no. 1, pp. 111–122, DOI: 10.1007/s10762-015-0192-2

18. Корн Г., Корн Т. Справочник по математике для научных работников и инженеров / Г. Корн, Т. Корн. – М.: Наука, 1978. – 832 с.
19. Цимринг Ш.Е. Введение в высокочастотную вакуумную электронику и физику электронных пучков / Ш.Е. Цимринг. – Нижний Новгород: ИПФ РАН, 2012. – 576 с.

Семенов Евгений Сергеевич, к.ф.-м.н., н.с., Институт прикладной физики РАН, semes@ipfran.ru, Россия, г. Нижний Новгород, ул. Ульянова, 46

Зуев Андрей Сергеевич, м.н.с., Институт прикладной физики РАН, andrey.zuev@ipfran.ru, Россия, г. Нижний Новгород, ул. Ульянова, 46

Фокин Андрей Павлович, к.ф.-м.н., с.н.с., Институт прикладной физики РАН, fokin@ipfran.ru, Россия, г. Нижний Новгород, ул. Ульянова, 46

UDC 621.385.69

doi:10.38028/ESI.2022.25.1.003

Calculation of electrodynamic characteristics and electron-wave interaction in gyrotron resonators based on the ANGEL software package

Evgeny S.Semenov, Andrey S. Zuev, Andrey P. Fokin

Institute of Applied Physics RAS, Russia, Nizhny Novgorod, semes@ipfran.ru

Abstract. The article presents self-consistent models used at various stages of development of powerful sources of electromagnetic radiation - gyrotrons. Methods for solving them are presented, implemented in the ANGEL software package, which allows determining the characteristics of the electrodynamic system, calculating the starting currents and the efficiency of the electron-wave interaction.

Keywords: mathematical modeling, inhomogeneous string equation, Sturm–Liouville problem, electron-wave interaction, starting current, gyrotron

Acknowledgements: The authors are grateful to A.S. Sergeev for invaluable practical advice, M.Yu. Glyavin, V.E. Zapevalov, M.D. Proyavin and A.S. Sedov for valuable comments and attention to the work. This work was supported by a grant from the President of the Russian Federation No. MK-3413.2021.1.2.

REFERENCES

1. Flyagin V.A., Gaponov A.V., Petelin I., Yulpatov V.K. The gyrotron. IEEE Transactions on Microwave Theory and Techniques. 1977, vol. 25, no. 6, pp. 514–521, DOI: 10.1109/TMTT.1977.1129149.
2. Nusinovich G.S. Introduction to the Physics of Gyrotrons. Baltimore: The Johns Hopkins University, 2004, 336 p.
3. Ginzburg N.S., Zavol'skiy N.A., Nusinovich G.S., Sergeyev A.S. Ustanovleniye avtokolebaniy v elektronnykh SVCH generatorakh s difraktsionnym vyvodom izlucheniya [Establishment of self-oscillations in electronic microwave generators with diffraction output of radiation]. Izv. vuzov. Radiofizika [Izv. universities. Radiophysics]. 1986, v. 29, no. 1, pp. 106–114.

4. Botton M., Antonsen T.M., Levush B., Nguyen K.T., Vlasov A.N. MAGY: a time-dependent code for simulation of slow and fast microwave sources. *IEEE Trans, Plasma Sci*, 1998, vol. 26, no. 3, pp. 882–892, DOI: 10.1109/27.700860.
5. Moiseev M.A., Zapevalov V.E., Zavolsky N.A. Efficiency enhancement of the relativistic gyrotron. *Int. J. Infrared Millimeter Waves*. 2001, vol. 22, no. 6, pp. 813–833, DOI: 10.1023/A:1014954012067
6. Sabchevski S., Idehara T., Saito T., Ogawa I., Mitsudo S., Tatematsu Y. Physical Models and Computer Codes of the GYROSIM (GYROtron SIMulation) Software Package. FIR Center Report FIR FU-99, 2010, URL: http://fir.u-fukui.ac.jp/FIR_FU99S.pdf
7. Avramides K.A., Pagonakis I.Gr., Iatrou C.T., Vomvoridis J.L. EURIDICE: a code-package for gyrotron interaction simulations and cavity design. *EPJ Web of Conferences*, 2012, vol. 32, p. 04016, DOI: 10.1051/epjconf/20123204016
8. Wang P., Chen X., Xiao H., Dumbrajs O., Qi X., Li L. GYROCOMPU: toolbox designed for the analysis of gyrotron resonators. *IEEE Trans. Plasma Sci*. 2020, vol. 48, no. 9, pp. 3007–3016, DOI: 10.1109/TPS.2020.3013299
9. Semenov E., Zapevalov V., Zuev A. Methods for Simulation the Nonlinear Dynamics of Gyrotrons. *Communications in Computer and Information Science*. 2021, vol. 1413, pp. 49-62, DOI: 10.1007/978-3-030-78759-2_4
10. Zavol'skiy N.A., Zapevalov V.Ye., Moiseyev M.A. Chislennoye modelirovaniye protsessov elektronno-volnovogo vzaimodeystviya v rezonatorakh moshchnykh girotronov s chastotoy 300 GGts [Numerical modeling of the processes of electron-wave interaction in the resonators of high-power gyrotrons with a frequency of 300 GHz]. *Izv. vuzov. Radiofizika [Izv. universities. Radiophysics]*. 2021, v. 64, no. 3, pp. 192–205.
11. Kolosov S.V., Zaytseva I.Ye. Komp'yuternaya programma Gyro-K dlya razrabotki i proyektirovaniya girorezonansnykh priborov [Computer program Gyro-K for the development and design of gyroresonance devices]. *SVCH elektronika [Microwave electronics]*. 2017, v. 2, pp. 44-46.
12. Vaynshteyn L.A. Elektromagnitnyye volny [Electromagnetic waves]. *Radio i svyaz' [Radio and communication]*, 1988, 440 p.
13. Katsenelenbaum V.Z. Teoriya neregulyarnykh volnovodov s medlenno menyayushchimisya para-metrami [Theory of irregular waveguides with slowly changing parameters]. *Izd-vo AN SSSR*, 1961, 216 p.
14. Vlasov S.N., Zhislin G.M., Orlova I.M., Petelin M.I., Rogacheva G.G. Otkrytyye rezonatory v vide volnovodov peremennogo secheniya [Open resonators in the form of variable section waveguides]. *Izvestiya vuzov. Radiofizika [Izvestiya vuzov. Radiophysics]*. 1969, v. 12, no. 8, pp. 1236–1244.
15. Glyavin M.YU., Morozkin M.V., Tsvetkov A.I., Lubyako L.V., Golubyatnikov G.YU., Kuftin A.N., Zapevalov V.Ye., Kholoptsev V.V., Yermeyev A.G., Sedov A.S., Malygin V.I., Chirkov A.V., Fokin A.P., Sokolov Ye.V., Denisov G.G. Avtomatizirovanny mikrovolnovyy kompleks na osnove rabotayushchego v nepreryvnom rezhime girotrona s rabochey chastotoy 263 GGts i vykhodnoy moshchnost'yu 1 kVt [Automated microwave complex based on a gyrotron operating in a continuous mode with an operating frequency of 263 GHz and an output power of 1 kW]. *Izvestiya vuzov. Radiofizika [Izvestiya vuzov. Radiophysics]*. 2015, v. 58, no. 9, pp. 709–719.
16. Zapevalov V.Ye., Zuyev A.S., Parshin V.V., Semenov Ye.S., Serov Ye.A. Snizheniye omicheskikh poter' v rezonatorakh teragertsovykh girotronov maloy moshchnosti [Reduction of ohmic losses in the resonators of low-power terahertz gyrotrons]. *Izvestiya vuzov. Radiofizika [Izvestiya vuzov. Radiophysics]*. 2021, v. 64, no. 4, pp. 265–275.

17. Nusinovich G.S., Dumbrajs O. Field Formation in the Interaction Space of Gyrotrons. J. Infrared, Millimeter and Terahertz Waves. 2016, vol. 37, no. 1, pp. 111–122, DOI: 10.1007/s10762-015-0192-2
18. Korn G., Korn T. Spravochnik po matematike dlya nauchnykh rabotnikov i inzhenerov [Handbook of mathematics for scientists and engineers]. Moskva, Nauka, 1978, 832p.
19. Tsimring SH.Ye. Vvedeniye v vysokochastotnuyu vakuumnuyu elektroniku i fiziku elektronnykh puchkov [Introduction to high-frequency vacuum electronics and electron beam physics]. Nizhniy Novgorod: IPF RAN, 2012, 576p.

Evgeny S.Semenov, Candidate of Physical and Mathematical Sciences, Researcher, Institute of Applied Physics RAS, semes@ipfran.ru, Russia, Nizhny Novgorod, st. Ulyanova, 46

Andrey S. Zuev, junior researcher, Institute of Applied Physics RAS, andrey.zuev@ipfran.ru, Russia, Nizhny Novgorod, st. Ulyanova, 46

Andrey P. Fokin, Candidate of Physical and Mathematical Sciences, Senior Researcher, Institute of Applied Physics RAS, fokin@ipfran.ru, Russia, Nizhny Novgorod, st. Ulyanova, 46

Статья поступила в редакцию 15.03.2022; одобрена после рецензирования 28.03.2022; принята к публикации 30.03.2022.

The article was submitted 03.15.2022; approved after reviewing 03.28.2022; accepted for publication 03.30.2022

Методы искусственного интеллекта

УДК 519.6

doi:10.38028/ESI.2022.25.1.004

Типовые структуры популяционных алгоритмов глобальной оптимизации

Карпенко Анатолий Павлович

МГТУ им. Н.Э. Баумана,

Россия, Москва, apkarpenko@mail.ru

Аннотация. Представляем авторскую классификацию структур популяционных алгоритмов глобальной оптимизации (П-алгоритмов). На уровне популяции выделяем следующие структурные классы этих алгоритмов: одноуровневые одно популяционные; одноуровневые много популяционные; двух уровневые; трех уровневые. На следующем уровне иерархии различаем одностадийные (два макрооператора), двух стадийные (два макрооператора) и трех и более стадийные (три и более макрооператоров) алгоритмы. С другой стороны, подразделяем П-алгоритмы на стационарные, программно-динамические, параметрически и структурно-адаптивные динамические алгоритмы. Рассматриваем только одно популяционные алгоритмы.

Ключевые слова: глобальная оптимизация; популяционные алгоритмы; метаэвристические алгоритмы; алгоритмы, вдохновленные природой

Цитирование: Карпенко А.П. Типовые структуры популяционных алгоритмов глобальной оптимизации // Информационные и математические технологии в науке и управлении. – 2022. – № 1 (25). – С. 48-57. – DOI:10.38028/ESI.2022.25.1.004.

Введение. Рассматриваемые алгоритмы глобальной (непрерывной) оптимизации относятся к классу прямых (не использующих производные целевой функции) стохастических поисковых алгоритмов, которые в разных публикациях называют поведенческими, интеллектуальными, метаэвристическими, вдохновленными (инспирированными) природой, роевыми, многоагентными, популяционными и т.д. Вслед за [1] используем термин “популяционные алгоритмы” (П-алгоритмы). Эти алгоритмы (population based algorithms) предполагают одновременную обработку нескольких вариантов решения задачи оптимизации и представляют собой альтернативу классическим траекторным поисковым алгоритмам, в которых в области поиска эволюционирует только один кандидат в решение этой задачи. Все П-алгоритмы относятся к классу эвристических алгоритмов (heuristic algorithms), то есть алгоритмов, для которых сходимость к глобальному решению не доказана, но экспериментально установлено, что в большинстве случаев они дают достаточно «хорошее» решение.

Аналогично [1], рассматриваем детерминированные стационарные задачи оптимизации. Первый термин предполагает детерминированность целевой функции задачи оптимизации, а второй – неизменность этой функции и области ее допустимых значений во времени (в процессе решения задачи). Отметим, что многие из П-алгоритмов могут быть модифицированы (и такие модификации известны) для решения динамических задач оптимизации.

Информационной «базой данных» для статьи является обзор наиболее известных П-алгоритмов, представленный в книге [2], которая содержит описание 134 популяционных алгоритмов (как непрерывных, так и дискретных), наиболее известных на момент ее опубликования (2014 г.). Из числа представленных в [2] алгоритмов в статью не вошли, прежде всего,

дискретные алгоритмы. Кроме того, в ней не нашли место немногие другие алгоритмы, степень детализации которых оказалась недостаточной по сравнению с другими алгоритмами (представленными как в [2], так и в исходных публикациях). По этой причине в статье не представлен, например, алгоритм голых землекопов (Blind Naked Mole-Rats, BNMR) [3].

Обзоры П-алгоритмов различного уровня широты, а также различные аспекты анализа и синтеза П-алгоритмов рассмотрены, например, в публикациях [4-13].

Говоря об итерационных формулах рассматриваемых П-алгоритмов, выделяем их следующие базовые составляющие: аттракторная (attractor, A); когнитивная (cognitive, C); историческая (historical, H); инерционная (inertial, I); феромонная (pheromone, Ph); случайная (random, R); социальная (social, S). Кроме того, в алгоритмах глобального поиска могут использоваться такие комбинированные составляющие, как когнитивно-социальная (cognitive-social, CS), историко-социальная (historical-social, HS) и другие.

1. Стационарные алгоритмы. В стационарных алгоритмах в процессе эволюции популяции не изменяются структура алгоритмов, а также значения их параметров.

1.1. Одностадийные алгоритмы. Итерационные формулы одностадийных алгоритмов могут содержать одну или более указанных составляющих. Например, в алгоритме поиска системой зарядов (CSS) каждую из особей $s_i \in S$ перемещаем в пространствах скоростей $\mathfrak{R}_{\Delta x}$ и положений $\mathfrak{R}_x$ по формулам, содержащим инерциальную и социальную составляющие. В алгоритме оптимизации гравитационных взаимодействий (GIO) положения и скорости каждой из особей $s_i \in S$ обновляем по итерационным формулам, включающим в себя инерционную и историко-социальную компоненты. Аналогично, в алгоритме гравитационного поиска (GSA) и алгоритме, моделирующем движение частиц в магнитном поле (MOA), используются инерционная и социальная составляющие. В алгоритме оптимизации роем частиц (PSO) особей $s_i \in S$ перемещаем в пространстве поиска по трех составной формуле, включающей в себя когнитивную, инерциальную и социальную составляющую. Спиральный оптимизационный алгоритм (SOA) использует однокомпонентной формулу, включающую в себя только социальную составляющую.

1.2. Двух стадийные алгоритмы. Известны двух стадийные П-алгоритмы, содержанием стадий которых являются: глобальный поиск 1 + глобальный поиск 2 (алгоритм опыления цветов (FPA)); глобальный поиск + формирование новой популяции (алгоритм сорняковой оптимизации (IWO)); скрещивание особей + мутация (алгоритм на основе модели биогеографии (BBO)); глобальный поиск + удовлетворение ограничений (алгоритм интеллектуальной физической оптимизации (APO)); сжатие популяции + расширение популяции (алгоритм большого взрыва – большого сжатия (BB-BC)); локальный поиск + глобальный поиск (электромагнитный алгоритм (EMA)); глобальный поиск + скрещивание (алгоритм рассеянного поиска (SS)); сортировка особей + глобальный поиск (гусиный алгоритм оптимизации (GOA)).

Отметим, что алгоритмы, схема которых имеет вид «локальный поиск + глобальный поиск» и «глобальный поиск + локальный поиск», различны (даже если речь идет об одинаковых алгоритмах локального и глобального поиска соответственно).

Приведем два более детальных примера.

Алгоритм эволюционной стратегии (ES):

1. Мутация – всех особей $s_i \in S$ подвергаем гауссовой мутации: $S \rightarrow S'$;
2. формирование новой популяции – объединяем популяции S, S' в популяцию S'' , выбираем из популяции S'' наиболее приспособленных особей числом $|S|$ и помещаем их в популяцию $S(t+1)$: $S \cup S' \rightarrow S'' \rightarrow S(t+1)$.

Гибридный алгоритм непрерывно взаимодействующей муравьиной колонии (НСИАС):

1. глобальный поиск – если выбран канал стигмергии, то для каждой из особей $s_i \in S$ по схеме, близкой соответствующему фрагменту схемы алгоритма СИАС, определяем пробную точку $\tilde{X}_i$ и, если $\varphi(\tilde{X}_i) > \varphi(X_i)$, то полагаем $X'_i = \tilde{X}_i$; в случае, если выбран канал прямого взаимодействия, то для каждой из особей $s_i \in S$ также по схеме алгоритма СИАС выбираем из памяти этой особи сообщение от случайной особи $s_{i_j} \in S$, $i_j \neq i$ и в зависимости от значений величин $\varphi(s_{i_j})$, $\varphi(s_i)$ перемещаем особь s_i в ту или иную случайную точку $X'_i: S \rightarrow S'$;
2. локальный поиск – в окрестности текущего положения каждой из особей $s'_i \in S'$ осуществляем локальный поиск с помощью алгоритма Нелдера-Мида: $S' \rightarrow S(t+1)$.

1.3. Трех стадийные алгоритмы. Известны трех стадийные П-алгоритмы со следующим содержанием стадий: глобальный поиск + локальный поиск + «встряска» популяции (алгоритм оптимизации роем тараканов (CSO)); глобальный поиск + локальный поиск + перемещение особей (алгоритм летучих мышей (BaA)); локальный поиск 1 + локальный поиск 2 + глобальный поиск (обезьяний алгоритм (MA)); локальный поиск + глобальный поиск + «встряска» популяции (алгоритм кукушкиного поиска (CSA)); сжатие популяции + локальный поиск + формирование новой популяции (алгоритм рисового поля (PFA)); глобальный поиск 1 + глобальный поиск 2 + глобальный поиск 3 (поисковый алгоритм искусственного роя (ASSA)); локальный поиск + переформирование популяции + «встряска» популяции (простой алгоритм эволюции разума (SMEC)); одноточечная локальная оптимизация + многоточечная локальная оптимизация + «встряска» популяции (алгоритм колонии искусственных пчел (ABC)).

Ниже в качестве примера представлены схемы двух трех стадийных стационарных П-алгоритмов.

Генетический алгоритм:

1. селекция и скрещивание – методом панмиксии выбираем из популяции S пары родительских особей заданным общим числом, и к каждой из пар $\{s_{j_1}, s_{j_2}\}$ отобранных особей применяем арифметический оператор скрещивания: $S \rightarrow S'$;
2. мутация – особей набора S' подвергаем мутации с помощью оператора случайной мутации: $S' \rightarrow S''$;
3. селекция – на основе популяции S и набора S'' формируем методом панмиксии популяцию $S(t+1)$: $\{S, S''\} \rightarrow S(t+1)$.

Алгоритм хемотаксиса колонии бактерий (BCC):

1. глобальный поиск 1 – для каждой из особей $s_i \in S$ на основе их предыдущих направлений перемещения находим новые возможные направления перемещения в пространстве сферических координат R_A , пересчитываем эти направления в направления пространства R_X и на этой основе определяем новые (первые) возможные положения этих особей: $S \rightarrow \tilde{S}'$;
2. глобальный поиск 2 – для всякой особи $s_i \in S$ в R_X -окрестности ее текущего положения находим лучшую соседнюю особь s_{i_j} и определяем в ее направлении второе возможное положение особи s_i : $S \rightarrow \tilde{S}''$;

3. перемещение особей – перемещаем особь S_i в лучшую из точек $\tilde{X}'_i, \tilde{X}''_i$: $\{\tilde{S}', \tilde{S}''\} \rightarrow S(t+1)$.

1.4. Четырех и более стадийные алгоритмы. Известно небольшое число стационарных четырех и более стадийных алгоритмов. Схемы основных таких алгоритмов имеют следующие вид: локальный поиск + глобальный поиск 1 + глобальный поиск 2 + «встряска» популяции (алгоритм поиска волчьей стаей (WPSA)); локальный поиск + глобальный поиск 1 + глобальный поиск 2 + глобальный поиск 3+ «встряска» популяции (алгоритм интеллектуальной стаи рыб (AFSA)).

Еще два примера представляют алгоритм выращивания саженцев (SGuA) и алгоритм искусственной иммунной системы (AIS).

Алгоритм выращивания саженцев (SGuA):

1. скрещивание 1 – выполняем скрещивание всех возможных пар s_i, s_j особей популяции $S: S \rightarrow S'$;
2. мутация – если координата $x'_{i,k}, k \in [1:|X|]$ особи $s'_i \in S'$ была изменена на предыдущем этапе алгоритма, то с вычисляемой вероятностью по тому или иному правилу подвергаем мутации компоненту $x'_{i,l}, l \in [1:|X|], l \neq k$ того же вектора: $S' \rightarrow S''$;
3. скрещивание 2 – если манхеттоновское расстояние между особями $s''_i, s''_j, i \neq j$ превышает заданный порог, то выполняем обмен между векторами X''_i, X''_j их k -ми компонентами: $S'' \rightarrow S'''$;
4. формирование новой популяции – объединяем популяции S, S''' в промежуточную популяцию, и популяцию $S(t+1)$ получаем путем отбора без повторения из промежуточной популяции $|S|$ лучших особей: $\{S, S'''\} \rightarrow S(t+1)$.

Алгоритм искусственной иммунной системы (AIS):

1. клонирование – из текущей популяции S выбираем заданное число лучших особей и клонируем их с фиксированным коэффициентом клонирования: $S \rightarrow S^m$;
2. мутация – реализуем мутацию отобранных особей $s^m_i \in S^m: S^m \rightarrow S'^m$;
3. сжатие – осуществляем сжатие множества S'^m путем исключения из него всех особей, кроме заданного числа лучших: $S'^m \rightarrow S''^m$;
4. формирование новой популяции $S \rightarrow S'^m$ объединяем множества особей S, S''^m и выполняем сжатие полученного набора особей до множества $S(t+1)$ мощности $|S|$ с сохранением наилучших решений: $\{S, S''^m\} \rightarrow S(t+1)$.

2. Программно динамические алгоритмы. В П-алгоритмах данного класса в процессе эволюции популяции программно (в зависимости от номера итерации) меняются структура и/или значения параметров алгоритма.

2.1. Одностадийные алгоритмы. Из числа алгоритмов, представленных в [2], к этому классу относится только оптимизационный алгоритм искателя (SeOA). В этом алгоритме для каждой из особей $s_i \in S$ ее новое положение находим по формуле, компонентами которой являются инерциальная (инициативная), когнитивная (эгоистичная) и социальная (альтруистическая) составляющие.

2.2. Двух стадийные алгоритмы. В [2] представлены два следующих программно динамических алгоритма (в которых программно изменяются значения параметров).

Метод взрыва гранты (GEM):

1. локальный поиск – в окрестности каждой из особей $s_i \in S$ реализуем оригинальный многоточечный одношаговый алгоритм локального поиска (порождающий только допустимых особей): $S \rightarrow \{\tilde{s}_{i,j}\}$;
2. формирование новой популяции – если $\varphi(X_i) < \varphi(X_{i,j^*})$, то полагаем $X_i(t+1) = X_{i,j^*}$. В противном случае полагаем $X_i(t+1) = X_i$: $\{S, \{\tilde{s}_{i,j^*}\}\} \rightarrow S(t+1)$.

Алгоритм мозгового штурма (BSOA). Рассматриваем два следующих варианта этого алгоритма, отличающиеся содержанием второй стадии [2]:

1. кластеризация популяции – с помощью того или иного алгоритма кластеризации разбиваем в пространстве поиска популяцию S на заданное число кластеров $\{S_i\}$ и объявляем лучших особей кластеров их центрами; с заданной вероятностью в каждом из кластеров S_i замещаем его центральную особь на случайную по правилам инициализации начальной популяции: $S \rightarrow S'$;
- 2.1. глобальный поиск (одно кластерная стратегия поиска) – равномерно случайно выбираем из популяции S' заданное число особей и перемещаем каждую из них в пространстве поиска с заданной вероятностью в случайно возмущенном направлении центра случайно выбранного кластера, а в противном случае аналогично в направлении особи, выбранной из этого кластера случайным образом: $S' \rightarrow S(t+1)$;
- 2.2. глобальный поиск (двух кластерная стратегия поиска – аналогично первой стратегии выбираем из популяции S' заданное число особей и перемещаем каждую из них в пространстве поиска с заданной вероятностью в случайно возмущенных направлениях центров двух случайных кластеров, а в противном случае аналогично – в направлениях двух случайных особей указанных кластеров: $S' \rightarrow S(t+1)$.

2.3. Трех и более стадийные алгоритмы. В [2] нашли место четыре П-алгоритма данного класса. Стадии эволюционного алгоритма поиска с возвратом (BSA): селекция 1 + скрещивание 1 + скрещивание 2 + селекция 2. Аналогично, стадии поискового алгоритма охоты (HuS) имеют вид: глобальный поиск 1 + мутация + глобальный поиск 2.

Ниже представлены схемы еще двух П-алгоритмов данного класса.

Алгоритм лучевой оптимизации (ROA):

1. селекция – формируем набор $C \subset S$ заданного числа лучших особей текущей популяции: $S \rightarrow C$;
2. глобальный поиск – определяем возможные положения каждой из особей $s_i \in S$, используя с заданной вероятностью случайные составляющие вектора скорости особи или строим их на основе аттракторной и инерционной компонент, где аттракторная компонента есть линейная комбинация положений лучшей особи популяции и случайной особи набора C : $S \rightarrow \tilde{S}$;
3. удовлетворение ограничений – если положение особи $\tilde{s}_i \in \tilde{S}$ является недопустимым, то возвращаем ее в эту область по оригинальным правилам: $\tilde{S} \rightarrow S(t+1)$.

Алгоритм динамических сеток (VMO):

1. локальный поиск – в некоторой $\mathcal{R}_X$ -окрестности каждой из особей $s_i \in S$ выбираем лучшую особь s_{i,j^*} и с вероятностью, определяемую близостью указанных особей в пространстве $\mathcal{R}_\varphi$, создаем новую особь s_i^z в направлении особи s_{i,j^*} : $S \rightarrow S^z$;

2. глобальный поиск 1 – для каждой из особей $s_i \in S$, исключая глобально лучшую особь популяции s^{gbest} , с вероятностью, определяемую близостью указанных особей в пространстве $\mathcal{R}_\varphi$, создаем новую особь в направлении особи s^{gbest} : $S \rightarrow S^w$;
3. глобальный поиск 2 – вычисляем числа *внутренних* особей $\{s_j^d\}$, ближайших в пространстве $\mathcal{R}_X$ к центру области поиска O^c , а также *внешних* особей $\{s_j^o\}$, наиболее удаленных от этой точки, и для каждой из особей наборов $\{s_j^d\}$, $\{s_j^o\}$ создаем новую особь на программно уменьшающемся расстоянии от этой особи в пространстве $\mathcal{R}_X$: $S \rightarrow S^y$;
4. формирование новой популяции – объединяем особей популяции S с наборами особей S^z, S^w, S^y и сортируем полученный набор особей в порядке убывания значений фитнес-функции; прореживаем полученную совокупность особей, удаляя близких в пространстве $\mathcal{R}_X$ особей. Если мощность полученного множества особей меньше величины $|S|$, то недостающих особей генерируем по схеме инициализации начальной популяции; в противном случае в новую популяцию включаем $|S|$ лучших с точки зрения значений фитнес-функции особей указанного множества: $\{S, S^z, S^w, S^y\} \rightarrow S(t+1)$.

3. Адаптивные динамические алгоритмы. В алгоритмах данного класса структура и/или значения параметров алгоритма могут изменяться в процессе эволюции популяции в зависимости от значений *характеристик* некоторых сущностей популяции.

3.1. Одностадийные алгоритмы. В [2] представлены четыре следующих алгоритма этого класса.

Алгоритм светлячков (FA) использует характеристику данной особи, имеющую смысл ее *привлекательности* для других особей популяции. Если $\varphi(X_i) < \varphi(X_j)$, то перемещаем особь $s_i \in S$ в рандомизированном направлении особи s_j с шагом, равном привлекательности особи s_i для особи s_j .

Алгоритм оптимизации роем светлячков (GSO) использует характеристику особей популяции, называемую *уровнем светимости*. Выбираем из множества $\mathcal{R}_X$ соседей каждой данной особи $s_i \in S$ «притягивающую» особь с вероятностью, пропорциональной уровню ее светимости, и перемещаем особь s_i в направлении выбранной особи.

Интегрированная радиационная оптимизация (IRO). Обновляем положение каждой из особей $s_i \in S$ по параметрически адаптивной итерационной формуле, включающей в себя социальную компоненту.

Алгоритм лиги чемпионов (LCA) использует следующие основные обозначения: особь – команда; популяция особей – лига; сравнение значений фитнес-функции двух команд – матч; итерация алгоритма – тур, который состоит из некоторого числа раундов, в каждом из которых каждая из команд проводит один матч. Моделируем заданное число матчей между командами в соответствии с предварительно сформированным расписанием LS . Полагаем, что вероятность победы команды $s_{i_\theta} \in S$ (и проигрыша команды $s_{j_\theta} \in S$) в матче между этими командами пропорциональна текущему нормированному значению фитнес-функции φ_{i_θ} . На основе результатов матчей между командами определяем новые возможные положения особей популяции S по стохастическим трех компонентным формулам и перемещаем их в лучшее из найденных возможных положений.

3.2. Двух стадийные алгоритмы. Примерами стадий этого класса адаптивных П-алгоритмов являются: расщепление популяции + глобальный поиск (стохастический фокусирующий поиск (SFS)); скрещивание + глобальный поиск (алгоритм искусственного племени (АТА)); глобальный поиск + «встряска» популяции (алгоритм социально-эмоциональной оптимизации (SEOA)); локальный поиск + формирование новой популяции (простой оптимизационный алгоритм (SOPT)); глобальный поиск + локальный поиск (алгоритм броуновского движения газа (GBMO)).

Представляем также схемы двух следующих П-алгоритмов данного класса.

Оптимизационный алгоритм распространения тараканов (RIO). Алгоритм использует такую характеристику особей популяции, как уровень их *голода*.

1. глобальный поиск – если текущий «голод» особи $s_i \in S$ не превышает заданный уровень, то перемещаем ее в пространстве поиска по правилам, близким к правилам, используемым алгоритмом роя частиц PSO: $S \rightarrow S(t+1)$;
2. «встряска» популяции – в противном случае перемещаем особь $s_i \in S$ в случайную точку начального параллелепипеда Π , созданную по правилам инициализации начальной популяции: $S \rightarrow S(t+1)$.

Оптимизационный алгоритм стволовых клеток (SCOA). Задача безусловной максимизации.

1. репликация – если $\varphi(X^{gbest}) > \varphi(X^{gbest}(t-1))$ (приспособленность особи s^{gbest} лучше приспособленности глобально оптимальной особи на предыдущей итерации), то реплицируем особь s^{gbest} и если мощность набора $S \cup s^{gbest}$ превышает заданную величину, то удаляем из этого набора худшую особь: $S \rightarrow S'$;
2. скрещивание – каждую из особей $s' \in S'$ с вероятностью ξ_i , пропорциональной приспособленности этой особи, скрещиваем с особью s'_j , $j \neq i$, равномерно случайно выбранной из популяции S' : $S' \rightarrow S''$.

3.3. Трех и более стадийные алгоритмы. Алгоритмы данного класса могут, например, иметь следующие стадии: локальный поиск + глобальный поиск 1 + глобальный поиск 2 (алгоритм поиска косяком рыб (FSS)); глобальный поиск + скрещивание + мутация (алгоритм стаи криля (КН)); создание особей + объединение + формирование новой популяции (оптимизационный алгоритм фейерверков (FOA)); создание новых особей 1 + создание новых особей 2 + создание новых особей 3 + скрещивание + удовлетворение ограничений (непрерывная оптимизация химической реакцией (RCCRO)); вычисление характеристик подобластей + создание новых особей + локальный поиск + формирование новой популяции (алгоритм бессознательного поиска (USA)); глобальный поиск + локальный поиск + «встряска» регионов поиска (непрерывный ортогональный алгоритм муравьиной колонии (COAC)). Отметим, что последний алгоритм использует окрестности поиска (регионы), которые эволюционируют в процессе эволюции популяции на основе оценки их такой характеристики, как *привлекательность*, вычисляемой на основе текущей интенсивности всех феромонных пятен, принадлежащих региону.

В качестве примера адаптивных П-алгоритмов данного класса представляет схемы алгоритма оптимизации колонией бактерий (BCO) и алгоритма поиска бактериями пищи (BFO).

Алгоритм оптимизации колонией бактерий (BCO). Алгоритм использует вещественную характеристику особей популяции, называемую *энергией*, а также бинарную характеристику *здоровье*.

1. глобальный поиск – каждую из особей $s_i \in S$ перемещаем в пространстве поиска на один шаг по формуле, содержащей компоненты, соответствующие текущим глобально лучшей и исторически лучшей особей, а также случайную компоненту. Затем реализуем заданное число шагов перемещения особей $s_i \in S$ в направлениях текущих глобально лучшей и исторически лучшей особей: $S \rightarrow S' \rightarrow S''$;
2. клонирование – если приспособленность φ_i'' особи превышает рассчитанный уровень «энергии» и положение особи является допустимым, то клонируем (реплицируем) эту особь: $s_i'' \rightarrow \{s_i'', s_i'''\}$;
3. «встряска» популяции – если приспособленность φ_i'' особи не превышает указанный уровень «энергии», то независимо от допустимости или не допустимости ее текущего положения удаляем эту особь из популяции, а взамен создаем новую особь s_i''' по правилам инициализации начальной популяции: $s_i'' \rightarrow \{s_i'', s_i'''\}$;
4. формирование новой популяции – если число особей популяции S''' превышает величину $|S|$, то удаляем из этой популяции худших особей числом $|S'''| - |S|$: $S''' \rightarrow S(t+1)$.

Алгоритм поиска бактериями пищи (BFO). Алгоритм использует такую характеристику особи, как ее *состояние здоровья*, равную сумме значений фитнес-функции во всех точках следа этой особи от начальной до текущей итерации.

1. локальный поиск – для каждой особи $s_i \in S$ осуществляем заданное число шагов локальной оптимизации: $S \rightarrow S'$;
2. обновление – после каждых $\hat{t}_r$ шагов эволюции популяции S половину наиболее слабых особей (имеющих худшее состояние здоровья) исключаем из популяции, а каждую из особей оставшейся половины расщепляем на две особи: $S' \rightarrow S''$;
3. «встряска» популяции – после каждых $\hat{t}_e$ шагов эволюции популяции S с заданной вероятностью случайным образом выбираем из текущей популяции фиксированное число особей и уничтожаем их, а вместо каждой из уничтоженных особей в случайно выбранной точке области Π создаем новую особь с тем же номером: $S'' \rightarrow S(t+1)$.

Заключение. В статье не нашли место одно и много уровневые много популяционные П-алгоритмы. Предполагается, что эти алгоритмы составят предмет самостоятельной публикации.

СПИСОК ИСТОЧНИКОВ

1. Карпенко А.П. Современные алгоритмы поисковой оптимизации. Алгоритмы, вдохновленные природой / А.П. Карпенко. – М.: Издательство МГТУ им. Н.Э. Баумана, 2014. – 446 с.
2. Bo Xing, Wen-Jing Gao Innovative Computational Intelligence: A Rough Guide to 134 Clever Algorithms. Springer Cham. Heidelberg, New York, Dordrecht, London, 451 p.
3. Taherdangkoo M., Shirzadi M. H., Bagher, M. H. A novel meta-heuristic algorithm for numerical function optimization Blind Naked Mole-Rats (BNMR) algorithm. Scientific Research and Essays, 2012, vol. 7, pp. 3566–3583.
4. Гладков Л. А., Курейчик В. В., Курейчик В. М. и др. Биоинспирированные методы в оптимизации: монография / Л. А. Гладков, В. В. Курейчик, В. М. Курейчик и др. – М: Физматлит, 2009. – 384 с.

5. Биоэвристики: теория, алгоритмы и приложения: монография / С.И. Родзин, Ю.А. Скобцов, С.А. Эль-Хатиб. – Чебоксары: ИД «Среда», 2019. – 224 с.
6. Пантелеев А.В., Метлицкая Д.В., Алешина Е.А. Методы глобальной оптимизации: Метаэвристические стратегии и алгоритмы / А.В. Пантелеев, Д.В. Метлицкая, Е.А. Алешина. – М.: Вузовская книга, 2013. – 44 с.
7. Can Umit, Alatas Bilas Physics Based Metaheuristic Algorithms for Global Optimization. American Journal of Information Science and Computer Engineering, 2015, vol. 1, no. 3, pp. 94-106.
8. Engelbrecht Andries P. Computational Intelligence. An Introduction. John Wiley & Sons Ltd, England, 2007, 597 p.
9. Evolutionary Computation 1. Basic Algorithms and Operators. Edited by Thomas Back, David B Fogel and Zbigniew Michalewicz. Institute of physics publishing Bristol and Philadelphia, 2000, 339 p.
10. Evolutionary Computation 2. Advanced Algorithms and Operators. Edited by Thomas Back, David B Fogel and Zbigniew Michalewicz. Institute of physics publishing Bristol and Philadelphia, 2000, 270 p.
11. Glover F., Kochenberger G.A. Handbook of metaheuristics. Springer, 2010, 648 p.
12. Nazmul Siddique, Hojjat Adeli. Nature Inspired Computing: An Overview and Some Future Directions. Cognitive Computation, 2015, vol. 7, pp. 706–714.
13. Xin-She Yang Nature-Inspired Metaheuristic Algorithms. Second Edition. Luniver Press, United Kingdom, 2010, 115 p.

Карпенко Анатолий Павлович, д.ф.-м.н., профессор, зав. кафедрой САПР, МГТУ им. Н.Э. Баумана, apkarpenko@mail.ru, Россия, г. Москва, ул. 2-я Бауманская, д.5

UDC 519.6

doi: 10.38028/ESI.2022.25.1.004

Typical structures of population algorithms for global optimization

Anatoly P. Karpenko

N.E. Bauman MSTU,

Russia, Moscow, apkarpenko@mail.ru

Abstract. We present the author's classification of structures of population algorithms for global optimization (P-algorithms). At the population level, we single out the following structural classes of these algorithms: one-level one-population algorithms; one-level multi-population algorithms; two-level algorithms; three level algorithms. At the next level of the hierarchy, we distinguish one-stage (two macro-operators), two-stage (two macro-operators) and three or more stages (three or more macro-operators) algorithms. On the other hand, we subdivide P-algorithms into stationary, program-dynamic, parametric and structural-adaptive dynamic algorithms. We consider only one population algorithms.

Keywords: global optimization; population algorithms; metaheuristic algorithms; algorithms inspired by nature

REFERENCES

1. Karpenko A.P. Sovremennyye algoritmy poiskovoy optimizatsii. Algoritmy, vdokh-novlen-nyye prirodoy [Modern search engine optimization algorithms. Algorithms inspired by nature]. M., Izdatel'stvo MGTU im. N.E. Baumana, 2014, – 446 p.
2. Bo Xing, Wen-Jing Gao Innovative Computational Intelligence: A Rough Guide to 134 Clever Algorithms. Springer Cham. Heidelberg, New York, Dordrecht, London, 451 p.
3. Taherdangkoo M., Shirzadi M. H., Bagher, M. H. A novel meta-heuristic algorithm for numerical function optimization Blind Naked Mole-Rats (BNMR) algorithm. Scientific Research and Essays, 2012, vol. 7, pp. 3566–3583.
4. Gladkov L. A., Kurejchik V. V, Kurejchik V. M. i dr. Bioinspirovannyye metody v optimizatsii: monografiya [Bioinspired methods in optimization: monograph]. M: Fizmatlit, 2009, 384p.
5. Rodzin S.I., Skobcov YU.A., El'-Hatib S.A. Bioevristiki: teoriya, algoritmy i prilozheniya: monografiya [Bioheuristics: theory, algorithms and applications: monograph]. Cheboksary, ID «Sreda», 2019, 224p.
6. Panteleev A.V., Metlickaya D.V., Aleshina E.A. Metody global'noj optimizatsii: Metaevristicheskie strategii i algoritmy [Methods of global optimization: Metaheuristic strategies and algorithms]. M., Vuzovskaya kniga, 2013, 44p.
7. Can Umit, Alatas Bilas Physics Based Metaheuristic Algorithms for Global Optimization. American Journal of Information Science and Computer Engineering, 2015, vol. 1, no. 3, pp. 94-106.
8. Engelbrecht Andries P. Computational Intelligence. An Introduction. John Wiley & Sons Ltd, England, 2007, 597 p.
9. Evolutionary Computation 1. Basic Algorithms and Operators. Edited by Thomas Back, David B Fogel and Zbigniew Michalewicz Institute of physics publishing Bristol and Philadelphia, 2000, 339 p.
10. Evolutionary Computation 2. Advanced Algorithms and Operators. Edited by Thomas Back, David B Fogel and Zbigniew Michalewicz. Institute of physics publishing Bristol and Philadelphia, 2000, 270 p.
11. Glover F., Kochenberger G.A. Handbook of metaheuristics. Springer, 2010, 648 p.
12. Nazmul Siddique, Hojjat Adeli. Nature Inspired Computing: An Overview and Some Future Directions. Cognitive Computation, 2015, vol. 7, pp. 706–714.
13. Xin-She Yang Nature-Inspired Metaheuristic Algorithms. Second Edition. Luniver Press, United Kingdom, 2010, 115 p.

Anatoly P. Karpenko, dr., professor, Head of CAD department, N.E. Bauman MSTU, ap-karpenko@mail.ru, Russia, Moscow, st. 2nd Baumanskaya, 5

Статья поступила в редакцию 03.03.2022; одобрена после рецензирования 23.03.2022; принята к публикации 28.03.2022.

The article was submitted 03.03.2022; approved after reviewing 03.23.2022; accepted for publication 03.28.2022.

Отслеживание мобильных объектов средствами беспроводных сенсорных сетей

Виноградов Геннадий Павлович

НИИ «Центрпрограммсистем»,

Россия, Тверь, wg272ng@mail.ru

Аннотация. Целью исследования является повышение эффективности алгоритмов отслеживания мобильных объектов (целей CSP) в распределенных сенсорных сетях в составе киберфизических систем на базе организации их взаимодействия с алгоритмами инициализации/маршрутизации. Задача локализации и отслеживания объектов вторжения в защищаемую зону составляет основу системы защиты особо важных объектов. Отслеживание нескольких целей с помощью беспроводной сенсорной сети является актуальной проблемой. Успешное ее решение средствами сенсорных сетей предполагает решение ряда проблем, среди которых следует выделить две основные, определяющие успешную работу сенсорных сетей. Первая – это наличие эффективных методов обмена информацией между локальными узлами, вторая – организация совместной обработки сигналов (CSP) группой узлов на основе собранной информации о состоянии среды в зоне их ответственности в результате наступления некоторых событий. В качестве модельного примера решается задача обнаружения, локализации и отслеживания одной цели в защищаемой зоне. На этом примере рассмотрены основные идеи, лежащие в основе интеграции алгоритмов, а также основные аспекты реализации алгоритмов обнаружения, локализации и отслеживания. Предлагаемые решения учитывают ограничения, связанные с возможностями локальных узлов, сетью в целом и маршрутизацией. Источником данных для предлагаемых алгоритмов являются сигналы от звуковых, сейсмических, тепловых и т.п. сенсоров, у которых мощность сигнала имеет выраженный максимум в зависимости от расстояния от цели и до узла сети. Полученные результаты распространяются на проблему отслеживания множества целей, что предполагает оценку применимости методов идентификации и классификации в условиях, когда наблюдается наложение воспринимаемых сигналов сенсорами. Рассматривается архитектура системы совместной обработки данных в распределенных сенсорных сетях. Предметом рассмотрения являются методы отслеживания множества движущихся объектов в защищаемой зоне. Основные этапы процедуры отслеживания состоят из: обнаружения целей, их классификации, оценок местоположения и прогнозирования траектории передвижения цели. Обсуждаются алгоритмы для решения этих задач.

Ключевые слова: сенсорные сети, обнаружение, локализация, классификация, отслеживание

Цитирование: Виноградов Г.П. Отслеживание мобильных объектов средствами беспроводных сенсорных сетей // Информационные и математические технологии в науке и управлении. – 2022. – № 1 (25). – С. 58-69. – DOI:10.38028/ESI.2022.25.1.005.

Введение. Беспроводные сенсорные сети (БСС) в настоящее время рассматриваются как составная часть киберфизических систем, в которых они осуществляют интерфейс между

физическим миром и информационными технологиями [1-5]. Как правило, узел БСС оснащен датчиками контроля нескольких параметров среды, но имеет ограниченные энергетические и вычислительные ресурсы. Разработка системы защиты территорий и объектов средствами киберфизических систем, использующих сенсорные сети, предполагает решение ряда проблем, требующих значительных вычислительных затрат. Среди них следует выделить две основные. Первая – это создание эффективных методов обмена информацией между локальными узлами при локализации и отслеживании цели, вторая – организация *совместной обработки сигналов* (англ. *CSP*) группой узлов на основе собранной информации о состоянии среды в зоне их ответственности по результатам наступления событий. Их решение связано с необходимостью снижения энергопотребления в сети и повышения живучести системы. Особенно это важно при отслеживании нескольких целей. Возможные подходы и варианты решения отслеживания одной цели рассматривались в [5, 6-16, 17]. В этих работах рассматривались алгоритмы обработки только пространственной одномоментной информации. Мониторинг зоны защиты объектов средствами БСС требует решения задач обнаружения, классификации, идентификации и отслеживания множества целей. Задачи локализации и отслеживания множества целей требуют совместной работы группы узлов. Это означает значительное усложнение задач обработки сигналов в сенсорной сети и повышение энергопотребления [18]. Информация от узлов должна обобщаться и анализироваться в координатах пространства-времени. Это требует достаточно интенсивной работы вычислительных средств, как отдельного узла, так и сети в целом. Известно, что энергопотребление является критическим фактором в беспроводной сенсорной сети. Ограниченное количество энергии, содержащейся в аккумуляторе каждого узла, должно обеспечивать множество функций, таких, как работа датчиков, обработка сигналов на борту и связь с соседними узлами. Экономия энергии предполагает разработку энергоэффективных методов и алгоритмов измерения, обработки сигналов, низкую частоту дискретизации, малое число энергоемких протоколов обмена информацией между узлами. Следовательно, должна быть организована распределенная обработка пространственно-временных сигналов путем организации взаимодействия между узлами, а алгоритм работы CSP должен быть наделен набором следующих функциональностей: 1) *распределенная обработка* – сигналы сенсоров должны после считывания предварительно обрабатываться на отдельных узлах, а не передаваться напрямую по беспроводному каналу; результат обработки должен храниться на локальных узлах и передаваться другим узлам *по запросу* или *после обработки событий*; 2) *обработка по событию* – для экономии энергии каждый узел выполняет только те задачи обработки сигналов, которые имеют отношение к текущему *событию*; при отсутствии события/запроса узел должен переходить в режим «сна», чтобы минимизировать потребление энергии; информация от конкретного узла в пункт управления пересылается только после обработки события “*Определение факта проникновения*” или по запросу, а не автоматически с заданным шагом дискретизации; 3) *агрегирование информации* – для обобщения информации о состоянии защищаемой зоны на основе локальных данных пространственно-временной природы CSP должна выполнять эффективное иерархическое агрегирование информации, тогда низкая пропускная способность сети будет достаточна для передачи информации от отдельных узлов к пункту управления и через него к узлам сети, при отслеживании целей; более высокая пропускная способность потребуется при выполнении задачи классификации целей по данным временных рядов, которыми должны обмениваться соседние узлы сети; 4) *динамическая кластеризация* – в зависимости от характера решаемой задачи, например, локализации объектов вторжения, может потребоваться использование кластеров меньшего размера, чем полученные при

инициализации сети и более высокая частота опроса узлов; шаг дискретизации по времени и пространственным координатам должен либо настраиваться, либо иметь достаточное количество возможных значений.

1. Динамическая кластеризация. Объект проникновения в процессе движения генерирует акустические, сейсмические, тепловые и другие сигналы. Величина уровня мощности воспринимаемого сенсорами сигнала (англ. *RSSI – Received Signal Strength Indicator*) определяется расстоянием “узел-цель”. Спектр сигналов будет иметь максимум при прохождении цели над узлом или в непосредственной близости, и является индивидуальной характеристикой типа цели. Ее следует рассматривать как сигнатуру и использовать при определении типа цели. Этот факт обсуждался в работах [4, 5, 19, 17, 20]. Положение узла, фиксирующего максимум сигнала, можно использовать для определения координат цели, а отслеживание объекта равносильно отслеживанию местоположения максимума во времени и пространстве при прохождении целью точек доступа сети. В поле пространственного сигнала движущийся объект может идентифицироваться пиком сигнатуры. Значение сигнала сенсора конкретного узла изменяется при перемещении объекта в зоне ответственности узла и воспринимается широкоэвентально, то есть все узлы, в радиусе восприятия которых находится цель, ее «слышат». Узлы, «слышащие» цель, осуществляют считывание значений спектра сигнала цели в определенные моменты времени при ее движении. Тем самым формируется пространственная и временная выборка поля фактической сигнатуры цели. В этом случае плотность размещения узлов должна быть достаточной для определения зависимости уровня сигнала от расстояния узел-цель. По аналогии, временные ряды значений уровня сигнала от датчиков каждого узла должны формироваться с частотой, соизмеримой со скоростью движения цели. Таким образом, характер изменения поля пространственно-временной сигнатуры и характер запроса определяют требуемую частоту дискретизации в пространстве-времени.

В работе [20] для обеспечения отслеживания цели и эффективной локальной обработки данных в сенсорной сети предложено разделить область проникновения на пространственно-временные ячейки (окна). Их размер зависит от скорости, направления движения цели и от показателя затухания прохождения сигнала в среде. Размер ячейки должен приблизительно соответствовать области, внутри которой в течение некоторого времени спектр считываемой сигнатуры остается практически постоянным и на границах не превышает некоторого порогового значения. Это предполагает учет пространственно-временных координат при анализе, а также динамическую коррекцию размера пространственно-временных ячеек (кластеров) на основе прогнозируемых местоположений, типов целей и показателей их движения. Усреднение пространственно-временного сигнала по узлам в каждой ячейке-кластере позволит уменьшить влияние шума и повысить эффективность алгоритмов локализации, определения типа цели и ее отслеживания, а также уменьшить объем передаваемых данных. В предлагаемом варианте событие «обнаружение цели» будет возникать тогда, когда выход сенсора узла превышает некоторый порог, величина которого регулируется так, чтобы величина частоты ложной тревоги не превышала некоторую норму. Узел передает головному узлу следующие данные: время начала события, при котором выход датчика превышает порог; показание датчика во время прохождения точки входа; значение максимума; время окончания события, когда выход датчика падает ниже порога. Отметим, что в существующих БСС частота дискретизации по времени определяется протоколом и рядом других факторов, таких, как рельеф местности, наличие преград, градиенты температуры и характер сигнала. Шаг дискретизации может превышать 20 секунд для экономии энергии и определяется также требованиями скрытности сети. Таким образом,

возникает проблема нахождения компромисса между сохранением живучести сети и ее эффективностью.

2. Взаимодействие узлов при обнаружении и отслеживании цели. Решение задачи обнаружения, локализации и отслеживания одной цели в распределенной сенсорной сети предполагает взаимодействие между сетевыми протоколами, протоколами маршрутизации и алгоритмами CSP. Маршруты передачи информации сетевыми протоколами в сенсорной сети определяются по минимуму расстояния прохождения сигнала [15, 21], то есть структура сети не ориентирована на обработку событий в локальной области зоны и на распределенную форму решения задач локализации, классификации и отслеживания группы узлов для снижения расхода энергии аккумуляторов. Пространственно-временная коррекция размера ячеек в регионе проникновения защищаемой зоны путем разделения на более мелкие области (пространственные ячейки), в которых выполняется предположение о постоянстве сигнатуры, позволяет исключить эту ситуацию.

Одиночная цель [20]. Пусть область, где фиксируется факт появления цели, имеет вид прямоугольника. Определим на углах этой области четыре ячейки А, В, С и D, и предположим, что цель может войти в контролируемую область через один из четырех углов области, например, через ячейку А. Узлы этой ячейки активируются при обнаружении цели. Активированный узел запускает алгоритм обнаружения спектра уровня сигнала объекта вторжения. Считывание этих данных производится с априорно фиксированной скоростью в зависимости от характеристик вероятной цели. Алгоритм отслеживания цели состоит из следующих пяти шагов:

1. Часть, возможно, все узлы в ячейке А обнаруживают цель. Эти узлы становятся активными узлами, а ячейка А – активным кластером. Активные узлы передают информацию о времени проникновения и результатах измерений спектра сигнала головному узлу в N последовательные моменты времени.
2. В каждый момент времени головные узлы определяют местоположение цели по выходам сенсоров активных узлов. Простейшей оценкой местоположения цели в данный момент является местоположение узла с самым сильным сигналом в этот момент. Могут быть использованы более сложные алгоритмы для локализации цели. Применение сложных алгоритмов локализации будет оправдано только в том случае, если точность определения местоположения цели будет выше, чем расстояние между узлами.
3. Головные узлы используют местоположения цели в N последовательных моментов времени для прогнозирования местоположения цели в M ($< N$) будущие моменты времени.
4. Предсказанные положения цели используются протоколами для активизации или создания новых ячеек, в которые цель, скорее всего, может попасть. Так определяются области из ячеек, в которые цель попадет после того, как покинет текущую активную ячейку. Подмножество этих ячеек активируется протоколами для продолжения обнаружения и отслеживания цели.
5. Как только цель обнаружена в одной из новых ячеек, она обозначается как новая активная ячейка, а узлы в исходной активной ячейке переводятся в состояние сна для экономии энергии.

Шаги 1 – 5 повторяются для новой активной ячейки, и это составляет основу процесса обнаружения и отслеживания одной цели. Для каждой обнаруженной цели передается сообщение, содержащее информацию об отслеживании – местоположение цели, время обнаружения, время прохождения зоны ответственности – передается от текущей активной ячейки в следующую. Аналогичные алгоритмы рассматриваются в [6-8].

Множественные цели. Пусть несколько целей проникают в разные области защищаемой зоны в разные моменты времени, и эти области достаточно разделены пространственно. Это допущение означает, что выходной сигнал датчика спектра конкретного узла будет содержать различимые пики в конкретные моменты, соответствующие сигнатурам разных целей. Тогда может быть использована процедура, описанная выше: для каждой цели иницируется и поддерживается свой трек отслеживания. Если предположение о достаточном разделении положения цели в пространстве и/или времени не соответствует реальности, то в этом случае необходимы алгоритмы классификации, которые используют базу пространственно-временных сигнатур целей. Ее формирование предполагает формирование априорных знаний о типовых сигнатурах для различных классов целей. Использование этих знаний позволяет предложить алгоритм определения типа цели одним узлом (без взаимодействия между узлами) на основе отслеживания сигнатур целей во времени. Узел формирует сегмент временного ряда обнаруженного события и направляет его для классификации в головной узел. Соответствующий алгоритм выполняет обработку, например, с помощью быстрого преобразования Фурье (БПФ), и преобразованный вектор сравнивается с имеющимися в базе и определяется наиболее подходящая сигнатура. Выходные данные классификатора о такой сигнатуре, соответствующей цели, передаются на головные узлы для определения отличий от выходных данных результатов измерения другими узлами. Этапы 1–5 повторяется для создания и поддержания треков для построения и прогнозирования направления и будущего местоположения нескольких классифицированных целей.

3. Алгоритмы обработки сигналов. Обнаружение выполняется путем измерения спектра сигнала цели. Для этого выходной сигнал активных узлов отбирается с заданным шагом. Размер окна и частота дискретизации определяются характеристиками цели, такими как скорость движения сигнатуры, ожидаемая продолжительность ее нахождения в активной ячейке и конкретный способ измерения. Событие обнаруживается, когда выходной сигнал детектора превышает пороговое значение. Выходные параметры узла, которые передаются головным узлам, состоят из: 1) начального времени, когда выход детектора превышает пороговое значение; 2) времени фиксации максимума энергии сигнала (СРА); 3) текущие значения выходного сигнала детектора во время СРА; 4) времени ухода, когда выход детектора узла падает ниже порогового значения.

Для локализации цели выходной сигнал узла фиксируется и передается в моменты времени в пределах времени начала проникновения и ухода цели из зоны узла. Определение местоположения цели в последовательные моменты времени может быть реализовано различными алгоритмами, описанными в [5, 12, 17, 19, 22, 23]. Задачи минимизации расхода энергии сети предполагают использование простых алгоритмов оценки местоположения цели в конкретный момент времени с помощью измерений спектра в нескольких (*четырёх и более*) узлах. Например, использование алгоритмов, основанных на оценке временной задержки сигнала, требует применения точной временной синхронизации узлов по времени и дорогого оборудования. Применение алгоритмов, основанных на использовании нейронных сетей, предполагает: а) измерение уровня сигнала от всех доступных по уровню сигнала точек доступа, называемых опорными; б) создание математической модели позиционирования; в) обучения модели [12]. Поддержание работоспособности устройств на основе данного метода требует значительных затрат. Методы когерентной локализации, например, метод формирования луча [3], требуют допущений о плоской волны. Такие предположения часто нарушаются в сенсорных сетях; например, цели могут находиться довольно близко к узлам во время опроса. Альтернативные методы, которые устраняют такие требования, также могут быть использованы, но предполагают необходимость

получения дополнительной статистической информации [5, 13, 23]. Ряд алгоритмов классификации предполагает передачу больших объемов данных о временных рядах от локальных узлов кластера к головному, что приводит к потреблению слишком большого количества энергии. Это делает их реализацию проблематичной.

Известен алгоритм локализации цели, основанный на измерении уровня сигнала, излучаемого целью, и на предположении об экспоненциальном затухании уровня сигнала с расстоянием от цели: $y_i(t) = \frac{s(t)}{\|r(t) - r_i\|^\alpha}$, где $y_i(t)$ – это значение уровня сигнала, считываемое i -м датчиком, $r(t)$ обозначает неизвестные координаты цели с относительно известного репера, r_i – координаты i -го датчика, $s(t)$ – уровень сигнала цели, α – показатель степени затухания, который предполагается известным (или его можно оценить экспериментально [24]). Доказано, что при отсутствии шума четыре и более окружности, определенные на показаниях 4-х или более датчиков, пересекаются в одной точке. Для измерений с шумом для надежности может использоваться более $n-1$ показаний, а неизвестное местоположение цели $r(t)$ определено методом наименьших квадратов:

$$J(x, y) = \sum_{i=1}^m \|(x - o_{i,x})^2 + (y - o_{i,y})^2 - \rho_i^2\|^2,$$

где (x, y) являются неизвестными координатами цели, $o_{i,x}$, $o_{i,y}$ – координаты i -го узла, а ρ_i – радиус окружности, связанный с i -м соотношением. Например, на основе 6-ти показаний датчиков можно построить семейство окружностей, внутри которых, возможно, находится цель. Три различные точки, в которых пересекаются 3 из 6 окружностей, будут указывать на местоположение цели. Точность оценок положения зависит от точности, с которой могут быть измерены координаты узлов и показатель затухания.

На основе данных о местоположении целей в различные моменты времени можно построить, с использованием метода наименьших квадратов, модель для прогнозирования их будущих местоположений [6, 25]. Для одной движущейся цели достаточно точное отслеживание может быть достигнуто с помощью линейной или полиномиальной модели. При наличии нескольких целей отслеживания такой метод определения траектории движения целей может приводить к необходимости нахождения решения системы одновременных уравнений. В реальных условиях решение проблемы локализации и отслеживания целей будут осложнять вариабельность константы затухания, зависимость уровня сигнала от направления движения и наложения сигналов от различных целей.

4. Классификация целей. Прежде всего, следует отметить, что использование спектра уровня сигналов в качестве сигнатур цели является оправданным для классификации, так как они значительно отличаются для разных классов целей. Сигнатуры строятся на основе временных рядов измерений в пределах интервала между временем начала фиксации и ухода цели. Точность классификации повышается за счет того, что каждый узел пространственно-временной ячейки определяет тип цели, используя данные, фиксируемые собственными сенсорами. Это позволяет реализовать алгоритмы совместной классификации. Для повышения точности классификации измеренные пространственно-временные сигнатуры, соответствующие нескольким модальностям, таким как акустические и сейсмические, могут быть объединены для повышения эффективности [10]. Однако этот прием вызывает значительную нагрузку на сеть, связанную с необходимостью передачи в головные узлы кластера данных временных рядов с используемыми модальностями от нескольких узлов ячейки. Как будет показано ниже, изменчивость временных сигнатур для конкретного класса целей создает серьезную проблему при разработке классификатора. В общем случае, требуется некоторое априорное знание характеристик сигнатур для различных классов целей [10]. Сформулируем задачу классификации в общем виде. Задача классификации в

машинном обучении – это задача отнесения объекта к одному из заранее определенных классов на основании данных его признаков, то есть классификатор сопоставляет каждый вектор x элементу $\omega \in \Omega$ таким образом, чтобы вероятность неправильной классификации – вероятность того, что метка классификатора отличается от истинной метки – была бы сведена к минимуму. Для достижения этой минимальной вероятности ошибок оптимальный классификатор использует решающее правило вида: x имеет метку ω_i , если $p(\omega_i|x) > p(\omega_j|x)$ для всех $j \neq i$, $\omega_i, \omega_j \in \Omega$. Практически оценку апостериорной вероятности $p(*)$ получить достаточно очень сложно. Вместо этого вводят дискриминантную функцию $g_i(x)$, для которой справедливо: неравенство $g_i(x) > g_j(x)$ имеет место, если $p(\omega_i|x) > p(\omega_j|x)$ для $j \neq i$, и для всех x . Тогда минимальная ошибка классификации может быть получена по правилу: принять, что x имеет метку ω_i , если $g_i(x) > g_j(x)$ для $j \neq i$. Минимальная вероятность неправильной классификации известна, как ошибка Байеса, а классификатор минимальной ошибки известен как классификатор Байеса или классификатор максимальной апостериорной вероятности (МАР). Ниже будут кратко рассмотрены три классификатора, которые требуют минимальной вычислительной работы и по характеристикам приближаются к классификатору Байеса: классификатор k -ближайших соседей (kNN), классификатор максимального правдоподобия (МП), и классификатор машины опорных векторов (SVM).

4.1. Классификатор k -ближайших соседей (kNN). Это самый простой классификатор. Классификатор kNN использует набор объектов, для которых заранее определены классы или данные в виде набора прототипов $\{p_k\}$. На этапе тестирования вычисляется расстояние между каждым тестовым вектором и каждым прототипом, и определяются k векторов прототипов, наиболее близких к тестовому вектору. Метки классов этих k -ближайших векторов-прототипов затем объединяются с использованием большинства голосов или какого-либо другого метода для определения метки класса тестового вектора. Когда $k = 1$, классификатор kNN называется классификатором ближайшего соседа. В [26] показано, что асимптотически (по числу обучающих векторов) вероятность неправильной классификации классификатора ближайшего соседа приближается к удвоенной (оптимальной) ошибке Байеса. Следовательно, оценка эффективности этого классификатора может быть использована в качестве базовой для оценки производительности других классификаторов, но по мере увеличения числа прототипов классификатор kNN начинает требовать для работы достаточно много памяти и вычислительной мощности.

4.2. Классификатор максимального правдоподобия. Это наиболее универсальный из «классических» методов классификации с обучением, поскольку позволяет не только разделять классы с различными типами функций плотности распределения признаков, но и минимизировать в среднем ошибки классификации. Приемлемым приближением следует считать, что распределение обучающих векторов соответствует смеси функций плотности распределения Гаусса. Это позволяет на основе известных результатов определить

$$p(x|\omega_i) \rightarrow G_i(x|\theta_i) = \sum_k |\Lambda_{ik}|^{-N/2} \exp\left(-\frac{1}{2}(x - m_{ik})^T \Lambda_{ik}^{-1}(x - m_{ik})\right),$$

где $\theta_i = [m_{i1}, \dots, m_{ip}, \Lambda_{i1}, \dots, \Lambda_{ip}]$ – параметры средней и ковариационной матрицы плотностей смеси P , соответствующих классу ω_i . Эти параметры модели могут быть идентифицированы путем применения соответствующего алгоритма кластеризации, такого, как алгоритм k -средних [26], к обучающим векторам каждого класса. Дискриминантная функция теперь вычисляется как $g_i(x) = G_i(x|\theta_i)p(\omega_i)$, где априорная вероятность $p(\omega_i)$ аппроксимируется относительным числом обучающих векторов в классе i .

4.3. Классификатор машины опорных векторов (SVM). Основная идея метода заключается в построении гиперплоскости, разделяющей объекты выборки оптимальным способом. Алгоритм работает в предположении, что чем больше расстояние (зазор) между разделяющей гиперплоскостью и объектами разделяемых классов, тем меньше будет средняя ошибка классификатора.

Все активные узлы в ячейке отправляют данные временных рядов, соответствующие обнаруженному событию, в головной узел. Головной узел обрабатывает данные временных рядов для классификации цели. Согласованная во времени обработка будет возможна, если известны и используются относительные временные интервалы различных временных рядов, например, в алгоритмах оценки пеленга на основе времени прибытия [9]. Аналогично, пространственно-временная обработка может быть возможна, если известны и используются относительные местоположения узлов. Например, в таких случаях может быть эффективно реализовано распределенное пространственное формирование луча для локализации целей, а также для подавления помех от нежелательных целей.

Доплеровские эффекты – это основная причина изменчивости сигналов. Они могут быть явно выявлены и оценены с помощью подхода, описанного в [22]. Показано, что влияние доплеровских эффектов на сигнатуру цели для m -го класса может быть смоделировано, как реализация стационарного процесса с нулевым средним значением, модулируемого изменяющейся во времени мгновенной частотой $f(\theta_m)$ с параметризованным вектором $\theta_m = [\alpha_m \ v_m \ d_m]$. Тогда классификатор осуществляет операции следующим образом: сначала для каждой гипотезы получают оценку ML доплеровских параметров $\widehat{\theta}_m$, а затем функции апостериорной плотности вероятности, соответствующие этим оценкам $p(\omega m | x, \widehat{\theta}_m)$, используются для выбора между классами. Такой подход, основанный на GLRT, требует “выравнивания” данных с некоторой фиксированной частотой на этапе обучения, что может быть сделано различными способами [4]. Выровненные данные затем могут быть использованы для обучения любого из трех классификаторов. Очевидно, что такие алгоритмы требуют больших вычислительных затрат.

Заключение. В статье представлены соображения, которые можно положить в основу разработки системы отслеживания нескольких целей средствами сенсорной сети. Основные компоненты системы включают обнаружение событий, оценку, прогнозирование местоположения и классификацию целей. На примере отслеживания одной цели и измерениях акустического или сейсмического сигналов от цели такая задача может быть решена. Определение типа и отслеживание нескольких целей одновременно требуют развития простых алгоритмов классификации. Спектры уровней сигналов от целей могут быть использованы для достаточно точного определения типа цели. Влияние ряда факторов, включая доплеровские эффекты, приводят к значительному искажению спектров сигналов, что создает серьезную проблему при определении типа цели. Это предполагает формирование базы сигнатур для классификации, например, колесных и гусеничных транспортных средств и сравнение на этой основе эффективности алгоритмов классификации. Алгоритмы, представленные в статье, представляют собой только один из подходов решения сложной проблемы отслеживания нескольких целей беспроводной сетью датчиков. В заключение, следует отметить, что распределенные беспроводные сенсорные сети превращаются в мощную технологию для дистанционного управления и мониторинга в различных предметных областях.

СПИСОК ИСТОЧНИКОВ

1. Акулдиз И.Ф. Беспроводные сенсорные сети: обзор /И.Ф. Акулдиз. – URL: <http://www.sciencedirect.com/science/article/pii/S1389128601003024>. Дата обращения: 09.01.2021.
2. A.R. Al-Ali, Ragini Gupta, Ahmad Al Nabulsi Cyber physical systems role in manufacturing technologies. AIP Conference Proceedings 1957, 2018, URL: <https://aip.scitation.org/doi/pdf/10.1063/1.5034337>.
3. L. P. Clare, G. J. Pottie, and J. R. Agre Self-Organizing Distributed Sensor Networks. Proc. SPIE, vol. 3713, April 199, pp. 229-238.
4. M. Chu, H. Haussecker, F. Zhao Scalable Information-Driven Sensor Querying and Routing for Ad Hoc Heterogeneous Sensor Networks. International Journal on High Performance Computing Applications, Special issue on Sensor Networks, 2002.
5. E. N. Onggosanusi, A. M. Sayeed, B. D. Van Veen Canonical Space-Time Processing for Wireless Communications. IEEE Trans. Commun., pp. 1669-1680, Oct. 2000.
6. R. Brooks, C. Griffin, and D. S. Friedlander Self-Organized distributed sensor network entity tracking. International Journal of High Performance Computing Applications, Special Issue on Sensor Networks, 2002.
7. R. Brooks and C. Griffin Traffic model evaluation of ad hoc target tracking algorithms. International Journal of High Performance Computing Applications, Special Issue on Sensor Networks, 2002.
8. J. C. Chen, R. E. Hudson, and K. Yao A Maximum likelihood parametric approach to source localization. Proc. ICASSP'2001, Salt Lake City, UT, 2001, pp. 1043-1046.
9. L. M. Kaplan, Q. Le, and P. Molnar Maximum likelihood methods for bearings-only target localization. Proc. ICASSP'2001, Salt Lake City, UT, 2001, pp. 554-557.
10. P. W. Boettcher and G. A. Shaw A Distributed Time-Difference of Arrival Algorithm for Acoustic Bearing Estimation. Proc. 4th International Conf. on Information Fusion, vol. 1, Montreal, August 2001.
11. A. M. Sayeed and D. L. Jones Optimal Detection Using Bilinear Time-Frequency and Time-Scale Representations. IEEE Transactions on Signal Processing, vol. 43, no. 12, pp. 2872-2883, December, 1995.
12. K. Yao, R. E. Hudson, C. W. Reed, D. Chen, and F. Lorenzelli Blind Beamforming on a Randomly Distributed Sensor Array System. IEEE Journal on Selected Areas in Communications. Special Issue on Signal Processing for Wireless Communications, vol. 16, no. 8, pp. 1555-1567, Oct. 1998.
13. E.A. Lee, S.A. Seshia A cyber-physical systems approach. Introduction to embedded systems, 2011
14. J. F. Scholl, J. R. Agre, L. P. Clare, and M. C. Gill A Low Power Impulse Signal Classifier Using the Haar Wavelet Transform. Proc. SPIE, vol. 3577, Nov. 1998, pp. 136-145.
15. J.M. Hsua, C.C. Chenb, C.C. Li POOT: An efficient object tracking strategy based on short-term optimistic predictions for face-structured sensor networks. Elsevier International Journal of Computers and Mathematics with Applications, vol. 63, issue 2, pp.391-406; 2012.
16. Paulo Tabuada Cyber-physical systems: Position paper. University of California at Los Angeles, 2010.
17. Рева И.А., Богданов А.А., Малахова Е.А. Применение точек доступа WI-Fi для регистрации движения на объекте / И.А. Рева, А.А. Богданов, Е.А. Малахова // Системы анализа и обработки данных, 2017. — № 3 (68). с. 104–125.

18. Xu Yingqi, Winter Julian, Lee Wang-Chien Prediction-based strategies for energy saving in object tracking sensor networks. Proceedings - 2004 IEEE International Conference on Mobile Data Management, 346- 357, 10.1109/MDM.2004.1263084.
19. Волков А.Н., Сиверс М.А., Сухов В.А. Позиционирование в сетях Wi-Fi / А.Н. Волков, М.А. Сиверс, В.А. Сухов // Вестник связи, 2010.
20. Dan Li, Kerry D. Wong, Yu H. Hu, Akbar M. Sayeed. Detection, Classification and Tracking of Targets in Distributed Sensor Networks, URL: <http://www.ecce.wisc.edu/~sensit>, accessed: 09.01.2021.
21. Кучерявый А.Е., Прокофьев А.В., Кучерявый Е.А. Самоорганизующиеся сети / А.Е. Кучерявый, А.В. Прокофьев, Е.А. Кучерявый. – СПб.: Любавич, 2011. – 312 с.
22. H. V. Poor. An Introduction to Signal Detection and Estimation. Springer Verlag, 1988.
23. C. W. Reed, R. Hudson, and K. Yao Direct Joint Source Localization and Propagation Speed Estimation. Proc. IEEE ICASSP, March 1999, pp. 1169-1172.
24. Martin Mikusz Towards an understanding of cyber-physical systems as industrial software-product-service systems. Conference paper in Procedia CIRP, 2014.
25. Виноградов Г.П., Емцев А.С., Федотов И.С. Беспроводные сенсорные сети в защищаемых зонах / Г.П. Виноградов, А.С. Емцев, И.С. Федотов // Известия ЮФУ. Технические науки, 2021. – №1. – с. 19–30.
26. Виноградов Г.П., Мальков А.А. Модели поиска структур данных на основе конкуренции и кооперации / Г.П. Виноградов, А.А. Мальков // Сб. трудов “Управление большими системами”. М.: ИПУ РАН, 2008. – Вып. 22. – С. 98-110

Виноградов Геннадий Павлович, д.т.н., профессор, зав. лабораторией

«Интеллектуальные системы», НИИ «Центрпрограммсистем», wgp272ng@mail.ru, г. Тверь, Россия, пр. Н. Корыткова, 3а

UDC 007.52:004.896:004:942

doi:10.38028/ESI.2022.25.1.005

Tracking mobile objects using wireless sensor networks

Gennady P. Vinogradov

Research Institute Centerprogramsystems,

Russia, Tver, wgp272ng@mail.ru

Abstract. The aim of the study is to improve the efficiency of the functioning of mobile object tracking algorithms (CSP targets) in distributed sensor networks as part of cyber-physical systems based on the organization of their interaction with initialization/routing algorithms. The task of localizing and tracking objects of intrusion into the protected zone is the basis of the protection system for critical objects. Tracking multiple targets with a wireless sensor network is a hot topic. Its successful solution by means of sensor networks involves solving a number of problems, among which there are two main ones that determine the successful operation of sensor networks. The first is the availability of effective methods of information exchange between local nodes, the second is the organization of joint signal processing (CSP) by a group of nodes based on the collected information about the state of the environment in their area of responsibility as a result of the occurrence of certain events.

The problem of detection, localization and tracking of one target in the protected zone is accepted as a model example. This example discusses the main ideas underlying the integration of algorithms, as well as the main aspects of the implementation of detection, localization and tracking algorithms. The proposed solutions take into account the limitations associated with the capabilities of local nodes, the network as a whole and routing. The source of data for the proposed algorithms are signals from sound, seismic, thermal, etc. sensors, in which the signal power has a pronounced maximum depending on the distance from the target to the network node. The results obtained extend to the problem of tracking multiple targets, which involves assessing the applicability of identification and classification methods in conditions where there is an overlap of perceived signals by sensors. The architecture of a system for joint data processing in distributed sensor networks is considered. The subject of consideration are methods for tracking a plurality of moving objects in a protected area. The main steps in the tracking procedure consist of: target detection, target classification, location estimates, and target movement trajectory prediction. Algorithms for solving these problems are discussed.

Keywords: sensor networks, detection, localization; classification, tracking

REFERENCES

1. Akuldiz I.F. Besprovodnyye sensornyye seti: obzor [Wireless sensor networks: a review], URL: <http://www.sciencedirect.com/science/article/pii/S1389128601003024>, accessed: 09.01.2021.
2. A.R. Al-Ali, Ragini Gupta, Ahmad Al Nabulsi Cyber physical systems role in manufacturing technologies. AIP Conference Proceedings 1957, 2018, URL: <https://aip.scitation.org/doi/pdf/10.1063/1.5034337>.
3. L. P. Clare, G. J. Pottie, and J. R. Agre Self-Organizing Distributed Sensor Networks. Proc. SPIE, vol. 3713, April 199, pp. 229-238.
4. M. Chu, H. Haussecker, F. Zhao Scalable Information-Driven Sensor Querying and Routing for Ad Hoc Heterogeneous Sensor Networks. International Journal on High Performance Computing Applications, Special issue on Sensor Networks, 2002.
5. E. N. Onggosanusi, A. M. Sayeed, B. D. Van Veen Canonical Space-Time Processing for Wireless Communications. IEEE Trans. Commun., pp. 1669-1680, Oct. 2000.
6. R. Brooks, C. Griffin, and D. S. Friedlander Self-Organized distributed sensor network entity tracking. International Journal of High Performance Computing Applications, Special Issue on Sensor Networks, 2002.
7. R. Brooks and C. Griffin Traffic model evaluation of ad hoc target tracking algorithms. International Journal of High Performance Computing Applications, Special Issue on Sensor Networks, 2002.
8. J. C. Chen, R. E. Hudson, and K. Yao A Maximum likelihood parametric approach to source localization. Proc. ICASSP'2001, Salt Lake City, UT, 2001, pp. 1043-1046.
9. L. M. Kaplan, Q. Le, and P. Molnar Maximum likelihood methods for bearings-only target localization. Proc. ICASSP'2001, Salt Lake City, UT, 2001, pp. 554-557.
10. P. W. Boettcher and G. A. Shaw A Distributed Time-Difference of Arrival Algorithm for Acoustic Bearing Estimation. Proc. 4th International Conf. on Information Fusion, vol. 1, Montreal, August 2001.
11. A. M. Sayeed and D. L. Jones Optimal Detection Using Bilinear Time-Frequency and Time-Scale Representations. IEEE Transactions on Signal Processing, vol. 43, no. 12, pp. 2872-2883, December, 1995.

12. K. Yao, R. E. Hudson, C. W. Reed, D. Chen, and F. Lorenzelli Blind Beamforming on a Randomly Distributed Sensor Array System. *IEEE Journal on Selected Areas in Communications*. Special Issue on Signal Processing for Wireless Communications, vol. 16, no. 8, pp. 1555-1567, Oct. 1998.
13. E.A. Lee, S.A. Seshia A cyber-physical systems approach. *Introduction to embedded systems*, 2011
14. J. F. Scholl, J. R. Agre, L. P. Clare, and M. C. Gill A Low Power Impulse Signal Classifier Using the Haar Wavelet Transform. *Proc. SPIE*, vol. 3577, Nov. 1998, pp. 136-145.
15. J.M. Hsua, C.C. Chenb, C.C. Li POOT: An efficient object tracking strategy based on short-term optimistic predictions for face-structured sensor networks. *Elsevier International Journal of Computers and Mathematics with Applications*, vol. 63, issue 2, pp.391-406; 2012.
16. Paulo Tabuada Cyber-physical systems: Position paper. University of California at Los Angeles, 2010.
17. Reva I.A., Bogdanov A.A., Malakhova Ye.A. Primeneniye toчек dostupa WI-Fi dlya registratsii dvizheniya na ob"yekte [Application of WI-Fi access points for motion registration at the facility]. *Sistemy analiza i obrabotki dannykh* [Systems of data analysis and processing]. 2017, no. 3 (68), pp. 104–125.
18. Xu Yingqi, Winter Julian, Lee Wang-Chien Prediction-based strategies for energy saving in object tracking sensor networks. *Proceedings - 2004 IEEE International Conference on Mobile Data Management*, 346- 357, 10.1109/MDM.2004.1263084.
19. Volkov A.N., Sivers M.A., Sukhov V.A. Pozitsionirovaniye v setyakh Wi-Fi [Positioning in Wi-Fi networks]. *Vestnik svyazi* [Communication Bulletin], 2010.
20. Dan Li, Kerry D. Wong, Yu H. Hu, Akbar M. Sayeed. Detection, Classification and Tracking of Targets in Distributed Sensor Networks, URL: <http://www.ecce.wisc.edu/~sensit>, accessed: 09.01.2021.
21. Kucheryavyy A.Ye., Prokof'yev A.V., Kucheryavyy Ye.A. Samoorganizuyushchiyesya seti [Self-organizing networks]. SPb.:Lyubavich, 2011, 312 p
22. H. V. Poor. *An Introduction to Signal Detection and Estimation*. Springer Verlag, 1988.
23. C. W. Reed, R. Hudson, and K. Yao Direct Joint Source Localization and Propagation Speed Estimation. *Proc. IEEE ICASSP*, March 1999, pp. 1169-1172.
24. Martin Mikusz Towards an understanding of cyber-physical systems as industrial software-product-service systems. *Conference paper in Procedia CIRP*, 2014.
25. Vinogradov G.P., Yemtsev A.S., Fedotov I.S. Besprovodnyye sensornyye seti v zashchishcha-yemykh zonakh [Wireless sensor networks in protected areas]. *Izvestiya YUFU. Tekhnicheskoye nauki*, 2021, no. 1, pp. 19-30.
26. Vinogradov G.P., Mal'kov A.A. Modeli poiska struktur dannykh na osnove konkurentsii i kooperatsii [Search models for data structures based on competition and cooperation]. *Sb. trudov "Upravleniye bol'shimi sistemami"*. M.: IPU RAN, 2008, Issue. 22, pp. 98-110.

Gennady P. Vinogradov, Doctor of Technical Sciences, Professor, Head. Laboratory "Intelligent Systems", Research Institute "Tsentrprogrammsistem", wgp272ng@mail.ru, Tver, Russia, N. Korytkov Ave., 3a

Статья поступила в редакцию 18.03.2022; одобрена после рецензирования 28.03.2022; принята к публикации 31.03.2022.

The article was submitted 03.18.2022; approved after reviewing 03.28.2022; accepted for publication 03.31.2022.

Цифровые двойники

УДК 519.876.5

doi:10.38028/ESI.2022.25.1.006

Интеграция концептов цифровых двойников и ситуационной осведомленности в системе моделирования критических инфраструктур

Фридман Александр Яковлевич

Институт информатики и математического моделирования ФИЦ КНЦ РАН,
Россия, Апатиты, fridman@iimm.ru

Аннотация. Рассмотрена интеллектуальная информационная технология ситуационного моделирования промышленно-природных систем (ППС) различного масштаба в целях стратегического и оперативного управления ими, ориентированная на комплексную обработку разнородной информации о статических и динамических характеристиках подобных систем. Актуальность задачи вызвана: усложнением и взаимопроникновением ППС; ростом неопределенностей в мире из-за ускорения развития ППС; увеличением доступности данных в связи с рождением Интернета Вещей. Новизна подхода состоит в интеграции средств изучения слабо формализованных сложных нестационарных пространственных объектов в рамках причинно-следственных моделей и комплексном использовании экспертных знаний для формирования критериев и выбора предпочтительных вариантов реализации структуры ППС, изучаемых более детально в имитационном режиме с помощью цифровых двойников компонентов ППС.

Ключевые слова: промышленно-природная система комплекс, стратегическое и оперативное управление, количественная оценка ситуационной осведомленности, превентивная аналитика в задачах безопасности, цифровой двойник

Цитирование: Фридман А.Я. Интеграция концептов цифровых двойников и ситуационной осведомленности в системе моделирования критических инфраструктур // Информационные и математические технологии в науке и управлении. – 2022. – № 1 (25). – С. 70-78. – DOI:10.38028/ESI.2022.25.1.006.

Введение. В статье анализируются возможности повышения гибкости и эффективности моделирования нестационарных распределенных объектов, параметры которых меняются во времени и пространстве. Будем называть их промышленно-природными системами, поскольку они содержат как промышленные, а так и природные компоненты. Характеристики первых обычно требуется оптимизировать с помощью управляющих сигналов (решений разного уровня и масштаба), а условия жизни вторых формируют ограничения на допустимые параметры первых.

Такая задача актуальна всегда, но возможности ее решения непрерывно меняются с развитием технологий конструирования и методов управления. В последнее время новые подходы к решению указанной задачи связаны с Интернетом вещей (Internet of Things – IoT). Этот фактор привел к экспоненциальному росту количества данных о параметрах ППС, чаще всего в виде временных рядов. С другой стороны, наблюдающееся усложнение и взаимопроникновение (по функциональным характеристикам и географическим координатам) ППС существенно повышает риск возникновения и ущерб от последствий комплексных (зависимых [1]) отказов, вызывающих самые значительные последствия.

Методы моделирования сложных объектов чаще всего оценивают только вероятностные значения их параметров (энтропии, средних значений и других статистических свойств),

либо характеристики графов таких объектов (достижимость состояний, связность и т.п.). Однако в задачах, связанных с безопасностью, такие подходы неэффективны, поскольку инициирующие события, приводящие к нештатным и чрезвычайным ситуациям (НЧС), возникают при конкретных сочетаниях значений параметров и переменных на объекте, которые можно выявить только на причинно-следственных моделях, анализирующих реальные характеристики объекта. Аналитические модели редко адекватны сложным системам, для них наиболее пригодно имитационное моделирование. К первым моделям такого класса можно отнести системную динамику, но они не поддерживают иерархическую структуру описания объекта и поэтому становятся неконтролируемыми при росте размерности модели.

Для исследования иерархий предложены модели с различными принципами декомпозиции (в скобках даны названия соответствующего программного обеспечения каждого типа декомпозиции): функциональная (Structures Analysis and Design Technique – SADT [2]), организационная или структурная декомпозиция (Data Flow Diagram – DFD [3]), декомпозиция по состояниям (State Transition Diagram – STD [4]). Моделирование сложных объектов требует использовать все перечисленные типы декомпозиции совместно, что их программное обеспечение не поддерживает. Для решения этой задачи разрабатывался универсальный объектно-ориентированный язык моделирования Universal Modelling Language – UML [5], но в его среде требуется встраивать все программные модули в диаграмму классов и писать их тексты на этом языке, что не позволяет использовать интеллектуальные методы обработки данных и интегрировать в модель ранее разработанные модули.

1. Ситуационная система моделирования (ССМ). Видимо, самый близкий аналог рассматриваемой ниже ситуационной системы моделирования [6, 7] – интеллектуальная система имитационного моделирования «Ресурс-Действие-Объект» (РДО) [8], предназначенная для моделирования сложных систем, функционирование которых требуется представить в форме продукционных правил. У РДО и ССМ есть следующие общие аспекты: все переменные рассматриваются как дискретные; имитация производится в рамках дискретно-событийного метода; можно строить иерархическое описание объекта и подключать внешние программные модули, имитирующие работу компонентов объекта (в последнее время их называют «цифровыми двойниками» [9]). Основные минусы РДО для решения задачи исследования ППС таковы:

- не для всех объектов пригодна модель в виде системы продукций и моделирование с помощью логического вывода, особенно при высокой размерности модели;
- отсутствуют встроенные механизмы сравнения альтернатив реализации ППС;
- иерархические структуры модели не соответствуют структуре объекта: они предназначены либо для увеличения скорости логического вывода путем классификации продукционных правил, либо реализуют систему классов в стиле объектно-ориентированного программирования;
- РДО не позволяет работать с географическими параметрами объекта.

Все перечисленные и другие существующие системы моделирования имеют общий недостаток при решении рассматриваемой здесь задачи исследования ППС: они не предоставляют возможности эффективной реализации идей ситуационного подхода, который, по мнению автора, наиболее адекватен этой задаче.

Указанных недостатков лишена представленная ниже ССМ, разработанная в рамках работ М. Месаровича по иерархическим системам [10] и ситуационного управления, предложенного Д.А. Поспеловым [11]. Эта конкретизация ситуационного подхода для задачи моделирования ППС создавалась при участии и под руководством автора, ее отличия от прототипов заключаются в следующем:

- ядром ССМ является ситуационная концептуальная модель (СКМ) ППС, отражающая организационную структуру объекта и взаимодействия между его составными частями. Это иерархическая модель, описывающая все перечисленные выше типы декомпозиции, она допускает оперативную модификацию на любом этапе жизненного цикла, позволяет в деталях контролировать консистентность реализации ППС, осуществлять имитацию сценариев развития ППС, автоматизировать генерацию вычислительных экспериментов, интегрировать подмодели различных типов, в том числе экспертной системы и геоинформационной системы (ГИС);
- в отличие от всех аналогов, в ССМ даны формальные определения главных концептов ситуационного управления [11]: ситуации, сценария, пополнения и обобщения ситуаций и т.д., что позволило автоматизировать основные задачи моделирования: формирование ситуаций и сценариев, их сопоставление в статике и в режиме имитации с выбором на этой основе такой структуры ППС, которая соответствует предпочтениям ЛПР, а также генерацию «быстрых прототипов» в традициях ИИ для поддержки принятия решений, касающихся составных частей ППС.

СКМ относится к причинно-следственным моделям, в частности, в ней не применяется нечеткая логика, использованная в некоторых системах ситуационного управления. Это позволяет компьютеризировать вычислительные эксперименты по поиску «узких мест» в структуре объекта, которые порождают наиболее опасные зависимые отказы, появляющиеся при определенных комбинациях параметров ППС, а также анализировать чувствительность предполагаемых решений к вариациям внешних воздействий и внутренних характеристик системы. С этой же целью, НЧС моделируются как расширение модели нормальной работы ППС, с применением тех же процедур исполнения расчетов.

2. Основные особенности ССМ. Модель ППС в ССМ [7] включает элементы трех типов: объекты, процессы и ресурсы. Первый этап создания СКМ – это построение *дерева объектов* – компонентов ППС. Оно изоморфно организационной структуре ППС, а также формирует пространственные характеристики объектов: при конструировании СКМ один или более объектов из каждой ветви дерева сопоставляются с электронной картой средствами встроенной в ССМ ГИС. В результате всегда имеется взаимно-однозначная связь СКМ и отображения ППС на карте, что позволяет учитывать его параметры наряду с другими в ходе моделирования. *Ресурсы* отображают взаимодействия объектов и формализуют любые (как материальные, так и информационные) связи между объектами и приписанными к ним *процессами*, каждый из которых описывает преобразование определенного набора входных ресурсов этого процесса в набор выходных ресурсов. При формировании дерева объектов и назначении связей по ресурсам в СКМ вводятся возможные альтернативы структуры ППС путем декомпозиции объектов либо указанием альтернативных путей передачи наборов ресурсов между процессами и/или объектами. Компьютерная спецификация элементов СКМ определяется назначенными им *исполнителями*, которые для процессов задают способ доступа к имитатору (*цифровому двойнику*) процесса, для ресурсов определяют процедуру получения временного ряда значений, а для объектов организуют доступ к их графическим параметрам. Вызов исполнителя процесса при имитации переводит СКМ из одного состояния в другое, последовательность таких вызовов формирует конкретный сценарий имитации.

Математическое описание СКМ и ССМ, формулировки и доказательства теорем здесь не приводятся из-за ограниченного объема статьи, все это имеется в работах, на которые даны ссылки.

Ситуационный подход в СКМ опирается на следующие концепты. Атом данных в модели именуется *фактом*, он содержит имя и набор допустимых или запрещенных значений

одного ресурса. *Исходная ситуация* формируется пользователем в начале сеанса моделирования в виде набора интересующих его фактов, она рассматривается как цель анализа в данном сеансе. С помощью связей, заданных в СКМ, и правил пополнения ситуаций, которые хранятся во встроенной в ССМ экспертной системе (ЭС), исходная ситуация автоматически расширяется до тех пор, пока не будет найден связный фрагмент СКМ, включающий все ресурсы из исходной ситуации. Он именуется *полной ситуацией* (ПС) и трактуется как зона интересов ЛПР в данном сеансе. Объект принятия решения (ОПР) соответствует корневому объекту найденной ПС, его положение в дереве объектов задает организационный уровень решения этой ПС. Любая ПС может включать ранее заданные альтернативы и тем самым определять набор возможных в этой ПС структур ППС. Каждая такая (не включающая вариантов) допустимая структура называется *достаточной ситуацией* (ДС), так как она содержит конкретную структуру ППС, достаточную для корректного задания по крайней мере одного шага имитации. Если в некоторой ДС ЛПР принимает решение перевести модель из текущего класса ситуаций в новый, такая ДС называется *управляющей*. В таком случае ССМ предлагает ЛПР предпочтительную структуру из выбранного им класса путем статического сопоставления ситуаций для одного и того же ОПР. С этой целью при построении СКМ каждому объекту приписывается обобщенный критерий качества (ОКК) [7], представляющий собой взвешенную среднеквадратическую сумму относительных девиаций скалярных критериев качества (СКК) этого объекта от их номиналов, которые вводит ЛПР верхнего (относительно ОПР) уровня. Он же вводит допустимые отклонения СКК от номиналов, обратные значения этих допусков используются как веса в ОКК, то есть как коэффициенты важности СКК (чем данный СКК важнее для ЛПР, тем меньше допуск на его отклонение, что логично). Обобщенный критерий может стать равным нулю только при полном совпадении всех критериев с номиналами, то есть при идеальном для вышестоящего ЛПР функционировании ОПР). У такого ОКК есть еще одно полезное свойство: он равен 1, когда девиации всех СКК равны допускам, и существенно больше 1 в недопустимых режимах работы ОПР. Используя это свойство, легко найти в дереве объектов объект самого нижнего уровня с большим значением ОКК, именно он служит источником проблем для ОПР. Для статического сопоставления ДС одного ОПР используются обобщенные затраты на удовлетворение его ОКК, равные удельной чувствительности ОКК к изменениям СКК. Эти затраты дают основу для *классификации ситуаций*. В один класс входят те ДС, где слагаемое с определенным СКК минимально в ОКК, ДС из одного класса тем лучше, чем меньше вклад этого скалярного критерия в ОКК. Если статическое сравнение не дает однозначного результата, последствия реализации каждой ДС можно исследовать в режиме имитации, задание на которую генерируется автоматически путем последовательного подключения исполнителей элементов выбранного фрагмента модели и синтеза баз исходных данных и результатов имитации. *Сценарий* есть последовательность ДС для одного ОПР при одинаковых исходных данных.

Представленный выше ОКК показал свою эффективность также при решении в рамках ССМ задач планирования и координации управлений [12].

Теперь рассмотрим новые модификации ССМ в рамках темы настоящей работы.

3. Превентивная аналитика безопасности ППС. Для решения любых задач сопоставления ситуаций прежде всего необходимо унифицировать процедуры обработки переменных различной природы. Начальным шагом для этого может быть принятое с самого начала разработки ССМ решение рассматривать области значений любых параметров как списковую структуру. Для числовых переменных упорядочение этих списков естественным образом задается отношением «больше или равно», для других типов переменных требуется дополнительная работа. Первым делом нужно выбрать одно (в общем случае – любое) допустимое значение как номинальное ($З_{ид}$ на рис. 1). При реализации этого значения соответствующее

ему управление считается идеальным. Затем остальные значения упорядочиваются по степени их отличия от номинального: чем больше это отличие для некоторого значения параметра, тем дальше от номинального оно должно находиться в итоговом списке. На рис. 1 имена этих значений взяты в угловые скобки. При этом все значения, находящиеся по одну сторону от идеального, должны иметь тот же «знак» отклонения от него.

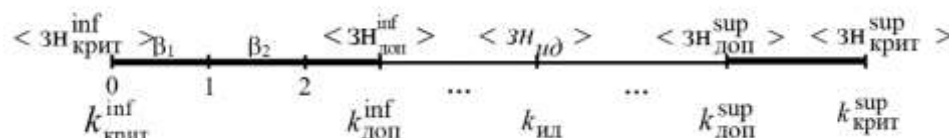


Рис. 1. Шкала значений параметра

Чтобы унифицировать процедуры анализа работы ППС в нормальных режимах и при НЧС [13], нежелательные и опасные значения параметров «дополняют» допустимые значения (от $<3H_{доп}^{inf}>$ до $<3H_{доп}^{sup}>$ на рис. 1) в соответствующие стороны. Неотрицательные числа показывают порядковые номера значений в стеке (списке) области значений параметра, их буквенные обозначения – символы k . Жирными отрезками оси выделены подмножества недопустимых значений параметра. Будем (без потери общности) считать, что все значения параметра за границами допуска, кроме крайних значений, нежелательны, но не критичны, а крайние значения ($<3H_{крит}^{inf}>$ и $<3H_{доп}^{sup}>$) соответствуют НЧС.

Для примера на рис. 2 показан безопасный диапазон значений строкового параметра «рост». При необходимости учета критических значений этот диапазон можно слева дополнить значением «карлик», а справа – значением «акромегал».

Если нельзя считать значения параметра равноудаленными, то при определении расстояния между ними можно учитывать экспертные положительные веса дуг между соседними значениями, показанные символами β с индексами на рис. 2.



Рис. 2. Допустимые значения параметра «рост»

В работе [13] предложены соотношения для вычисления относительных девиаций параметров в представленном выше обобщенном пространстве параметров таким образом, чтобы сохранить все положительные свойства ОКК, в том числе равенство его значения 1 при нахождении всех параметров на границах их допусков.

После описанной модификации в обобщенное пространство состояний ППС можно вводить любые типы параметров наравне с числовыми переменными и применять унифицированные процедуры анализа интегральных свойств поведения ППС в этом пространстве, которое используется при решении всех задач моделирования ППС.

Для выявления предпочтительных сценариев работы ППС в ходе стратегического моделирования (при разработке новых ППС) применяются следующие определения.

Определение 1. Сценарий *допустим* на определенном промежутке времени, если не приводит к критическим величинам компонентов состояния ППС, иначе он *запрещенный*.

Определение 2. Сценарий *безопасен* на определенном промежутке времени, если не приводит к появлению опасных значений компонентов состояния ППС, иначе он *опасный*.

Определение 3. Предпочтительность опасных сценариев тем выше, чем меньше величина суммы обобщенных затрат, превышающих единицу, либо степень отличия для опасных ситуаций от ближайших допустимых.

Определение 4. Сценарий *оптимален* на определенном промежутке времени при минимальной (среди всех допустимых сценариев) сумме обобщенных затрат на этом промежутке.

С учетом определений 1 – 4, цель исследования нормальной работы ППС заключается в поиске сценария, оптимального на заданном временном интервале при отсутствии в этом сценарии опасных значений состояния ППС, для чего необходимо наличие непустых отображений из текущей ДС в следующую на всех шагах сценария.

Моделирование НЧС требует перебора доступных по ресурсам мероприятий для нахождения такого, которое приводит к предпочтительному из опасных сценариев без появления критических значений состояния ППС.

Разумеется, нет возможности анализировать сценарии на реальных объектах, необходимо компьютерное моделирование с применением цифровых двойников таких объектов.

В задачах безопасности *ситуационные цифровые двойники* (СЦД), пригодные в качестве исполнителей процессов в ССМ, должны решать две основные задачи: корректное моделирование опасных воздействий новых ППС на окружающие ППС и обратно – реакции новых систем на опасные воздействия со стороны окружающих ППС. По этим критериям и должны разрабатываться и сертифицироваться СЦД. При появлении новых ППС может потребоваться и доработка СЦД уже существующих ППС, поскольку новым объектам могут быть опасны и те параметры существующих ППС, которые не были опасны для других ППС.

4. Ситуационная осведомленность в ППС. Парадигму ситуационной осведомленности (СО) [14] можно сегодня считать наиболее общим подходом к применению ситуационного подхода в задачах, связанных с нестационарными системами. Авторы этой концепции выделяют три уровня или этапа достижения СО: восприятие происходящего вокруг, понимание происходящего и прогноз будущего состояния. Концепция СО исходно разрабатывалась для операторов в таких видах деятельности, где количество информации велико, а последствия неверных решений серьезны (пилотирование, военные конфликты, скорая помощь и т.д.), однако, на наш взгляд, измерение СО полезно и в рассматриваемых проблемах исследования ППС, где фактор времени не так значителен, но многие аспекты СО важны и могут быть эффективно реализованы для СКМ следующим образом:

- восприятие моделируется заданием исходной ситуации, понимание – формированием полной ситуации и определением организационного уровня решения задачи, прогнозирование – имитационным экспериментом;
- цель моделирования задается ЛПР путем выбора желательного класса ситуаций в ходе моделирования;
- специфика информации учитывается выбором исходной ситуации и автоматической отбраковкой неперспективных альтернатив;
- ожидания ЛПР в ССМ формализуются выбором доминирующего частного критерия в ОКК и составом допустимых альтернатив, заложенных в СКМ при ее создании.

В результате впервые предложены числовые оценки СО для ССМ [15].

Заключение. Стратегическое моделирование ППС направлено на анализ структуры проектируемых потенциально опасных объектов и их взаимосвязей для выявления возможностей зависимых отказов, которые порождают цепочки событий, приводящие к самым значительным ущербам в нештатных и чрезвычайных ситуациях. Этот режим в ССМ использует ситуационные цифровые двойники ППС и решает задачу превентивной аналитики для поиска и сценарного прогнозирования «узких мест» в структуре ППС с возможностью оценки ущер-

бов от отказов и обоснования рациональных мер по их предупреждению. *Для повышения безопасности целесообразно до встраивания новых устройств и систем в реальный мир встроить их модели в мир виртуальный.*

Предложена метрика обобщенного пространства состояний, включающего как строковые, так и числовые переменные. Описано применение этой метрики в системе ситуационного моделирования статических и динамических характеристик промышленно-природных систем с учетом аспектов безопасности. Метрика допускает использование экспертных знаний о степени опасности различных состояний объекта моделирования.

Для оперативного моделирования ППС предложены количественные оценки ситуационной осведомленности ЛПР, участвующих в работе ППС, путем метризации пространства всех трех основных аспектов СО, что позволяет объективизировать оценку достигнутой степени СО и особенно важно, когда различные составные части одного ППС входят в зоны ответственности нескольких ЛПР. Для этого случая предложенный подход позволяет координировать действия ЛПР и предотвратить конфликты между ними.

ППС рассматриваются как иерархические системы, при росте их масштаба возможен переход к сетцентрической модели [16], основное отличие процесса моделирования от исследования иерархических структур в таком случае состоит в том, что зоны ответственности ЛПР формируются динамически по результатам оценки их СО.

СПИСОК ИСТОЧНИКОВ

1. ГОСТ 27.002-89: Надежность в технике. Основные понятия. Термины и определения.
2. Nam Pyo Suh. Axiomatic Design - Advances and Applications. Chapter 5, New York, Oxford University Press, 2007, p. 239-298.
3. Коваленко В.В. Проектирование информационных систем / В.В. Коваленко. – М.: Форум, 2011. – 357с.
4. Joe Celko's. Trees and Hierarchies in SQL for Smarties. The Morgan Kaufmann Series in Data Management Systems. Second Edition, Elsevier Inc., 2012.
5. Hunt J. The Unified Process for Practitioners: Object-oriented Design, UML and Java. Springer, 2000.
6. Фридман А.Я., Курбанов В.Г. Ситуационное моделирование надежности и безопасности промышленно-природных систем / А.Я. Фридман, В.Г. Курбанов // Информационно-управляющие системы № 4(71), 2014. – с. 1-10.
7. Фридман А.Я., Курбанов В.Г. Формальная концептуальная модель промышленно-природного комплекса как средство управления вычислительным экспериментом / А.Я. Фридман, В.Г. Курбанов // Труды СПИИРАН, 2014. № 6 (37), – с. 424–453.
8. Ясиновский С.И. Интеллектуальная гибридная система поддержки принятия решений РДО: вчера, сегодня, завтра / С.И. Ясиновский // Мягкие измерения и вычисления. № 12 (13), 2018. – с. 33-50.
9. Digital Twin: Transforming How We Make Sense of Data. URL: <https://www.ptc.com/ru/industry-insights/digital-twin> (accessed 11.03.2022).
10. Месарович М., Мако Д., Такахара И. Теория иерархических многоуровневых систем / М. Месарович, Д. Мако, И. Такахара. – М.: Мир, 1973. – 344с.
11. Поспелов Д.А. Ситуационное управление: теория и практика / Д.А. Поспелов. – М.: Наука, 1986. – 228с
12. Фридман А.Я. Координация и планирование управлений в локально организованных иерархических системах // Шестая Международная конференция «Системный анализ и информационные технологии» САИТ-2015 (15-20 июня 2015 г., г. Светлогорск, Россия): Труды конференции. В 2-х т, Т. 1. – М.: ИСА РАН, 2015. – с. 115-124.

13. Фридман А.Я. Экспертное пространство для ситуационного моделирования промышленно-природных систем / А.Я. Фридман // Вестник Московского университета им. С.Ю. Витте, № 1 (4), 2014. – с. 233-245.
14. Endsley M.R. Final Reflections: Situation Awareness Models and Measures. J. of Cognitive Engineering and Decision Making. Vol. 9, no. 1, 2015, pp. 101-111.
15. Fridman A., Kulik B. Situation Awareness in Modeling Industrial-Natural Complexes. Intern. Conf. Cyber-Physical Systems and Control (CPS&C'2019, St. Petersburg). In: Lecture Notes in Networks and Systems. Cyber-Physical Systems and Control. Vol. 95, 2020, pp. 247-256.
16. Fridman A., Oleynik A. Modelling of Situation Awareness in Net-Centric Commercial Systems. Proc. 34th Eur. Simulation and Modelling Conf. (ESM-2020) Toulouse, France, Oct. 21-23, 2020, pp. 64-67.

Фридман Александр Яковлевич, д.т.н., профессор, в.н.с. лаборатории управления промышленно-природными системами, Институт информатики и математического моделирования ФИЦ КНЦ РАН, fridman@iimm.ru, Россия, Апатиты Мурманской области, ул. Ферсмана 24А

UDC 519.876.5

doi:10.38028/ESI.2022.25.1.006

Incorporation of concepts of digital twins and situational awareness in the modelling system of critical infrastructures

Alexander Y. Fridman

Institute for Informatics and Mathematical Modelling, FRC KSC RAS,
Russia, Apatity, fridman@iimm.ru

Abstract. An intelligent information technology for situational modelling of industrial-natural complexes (INCs) of various scales for the purpose of strategic and operative management of them, focused on complex processing of heterogeneous information regarding static and dynamic characteristics of such systems, is considered. The relevance of this task results from the following: complication and interpenetration of INCs; growth of uncertainties in the world due to acceleration of INCs development; increasing data availability due to the birth of the Internet of Things. The novelty of the proposed approach lies in integration of means for studying poorly formalised complex non-stationary spatial objects within the framework of cause-and-effect models and integrated usage of expert knowledge to form criteria and select preferable options for implementing INCs structure, which are studied in more detail in a simulation mode with using digital twins of INCs components.

Keywords: industrial-natural complex, strategic and operative management, quantitative assessment of situational awareness, preventive analytics in safety tasks, digital twin

REFERENCES

1. GOST 27.002-89: Nadezhnost' v tehnikе. Osnovnye ponjatija. Terminy i opredelenija [Reliability in engineering. Basic concepts. Terms and Definitions].
2. Nam Pyo Suh. Axiomatic Design – Advances and Applications. Chapter 5, New York, Oxford University Press, 2007, p. 239-298.
3. Kovalenko V.V. Proektirovanie informacionnyh sistem [Design of information systems]. Moscow, Forum, 2011, 357p.
4. Joe Celko's. Trees and Hierarchies in SQL for Smarties. The Morgan Kaufmann Series in Data Management Systems. Second Edition, Elsevier Inc., 2012.

5. Hunt J. The Unified Process for Practitioners: Object-oriented Design, UML and Java. Springer, 2000.
6. Fridman A.Ja., Kurbanov V.G. Situacionnoe modelirovanie nadezhnosti i bezopasnosti promyshlenno-prirodnih sistem [Situational modeling of reliability and safety of industrial and natural systems]. Informacionno-upravljajushhie sistemy [Information and control systems], 2014, no. 4(71), p. 1-10.
7. Fridman A.Ja., Kurbanov V.G. Formal'naja konceptual'naja model' promyshlenno-prirodnogo kompleksa kak sredstvo upravlenija vychislitel'nyh jeksperimentom [Formal conceptual model of an industrial-natural complex as a means of managing a computational experiment]. Trudy SPIIRAN [Proceedings of SPIIRAS], 2014, no. 6 (37), p. 424–453.
8. Jasinovskij S.I. Intellektual'naja gibridnaja sistema podderzhki prinjatija reshenij RDO: vchera, segodnja, zavtra [Intelligent hybrid decision support system RDO: yesterday, today, tomorrow]. Mjagkie izmerenija i vychislenija [Soft Measurements and Computing], 2018, no. 12 (13), p. 33-50.
9. Digital Twin: Transforming How We Make Sense of Data. URL: <https://www.ptc.com/ru/industry-insights/digital-twin> (accessed 11.03.2022).
10. Mesapovich M., Mako D., Takahapa I. Teorija ierarhicheskikh mnogourovnevnyh sistem [Theory of hierarchical multilevel systems]. Moscow, Mir, 1973, 344p
11. Pospelov D.A. Situacionnoe upravlenie: teorija i praktika [Situational management: theory and practice]. Moscow, Nauka, 1986, 228p.
12. Fridman A.Ja. Koordinacija i planirovanie upravlenij v lokal'no organizovannyh ierarhicheskikh sistemah [Coordination and planning of management in locally organized hierarchical systems]. Shestaja Mezhdunarodnaja konferencija «Sistemnyj analiz i informacionnye tehnologii» SAIT-2015 (15-20 ijunja 2015 g., g. Svetlogorsk, Rossiya): Trudy konferencii. V 2-kh t, T. 1. [Sixth International Conference "System Analysis and Information Technologies" SAIT-2015 (June 15-20, 2015, Svetlogorsk, Russia): Proceedings of the conference. In 2 volumes, vol. 1]. Moscow, ISA RAN, p. 115-124.
13. Fridman A.Ja. Jekspertnoe prostranstvo dlja situacionnogo modelirovanija promyshlenno-prirodnih sistem [Expert space for situational modeling of industrial and natural systems]. Vestnik Moskovskogo universiteta im. S.Ju. Vitte [Bulletin of the Moscow University named after S.Yu. Witte], no.1 (4), 2014, p. 233-245.
14. Endsley M.R. Final Reflections: Situation Awareness Models and Measures. J. of Cognitive Engineering and Decision Making. Vol. 9, №. 1, 2015, pp. 101-111.
15. Fridman A., Kulik B. Situation Awareness in Modeling Industrial-Natural Complexes. Intern. Conf. Cyber-Physical Systems and Control (CPS&C'2019, St. Petersburg). In: Lecture Notes in Networks and Systems. Cyber-Physical Systems and Control. Vol. 95, 2020, pp. 247-256.
16. Fridman A., Oleynik A. Modelling of Situation Awareness in Net-Centric Commercial Systems // Proc. 34th Eur. Simulation and Modelling Conf. (ESM-2020) Toulouse, France, Oct. 21-23, 2020, pp. 64-67.

Alexander Y. Fridman, Doctor of Technical Sciences, Professor, Leading Researcher at the Laboratory for Industrial-Natural Systems Control, Institute for Informatics and Mathematical Modelling, FRC KSC RAS, fridman@iimm.ru, Russia, Apatity, Murmansk region, st. Fersmana 24A

Статья поступила в редакцию 17.02.2022; одобрена после рецензирования 25.03.2022; принята к публикации 28.03.2022.

The article was submitted 02.17.2022; approved after reviewing 03.25.2022; accepted for publication 03.28.2022.

УДК 004.89

doi:10.38028/ESI.2022.25.1.007

Разработка цифрового двойника ветровой электростанции: постановка задачи и проектирование

Массель Людмила Васильевна, Массель Алексей Геннадьевич, Щукин Никита Игоревич, Цыбиков Алексей Ринчинович

Институт систем энергетики им. Л.А. Мелентьева СО РАН,
Россия, Иркутск, niksha14@mail.ru

Аннотация. Статья посвящена разработке цифрового двойника (ЦД) ветровой электростанции. Выполнен обзор существующих решений в этой области. Подробно рассматривается подход к построению ЦД, основанный на онтологическом инжиниринге. Описаны основные шаги онтологического инжиниринга. Приводится фрагмент системы онтологий ветровых систем. Рассмотрена математическая модель для определения параметров работы ветроэнергетической установки, используемая при построении ЦД. В статье также уделено внимание проектированию архитектуры ЦД, включающей цифровую тень, цифровую модель и систему управления. Так же в статье приводится этап реализации прототипа ЦД ветровой электростанции.

Ключевые слова: ветроэнергетика, цифровой двойник, онтологии, базы данных

Цитирование: Массель Л.В., Массель А.Г., Щукин Н.И., Цыбиков А.Р. Разработка цифрового двойника ветровой электростанции: постановка задачи и проектирование // Информационные и математические технологии в науке и управлении. – 2022. – № 1 (25). – С. 79-90. – DOI:10.38028/ESI.2022.25.1.007.

Введение. Сейчас активно ведутся работы по цифровизации энергетики, обусловленные принятием Энергетической стратегия РФ на период до 2035 года [1]. Все чаще поднимаются вопросы применения возобновляемых источников энергии, в частности, с использованием энергии ветра для производства электроэнергии. Многие европейские страны активно используют ветроэнергетические электростанции.

Исследования показывают, что в большинстве случаев использование возобновляемых источников электроэнергии экономически целесообразно. При этом, учитывая нестабильность производства электроэнергии ветровыми электростанциями, целесообразно рассматривать также накопители электрической энергии. Актуальны также управление электропотреблением и активизация роли потребителей в этом процессе. Для рассмотрения всех вышеперечисленных задач целесообразно использовать технологию цифровых двойников.

В Российской Федерации актуальность использования возобновляемых источников энергии (ВИЭ) обуславливается тем, что значительная часть территории России не охвачена централизованным электроснабжением. Районы децентрализованного энергоснабжения занимают около 60% площади Российской Федерации и находятся в основном в северных регионах страны [2]. В этих районах расположено много небольших изолированных населенных пунктов. Их электроснабжение обеспечивается главным образом на базе дизельных электростанций, использующих дорогое привозное топливо.

В то же время при использовании ВИЭ возникает одна основная проблема – их поведение, в особенности солнечных и ветровых электростанций, очень сильно зависит от внешних факторов, в первую очередь от погодных условий. Соответственно, для эффективного использования ВИЭ необходимо разрабатывать не только системы управления, но и системы прогнозирования поведения этих объектов в изменяемых условиях. При цифровизации одной из используемых технологий является технология цифровых двойников. Цифровой

двойник является виртуальным прототипом реального объекта, с помощью которого можно проводить эксперименты и проверять гипотезы, прогнозировать поведение объекта и решать задачу управления его жизненным циклом [3].

Цифровой двойник позволяет снизить затраты на проектирование оборудования и системы в целом, а также позволяет сократить расходы на эксплуатацию за счет перехода от плановых ремонтов к текущим. Таким образом, можно сделать вывод, что использование цифровых двойников при проектировании ветровых электростанций (ВЭС) может положительно сказаться как на этапе проектирования, так и на этапе их эксплуатации.

В рамках этой статьи будут рассмотрены существующие работы в области построения ЦД ветровых электростанций, применение онтологического инжиниринга для проектирования ЦД ВЭС, предлагаемые информационная модель и архитектура ЦД ВЭС. В заключении описаны направления дальнейшего развития работы.

1. Анализ существующих решений в области применения цифровых двойников в ветроэнергетике. В условиях всеобщей конкуренции проблема перехода к высокотехнологичным и эффективным производствам в последние десятилетия оказалась особенно актуальной. Вследствие стремительного развития информационных технологий появилась возможность сбора, хранения, передачи и анализа больших массивов данных, собираемых с реальных объектов. Благодаря этому обнаружилась необходимость пересмотра стандартных подходов к управлению производственными процессами на предприятиях. Эти факторы побуждали в ряде стран разработку и принятие программ стратегического развития промышленности, таких, как «Платформа индустрия 4.0» (Германия), «Сделано в Китае 2025» (Китай), «Национальная технологическая инициатива» (Россия) и др. Все предложенные программы направлены на увеличение производительности труда, повышение экономической эффективности производств и внедрение современных наукоемких технологий [4].

Внедрение этих стратегий приводит к повышению уровня автоматизации предприятий и более широкой цифровизации производственных процессов. Подобные изменения обусловлены необходимостью быстро и точно моделировать продукт и его производственную технологию с целью увеличения рентабельности и экономии ресурсов в условиях современной конкуренции.

Решить эту задачу может помочь цифровой двойник, который объединяет как виртуальную среду предприятия (данные, поступающие с датчиков, математические и геометрические модели и др.), так и физическую (исполнительные механизмы, станки, схемы и др.), а также описывает процесс взаимодействия между этими средами и дополняет это технологиями автоматизации. В 2018 году компания «Gartner», являющаяся ведущей мировой исследовательской и консалтинговой компанией, включила цифровые двойники в список 10 стратегически важных технологий. Компания отмечала, что «цифровые двойники в контексте проектов Интернета вещей являются особенно многообещающими в течение следующих трех-пяти лет» [5].

В настоящее время технология цифровых двойников остается одной из самых многообещающих технологий, которую широко исследуют и пытаются применять [6-12].

Зарождение концепции цифровых двойников началось с авиакосмической промышленности. В одной из своих работ Глесген и Штаргель [13] приводят объяснение принципа использования цифровых двойников для сертификации транспортных средств и управления автопарком: «Цифровой двойник – это интегрированная мультифизическая, мультимасштабная, теоретико-вероятностная модель собранного транспортного средства или системы, в которой используются лучшие доступные физические модели, данные датчиков, а также история парка транспортных средств для моделирования состояния оригинала, работающего в реальных полевых условиях».

Можно отметить, что в настоящее время исследования и разработки в сфере цифровых двойников направлены на задачи функционирования [7-9, 11, 14-17], как, в общем, и наше текущее исследование.

Разработки и исследования в сфере цифровых двойников ветровых электростанций делятся на два направления: разработка цифровых двойников отдельных частей ветровой электростанции, в частности, разработка цифрового двойника структурной модели лопасти ротора ветрогенератора [9], и разработка цифрового двойника ветрогенератора [7, 8].

Основное отличие представленной работы заключается в том, что предложено использовать информацию, полученную на этапе решения задач функционирования объектов, для дальнейшего решения задач развития ветроэнергетических установок, интеграции их в общей изолированной энергосистеме на основе цифровых двойников, и последующей интеграции в едином цифровом пространстве. Помимо этой работы авторы параллельно решают задачи проектирования цифрового двойника солнечной электростанции.

За счет использования одного и того же подхода: онтологического инжиниринга, определения тех информационных потоков, которые необходимы для работы каждого ЦД в отдельности, и совместного функционирования в единой системе, предполагается решить вопрос их общей интеграции в едином информационном пространстве.

Анализ исследований, выполняемых в области использования концепции цифровых двойников [18, 19], подтверждает актуальность их построения для объектов ветровой электроэнергетики.

2. Онтологический инжиниринг при построении ЦД ВЭС. Применение онтологий в структуре ЦД приводится в работе [19]. Онтологический инжиниринг для построения ЦД был так же применен при построении ЦД солнечной электростанции (СЭС) [20]. В рамках данной работы были использованы классические этапы онтологического инжиниринга относительно разработки ЦД ВЭС.

Ниже представлены их результаты:

1. На этом этапе была поставлена цель – построить систему онтологий ветровой электростанции и интегрировать её в уже имеющуюся онтологию ТЭК, а также формализовать знания в области ветровой электроэнергетики для построения цифрового двойника. Определить область действия или «границы» онтологии и тип онтологии, прежде чем перейти к шагу 2. Разрабатываемая онтология относится к типу «легких онтологий» будет использована для построения цифровой тени и цифрового двойника.
2. На этом этапе была собрана вся необходимая информация, которая относится к описываемой области. Основная цель этого этапа – отбор и вербализация всех основных объектов и понятий в области. В результате были отобраны основные объекты и понятия, которые относятся к следующим областям:
 - система генерации;
 - возобновляемые источники энергии;
 - ветровые электростанции;
 - оборудование ВЭС.
3. Следующим этапом является определение всех основных объектов и концептов предметной области (основных уровней абстракции). Выявление структурированной иерархии является одной из основных целей на данном этапе. Иерархия, построенная с использованием результатов предыдущего шага, представлена ниже:
 - электроэнергетические системы;
 - системы генерации;
 - возобновляемые источники энергии;
 - ветровые электростанции;
 - оборудование ВЭС.

4. Последний этап посвящен верификации визуальной структуры путем исключения любых излишеств, синонимии и противоречий. На этом шаге были переработаны концепты и связи между ними, что привело к улучшению понятности и читаемости системы онтологий. В результате выполнения четырехступенчатого алгоритма была построена система онтологий ветровой электростанции. На рис. 1 приведена метаонтология, концепты которой, отмеченные значком, могут быть детализированы (пример для оборудования ВЭС приведен на рис. 2)

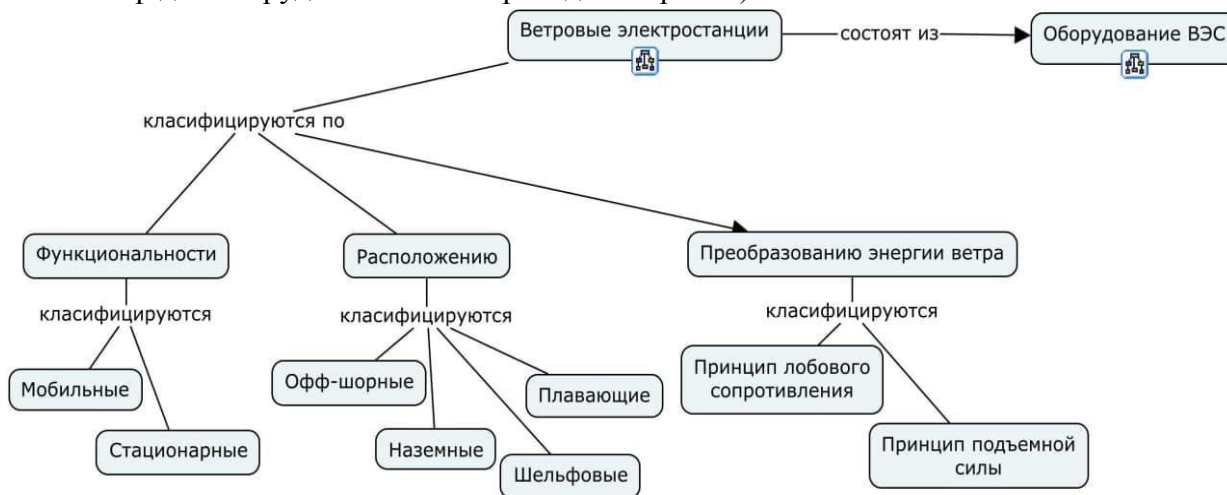


Рис. 1. Метаонтология ветровых электростанций

3. Математическая модель. В рамках данной работы для ЦД был выбран подход, предлагаемый Оганесяном Э.В., Бекировым Э.А. и др. [21]. В соответствии с ним приведем описание математической модели для определения параметров работы ветроэнергетической установки, которая необходима для построения ЦД ВЭС.

Для вычисления генерируемой мощности ветроэнергетической установки требуется использовать такие параметры: скорость ветра, коэффициент использования энергии ветра, быстроходность, КПД и площадь ометаемой поверхности.

В формуле (1) проводится расчет мощностной характеристики ВЭУ в зависимости от скорости ветра:

$$P_{ВЭУ}(V_i) = \rho C_P S_0 \eta \frac{V_i^3}{2} 10^{-3}, \quad (1)$$

Далее необходимо рассчитать характеристику коэффициента использования энергии ветра C_P от быстроходности Z ($Z_{opt} < Z < Z_{max}$) (2).

$$C_P = C_{Pmax} - \frac{C_{Pmax}}{(Z_{max} - Z_{opt})^2} (Z - Z_{opt})^2, \quad (2)$$

где C_{Pmax} – максимальный коэффициент использования энергии ветра.

При $Z \leq Z_{opt}$:

$$C_P = C_{Pmax} - \left(\frac{Z}{Z_{opt}}\right)^2 \left(3 - 2\frac{Z}{Z_{opt}}\right), \quad (3)$$

Для нахождения быстроходности используется следующее выражение.

$$Z = \frac{\omega R}{V_0}, \quad (4)$$

Где ω – угловая скорость; R – радиус ветроколеса; V_0 – скорость набегающего потока воздуха.

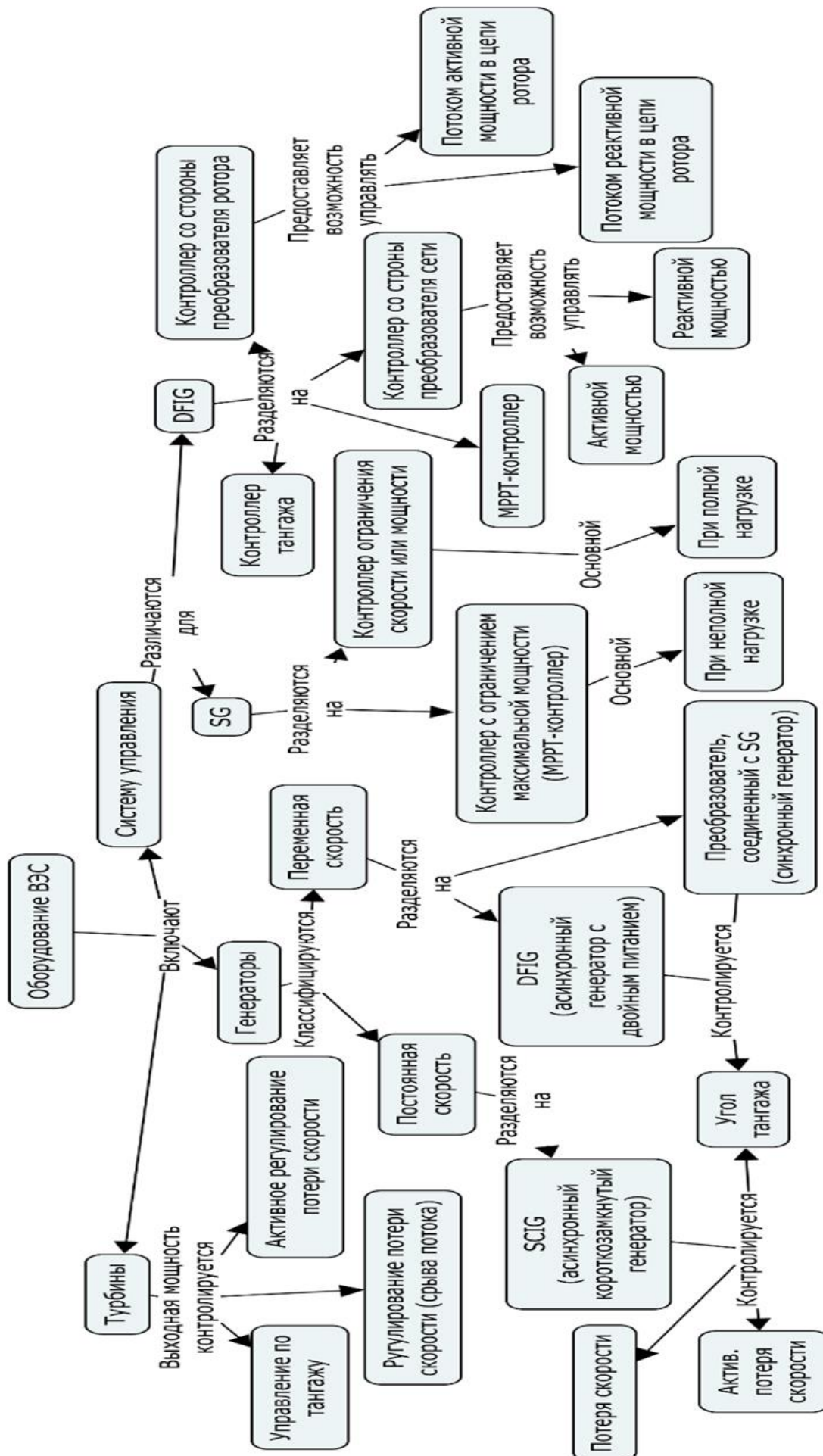


Рис. 2. Онтология оборудования ВЭС

Эффективная выходная удельная мощность ΔP на 1 м^2 ометаемой поверхности ветротурбины (ВТ) вычисляется по формуле:

$$\Delta P = \Delta P_{\text{ВП}} C_p \eta, \quad (5)$$

где $\Delta P_{\text{ВП}} = \rho 10^{-3} \frac{V_{\text{НОМ}}^3}{2}$ – удельная мощность ветрового потока при скорости ветра $V_{\text{НОМ}}$; ρ – плотность воздуха (зима – 1,25, лето 0,72 кг/м³); C_p – средний коэффициент использования энергии ветра (принят равным 0,31); $\eta = \eta_A \eta_{\text{РЕД}} \eta_{\text{ЭГ}}$ – суммарный КПД ВЭУ; η_A – аэродинамический КПД ВТ (принят в пределах 0,91..0,916); $\eta_{\text{РЕД}}$ – КПД редуктора (принят в пределах 0,95..0,96); $\eta_{\text{ЭГ}}$ – КПД электрогенератора, зависящий от его мощности.

После расчетов величины ΔP определяется площадь ометаемой поверхности ВТ:

$$S_0 = \frac{P_{\text{ВЭУНОМ}}}{\Delta P} \quad (6)$$

Представленная математическая модель определяет необходимые параметры и имеет преимущества, которые связаны с учетом изменения погодных условий и взаимодействия элементов ветроэнергетической установки.

Рассмотренная модель необходима для получения данных о поведении системы. Использование этих данных дает возможность цифровому двойнику прогнозировать поведение объекта и решать задачу управления его жизненным циклом

4. Проектирование ЦД ВЭС. При проектировании и функционировании цифрового двойника ветровой электростанции необходимо использовать большой массив данных, например, характеристики погодных условий, характеристики оборудования и т.д. Этот набор данных необходимо организовать и хранить. В нашем случае было решено использовать реляционную базу данных PostgreSQL.

Проектирование базы данных велось на основе онтологий (рис. 1 и 2), построенных в пункте “онтологический инжиниринг”. В результате проектирования были выделены следующие сущности:

- характеристики погоды;
- характеристики ветрогенератора;
- характеристики сетевого инвертора;
- изолированная система;
- сетевой инвертор;
- расчет сетевого инвертора;
- ветрогенератор;
- расчет ветрогенератора;
- изолированная система_Погода.

Логическая модель данных приведена на рис.3.

Для понимания роли и места базы данных в общей архитектуре ЦД ВЭС на основе онтологии построена архитектура типового ЦД, представленная на рис. 4.

Архитектура включает в себя следующие блоки.

- Цифровая тень:

Состоит из программы цифровой тени, которая отвечает за работу с базой данных и из методов машинного обучения, которые необходимы для воссоздания модели поведения системы, а также для воссоздания пропущенных данных, необходимых для работоспособности всего цифрового двойника;

- Программный комплекс:

Необходим для обеспечения работоспособности всей системы, так как реализует интерфейс приложения, математическую модель и систему сбора оперативной информации, использующуюся для внесения управляющего воздействия посредством системы управления;

5. Интерфейс для работы с базой данных. Работу с базой данных можно выполнять с использованием инструментов администрирования, предоставляемых СУБД PostgreSQL. Интерфейс для редактирования представлен на рис.5.

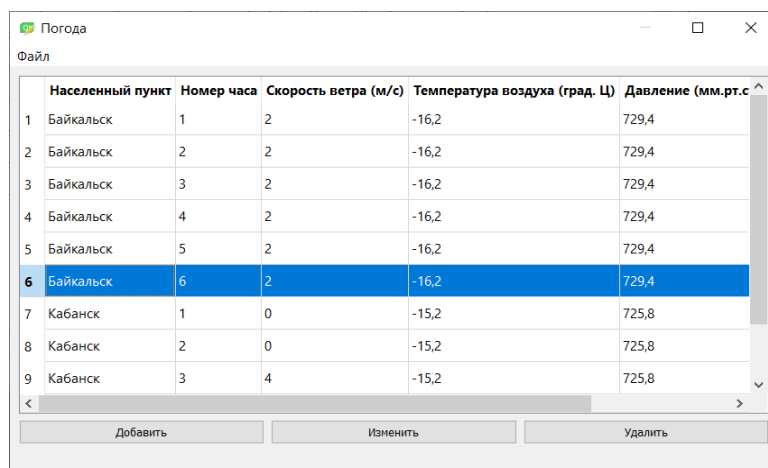
	locality character varying (15)	hour_number integer	wind_speed smallint	wind_temperature real	pressure real	direct_solar_rad real	diffuse_solar_rad real
84	Baikalsk	105037	1	-11.5	724.4	36.269047	250.0527
85	Baikalsk	105036	1	-11.5	724.4	5.906586	178.32928
86	Baikalsk	105035	2	-11.5	724.4	0	102.317955
87	Baikalsk	105034	3	-11.5	724.4	0	0

Рис. 5. Интерфейс PostgreSQL

Для добавления записей в таблицу и их удаления необходимо использовать SQL-запросы, что, в свою очередь, усложняет работу с ЦД. Вследствие этого было решено разработать свой интерфейс работы с базой данных, обеспечивающий следующие возможности:

- добавление форм для редактирования (добавления/изменения/удаления) таблиц;
- разработка формы, предоставляющей возможность подключения ветрогенератора к инверторам по особым условиям (количество портов в инверторе, подключение в инвертор только одинаковых ветрогенераторов и т.д.).

Пример разработанного интерфейса представлен на рис. 6.



	Населенный пункт	Номер часа	Скорость ветра (м/с)	Температура воздуха (град. Ц)	Давление (мм.рт.с)
1	Байкальск	1	2	-16,2	729,4
2	Байкальск	2	2	-16,2	729,4
3	Байкальск	3	2	-16,2	729,4
4	Байкальск	4	2	-16,2	729,4
5	Байкальск	5	2	-16,2	729,4
6	Байкальск	6	2	-16,2	729,4
7	Кабанск	1	0	-15,2	725,8
8	Кабанск	2	0	-15,2	725,8
9	Кабанск	3	4	-15,2	725,8

Рис. 6. Интерфейс программы для редактирования записей в таблице «Характеристики погоды»

Заключение. В статье рассмотрено существующее состояние в области разработки цифровых двойников, обозначены основные проблемы этих исследований, связанные, в первую очередь, с трудностями их интеграции с ЦД других энергетических систем. Детально описан онтологический подход, который использовался при проектировании цифрового двойника и базы данных для ветровой электростанции. Рассмотрен онтологический инжиниринг, как необходимый этап построения цифрового двойника ВЭС. Показана структура базы данных, её основные компоненты и их взаимосвязи на основе онтологических моделей. Основным назначением онтологий при разработке цифровых двойников является формальное описание и интеграция всех компонентов: математических моделей, баз данных и баз знаний (онтологий). В рамках развития этой работы предлагается усовершенствовать информационную модель, включающую базу данных, методы машинного обучения и систему сбора оперативной информации, за счет преобразования ее в цифровую тень. Преобразование приведет к улучшению сопровождения цифрового двойника, так как модуль с основными вычислениями (математической моделью) и модуль, предназначенный для работы с данными (цифровая тень) будут разделены.

В дальнейшем также необходимо подготовить систему к работе с большими массивами данных, собираемых с реального объекта. При работе с реальными объектами электроэнергетики появляется такая проблема, как большие объемы данных, поступающих с датчиков, которые необходимо хранить и обрабатывать. Для решения этой проблемы предлагается использовать методы Big Data, что, в свою очередь, позволит обрабатывать данные в доступное время.

В результате выполнения вышеперечисленного появится возможность верификации модели по данным, получаемым от реального объекта, и, в случае необходимости, ее дальнейшей корректировки.

Благодарности. Результаты получены в рамках выполнения базового проекта ИСЭМ СО РАН АААА-А21-121012090007-7 по госзаданию FWEU-2021-0007 и проекта по гранту РФФИ №20-07-00994

СПИСОК ИСТОЧНИКОВ

1. Энергетическая стратегия РФ на период до 2035 года. URL: <http://static.government.ru/media/files/w4sigFOiDjGVDYT4IgsApssm6mZRb7wx.pdf> (дата обращения 25.03.2022).
2. Суслов К.В. Развитие систем электроснабжения изолированных территорий России с использованием возобновляемых источников энергии / К.В. Суслов // Вестник Иркутского государственного технического университета, 2017. – Т. 21. – № 5. – С. 131-142. – DOI: 10.21285/1814-3520-2017-5-131-142
3. Никитина Е. Попали в сети: как работают цифровые двойники в электроэнергетике. / Е. Никитина. – URL: <https://pro.rbc.ru/news/5db1b59a9a79474bb142a3fe> (дата обращения 25.03.2022).
4. Цифровой двойник, Индустрия 4.0. Информатизация и системы управления в промышленности. – URL: <https://zen.yandex.ru/media/isup/cifrovoy-dvoynik-industriia-40-5b83b7155b279900a96c54e8>.
5. Garfinkel J. Gartner Identifies the Top 10 Strategic Technology Trends for 2019 // Gartner Tech. Rep, 2018, Oct.
6. Ebrahimi A. Challenges of developing a digital twin model of renewable energy generators // 2019 IEEE 28th International Symposium on Industrial Electronics (ISIE). 2019, pp. 1059-1066, DOI: 10.1109/ISIE.2019.8781529
7. Pimenta, F., Pacheco, J., Branco, C. M., Teixeira, C. M., & Magalhães, F. Development of a digital twin of an onshore wind turbine using monitoring data // Journal of Physics: Conference Series, 2020, 1618, 022065, DOI:10.1088/1742-6596/1618/2/022065
8. Jahanshahi Zeitouni, M., Parvaresh, A., Abrazeh, S., Mohseni, S.-R., Gheisarnejad, M., & Khooban, M.-H. Digital Twins-Assisted Design of Next-Generation Advanced Controllers for Power Systems and Electronics: Wind Turbine as a Case Study // Inventions, 2020, 5(2), 19, DOI:10.3390/inventions5020019
9. Chetan, M., Yao, S., & Griffith, D. T. Multi-fidelity digital twin structural model for a sub-scale downwind wind turbine rotor blade. Wind Energy. 2021, DOI:10.1002/we.2636
10. Karl Merz, Valentin Chabaud, Paula B. Garcia-Rosa and Konstanze Kolle A hierarchical supervisory wind power plant controller. Journal of Physics: Conference Serie. 2018, DOI: 10.1088/1742-6596/2018/1/012026
11. H. Solman, J. K. Kirkegaard, M. Smits, B. V. Vliet, S. Bush Digital twinning as an act of governance in the wind energy sector. Environmental Science & Policy. 2022, v. 127, pp. 272-279, DOI: 10.1016/j.envsci.2021.10.027

12. F. Dembski, U. Wössner, M. Letzgus, M. Ruddat and C. Yamu Urban Digital Twins for Smart Cities and Citizens: The Case Study of Herrenberg, Germany. Sustainability. 2020, 12(6), 2307, DOI: 10.3390/su12062307
13. E. Glaessgen and D. Stargel, "The Digital Twin Paradigm for Future NASA and U.S. Air Force Vehicles" // 53rd Structures, Structural Dynamics and Materials Conference. 2012.
14. Digital solutions for wind farms. URL: <https://www.ge.com/renewableenergy/wind-energy/onshore-wind/digital-wind-farm> (accessed: 20.03.2022)
15. H. Pargmann, D. Euhausen, R. Faber, Intelligent Big Data Processing for Wind Farm Monitoring and Analysis Based on Cloud-Technologies and Digital Twins. 2018 the 3rd IEEE International Conference on Cloud Computing and Big Data Analysis. 2018, pp. 233-237.
16. Obafemi O. Olatunji, Paul A. Adedeji, N. Madushele, Tien-Chien Jen, Overview of Digital Twin Technology in Wind Turbine Fault Diagnosis and Condition Monitoring. 2021 IEEE 12th International Conference on Mechanical and Intelligent Manufacturing Technologies. 2021, pp. 201-207, DOI: 10.1109/ICMIMT52186.2021.9476186
17. D.J. Wagg, K. Worden, R.J. Barthorpe, P. Gardner, Digital twins: State-of-the-art and future directions for modeling and simulation in engineering dynamics applications. ASCE-ASME J. Risk Uncertain. Eng. Syst. 2020, В 6 (3), pp. 1–17, DOI: 10.1115/1.4046739
18. Андриюшкевич, С.К. Подходы к разработке и применению цифровых двойников энергетических систем / С.К. Андриюшкевич, С.П. Ковалев, Е. Нефедов // Цифровая подстанция, 2019. – № 12. – С.38-43.
19. Ковалёв, С.П. Проектирование информационного обеспечения цифровых двойников энергетических систем / С.П. Ковалёв // Системы и средства информатики, 2020. – Т.30. – №1. – С. 66-81.
20. Ludmila Massel, Nikita Shchukin and Alexey Cybikov International Conference of Young Scientists "Energy Systems Research 2021". Vol. 289, Digital twin development of a solar power plant, DOI: 10.1051/e3sconf/202128903002
21. Оганесян Э.В., Бекиров Э.А., Асанов М.М. Математическая модель для определения параметров работы ветроэнергетической установки / Э.В. Оганесян, Э.А. Бекиров, М.М. Асанов // Строительство и техногенная безопасность, 2016. – №3 (55). – С. 82-86.

Массель Людмила Васильевна, профессор, доктор технических наук, зав. отделом систем искусственного интеллекта в энергетике, Институт систем энергетики им. Л. А. Мелентьева СО РАН, massel@isem.irk.ru, ORCID: 0000-0002-0351-0415, Россия, Иркутск, Лермонтова д.130

Массель Алексей Геннадьевич, кандидат технических наук, старший научный сотрудник отдела систем искусственного интеллекта в энергетике, Институт систем энергетики им. Л. А. Мелентьева СО РАН, amassel@gmail.com, AuthorID: 173012, SPIN-код: 2940-2510, Россия, Иркутск, Лермонтова д.130

Щукин Никита Игоревич, инженер-исследователь отдела систем искусственного интеллекта в энергетике, Институт систем энергетики им. Л. А. Мелентьева СО РАН, niksha14@mail.ru, SPIN-код: 6038-0410, ORCID: 0000-0001-8331-335X, Россия, Иркутск, Лермонтова д.130

Цыбиков Алексей Ринчинович, инженер-исследователь отдела систем искусственного интеллекта в энергетике, Институт систем энергетики им. Л. А. Мелентьева СО РАН, tsibikow@mail.ru, Россия, Иркутск, Лермонтова д.130

UDC 004.89

doi:10.38028/ESI.2022.25.1.007

Digital twin development of a wind farm: problem statement and construction**Liudmila V. Massel, Alexey G. Massel, Nikita I. Shchukin, Alexey R. Tsybikov**

Melentiev Energy Systems Institute SB RAS,

Russia, Irkutsk, *niksha14@mail.ru*

Abstract. The article is devoted to the development of a digital twin (DT) of a wind farm. A review of existing solutions in this area is made. The approach to the construction of DT, based on ontological engineering, is considered in detail. The main steps of ontological engineering are described. A fragment of the system of ontologies of wind systems is given. A mathematical model for determining the operation parameters of a wind power plant, used in the construction of DT, is considered. The article also focuses on the design of a DT architecture, including a digital shadow, a digital model, and a control system. Also in the article is the stage of implementation of the prototype DT wind farm.

Keywords: digital twin, ontologies, databases, wind energy

Acknowledgements: The results were obtained as part of the implementation of the basic project ISEM SB RAS AAAA-A21-121012090007-7 under the state order FWEU-2021-0007 and RFBR grant no. 20-07-00994

REFERENCES

1. Energeticheskaya strategiya RF na period do 2035 goda [Energy strategy of the Russian Federation for the period up to 2035]. URL: <http://static.government.ru/media/files/w4sigFOiDjGVDYT4IgsApssm6mZRb7wx.pdf> (data obrashcheniya 25.03.2022).
2. Suslov K.V. Razvitie sistem elektrosnabzheniya izolirovannyh territorij Rossii s ispol'zovaniem vozobnovlyaemyh istochnikov energii [Development of power supply systems of russian isolated territories using renewable energy sources]. Vestnik Irkutskogo gosudarstvennogo tekhnicheskogo universiteta [Proceedings of Irkutsk State Technical University], 2017, T. 21, no. 5, pp. 131-142, DOI: 10.21285/1814-3520-2017-5-131-142
3. Nikitina E. Popali v seti: kak rabotayut cifrovye dvojniki v elektroenergetike [Caught in the net: how digital twins work in the power industry], URL: <https://pro.rbc.ru/news/5db1b59a9a79474bb142a3fe> (data obrashcheniya 25.03.2022).
4. Cifrovoj dvojniki, Industriya 4.0. Informatizatsiya i sistemy upravleniya v pro-myshlennosti [Digital twin, Industry 4.0. Informatization and control systems in industry, URL: <https://zen.yandex.ru/media/isup/cifrovoi-dvojniki-industriia-40-83b7155b279900a96c54e8>.
5. Garfinkel J. Gartner Identifies the Top 10 Strategic Technology Trends for 2019. Gartner Tech. Rep, 2018, Oct.
6. Ebrahimi A. Challenges of developing a digital twin model of renewable energy generators // 2019 IEEE 28th International Symposium on Industrial Electronics (ISIE). 2019, pp. 1059-1066, DOI: 10.1109/ISIE.2019.8781529
7. Pimenta, F., Pacheco, J., Branco, C. M., Teixeira, C. M., & Magalhães, F. Development of a digital twin of an onshore wind turbine using monitoring data // Journal of Physics: Conference Series, 2020, 1618, 022065, DOI:10.1088/1742-6596/1618/2/022065
8. Jahanshahi Zeitouni, M., Parvaresh, A., Abrazeh, S., Mohseni, S.-R., Gheisarnejad, M., & Khooban, M.-H. Digital Twins-Assisted Design of Next-Generation Advanced Controllers for Power Systems and Electronics: Wind Turbine as a Case Study // Inventions, 2020, 5(2), 19, DOI:10.3390/inventions5020019
9. Chetan, M., Yao, S., & Griffith, D. T. Multi-fidelity digital twin structural model for a sub-scale downwind wind turbine rotor blade. Wind Energy. 2021, DOI:10.1002/we.2636
10. Karl Merz, Valentin Chabaud, Paula B. Garcia-Rosa and Konstanze Kolle A hierarchical supervisory wind power plant controller. Journal of Physics: Conference Serie. 2018,

- DOI: 10.1088/1742-6596/2018/1/012026
11. H. Solman, J. K. Kirkegaard, M. Smits, B. V. Vliet, S. Bush Digital twinning as an act of governance in the wind energy sector. *Environmental Science & Policy*. 2022, v. 127, pp. 272-279, DOI: 10.1016/j.envsci.2021.10.027
 12. F. Dembski, U. Wössner, M. Letzgun, M. Ruddat and C. Yamu Urban Digital Twins for Smart Cities and Citizens: The Case Study of Herrenberg, Germany. *Sustainability*. 2020, 12(6), 2307, DOI: 10.3390/su12062307
 13. E. Glaessgen and D. Stargel, "The Digital Twin Paradigm for Future NASA and U.S. Air Force Vehicles" // 53rd Structures, Structural Dynamics and Materials Conference. 2012.
 14. Digital solutions for wind farms. URL: <https://www.ge.com/renewableenergy/wind-energy/onshore-wind/digital-wind-farm> (accessed: 20.03.2022)
 15. H. Pargmann, D. Euhausen, R. Faber, Intelligent Big Data Processing for Wind Farm Monitoring and Analysis Based on Cloud-Technologies and Digital Twins. 2018 the 3rd IEEE International Conference on Cloud Computing and Big Data Analysis. 2018, pp. 233-237.
 16. Obafemi O. Olatunji, Paul A. Adediji, N. Madushele, Tien-Chien Jen, Overview of Digital Twin Technology in Wind Turbine Fault Diagnosis and Condition Monitoring. 2021 IEEE 12th International Conference on Mechanical and Intelligent Manufacturing Technologies. 2021, pp. 201-207, DOI: 10.1109/ICMIMT52186.2021.9476186
 17. D.J. Wagg, K. Worden, R.J. Barthorpe, P. Gardner, Digital twins: State-of-the-art and future directions for modeling and simulation in engineering dynamics applications. *ASCE-ASME J. Risk Uncertain. Eng. Syst.* 2020, B 6 (3), pp. 1–17, DOI: 10.1115/1.4046739
 18. Andryushkevich, S.K. Podhody k razrabotke i primeneniyu cifrovyyh dvoynikov energeticheskikh sistem [Approaches to the development and application of digital twins of energy systems]. *Tsifrovaya podstantsiya [Digital substation]*. 2019, no. 12, pp.38-43.
 19. Kovalyov, S.P. Proektirovanie informacionnogo obespecheniya cifrovyyh dvoynikov energeticheskikh system [Information architecture of the power system digital twin]. *Sistemy i sredstva informatiki [Systems and means of informatics]*. 2020, vol.30, no. 1, pp. 66-81.
 20. Ludmila Massel, Nikita Shchukin and Alexey Cybikov International Conference of Young Scientists "Energy Systems Research 2021". Vol. 289, Digital twin development of a solar power plant, DOI: 10.1051/e3sconf/202128903002
 21. Oganessian E.V., Bekirov E.A., Asanov M.M. Matematicheskaya model' dlya opredeleniya parametrov raboty vetroenergeticheskij ustanovki [Mathematical model for determination of wind turbine operation parameters]. *Stroitel'stvo i tekhnogennaya bezopasnost' [Construction and technogenic safety]*. 2016, no. 3 (55), pp. 82-86.

Liudmila V. Massel, professor, Head of the Department of Artificial Intelligence Systems in Energy, Institute of Energy Systems. L. A. Melentiev SB RAS, massel@isem.irk.ru, ORCID: 0000-0002-0351-0415, Russia, Irkutsk, st. Lermontova d.130

Alexey G. Massel, Ph. D, Senior Research Fellow of the Department of Artificial Intelligence Systems in the Energy Industry of the Institute of Energy Systems named after. L. A. Melentiev SB RAS, amassel@gmail.com, AuthorID:173012, SPIN: 2940-2510, Russia, Irkutsk, Lermontova d.130

Nikita I. Shchukin, Research Engineer of the Department of Artificial Intelligence Systems in the Energy Industry of the Institute of Energy Systems. L. A. Melentiev SB RAS, niksha14@mail.ru, SPIN: 6038-0410, ORCID: 0000-0001-8331-335X, Russia, Irkutsk, st. Lermontova d.130

Alexey R. Tsybikov, Research Engineer of the Department of Artificial Intelligence Systems in the Energy Industry of the Institute of Energy Systems named after. L. A. Melentiev SB RAS, tsibikow@mail.ru, Russia, Irkutsk, st. Lermontova d.130

Статья поступила в редакцию 20.03.2022; одобрена после рецензирования 25.03.2022; принята к публикации 29.03.2022.

The article was submitted 03.20.2022; approved after reviewing 03.25.2022; accepted for publication 03.29.2022

УДК 004.89

doi:10.38028/ESI.2022.25.1.008

Прогнозирование солнечной радиации и импутация данных для цифрового двойника солнечной электростанции

Лосев Алексей Сергеевич, Массель Алексей Геннадьевич

Институт систем энергетики им. Л.А. Мелентьева СО РАН,

Россия, Иркутск, losev.as@isem.irk.ru

Аннотация. В статье представлена реализация авторегрессионной модели для прогнозирования данных о солнечной радиации. Был проведен эксперимент с некоторыми подходами к импутации данных¹ и обоснован выбор подхода MICE (Multiple Imputation by Chained Equation – множественная импутация с помощью связанного уравнения). Также был проведен сравнительный эксперимент прогнозирования солнечной радиации на основе первоначальных данных и на основе данных, в которых была применена импутация.

Ключевые слова: цифровой двойник, импутация, солнечная радиация, авторегрессия, прогнозирование, ARIMA, SARIMA, MICE

Цитирование: Лосев А.С., Массель А.Г. Прогнозирование солнечной радиации и импутация данных для цифрового двойника солнечной электростанции // Информационные и математические технологии в науке и управлении. – 2022. – № 1 (25). – С. 91-101. – DOI:10.38028/ESI.2022.25.1.008.

Введение. Выделяют два этапа разработки цифровых двойников: 1) разработка прототипа с использованием научного инструментария исследователей, 2) отладка информационных взаимодействий с реальными информационными потоками для решения конкретных задач [1]. Работа с реальными данными допустима только на втором этапе, потому появляется необходимость в создании генератора правдоподобных синтетических данных, которые впоследствии могли бы использоваться во время тестирования цифрового двойника уже на первом этапе. В Институте систем энергетики им. Л.А. Мелентьева СО РАН было принято решение разработать модель генерации и импутации данных для цифрового двойника солнечной электростанции на основе имеющихся погодных данных, полученных от реального объекта.

В настоящее время продолжает возрастать роль солнечной энергии в производстве электроэнергии, поскольку возобновляемые источники энергии становятся многообещающей альтернативой для мира в целом [2]. Однако, если речь идет об установке солнечных электростанций в специфических регионах, например, на Крайнем севере, необходимо обосновать целесообразность такого решения. В качестве обоснования в статье предложен подход к прогнозированию солнечного излучения для лучшего представления потенциального генерируемого количества электроэнергии отдельно взятой системы солнечных панелей.

Существуют несколько моделей прогнозирования для предсказания будущего излучения, в том числе, включающие в себя технологии искусственного интеллекта, например, модель численного прогноза погоды (NWP), которая описывается нелинейными уравнениями в частных производных погодных переменных, с применением временных рядов в качестве фиксации динамики излучения [2, 3]. Эта технология, возможно, будет реализована в будущем, когда возникнет необходимость генерировать значения комплексно, с последующим вычислением мощности в зависимости не только от солнечного излучения, но и от влажности,

¹Процесс импутации - это замещение ошибочных, противоречивых и отсутствующих ответов в процессе редактирования данных другими ответами - значениями показателей.

скорости ветра и т.д. Приоритет первого прототипа будет больше сосредоточен на импутации пропущенных данных.

Оценка параметров в модели временных рядов определенно требует полного набора исторических данных. Данные зачастую содержат пропущенные значения. Более того, отсутствие данных может содержать пропуски за несколько дней подряд. Эту проблему возможно решить, заполнив значения записей с других датчиков [3, 2], однако нет никакой гарантии того, что запасные датчики также не выйдут из строя. Более продвинутые методы множественной импутации, например, анализ главных компонент, реализованы в MATLAB в качестве набора инструментов [4]. Однако, такие методы обычно предполагают, что данные являются статическими случайными величинами и не используют временную динамику данных применительно к временным рядам [3].

В этой работе сравниваются практические методы, подходящие для импутации большого объема пропущенных данных, реализованные на языке Python. Также предложен метод прогнозирования солнечной радиации как временного ряда с помощью модели авторегрессии ARIMA.

1. Прогнозирование солнечной радиации. Авторегрессионная интегрированная модель скользящего среднего (ARIMA) можно использовать для объяснения динамики нестационарных временных рядов. Эта модель популярна у исследователей за счет перспективы ее интеграции с нейронными сетями [5, 6].

Пусть $y(t)$ будет временным рядом солнечной радиации. Определим L как оператор запаздывания: $Ly(t) = y(t - 1)$, где шкала времени – каждый час. Модель ARIMA (p, d, q) описывается [7, 8]:

$$A(L)(1 - L)^d y(t) = C(L)v(t) \quad (1)$$

где полиномы авторегрессии и скользящего среднего

$$A(L) = 1 - (a_1 L + \dots + a_p L^p), C(L) = 1 + c_1 L + \dots + c_q L^q, \text{ соответственно.}$$

Член $(1 - L)^d$ является интегральной частью, где d – порядок разности, который может быть определен заранее из автокорреляционного анализа. Выполняем дифференцирование $\nabla y(t) = y(t) - y(t - 1)$ до тех пор, пока автокорреляционная функция (АКФ) результирующего временного ряда оказывается стационарной. Переменная $v(t) \in \mathbb{R}$ – белый шум с нулевым средним и дисперсией σ^2 . Интуитивно понятно, что солнечное излучение имеет суточный цикл. Предположим, что мы пренебрегаем данными ночных часов, поскольку значения освещенности равны нулю. Это приводит к временному ряду с периодом 12 часов. В качестве примера, который иллюстрирует метод определения сезонности данных с использованием декомпозиция команды в R, необходимо построить тренд временного ряда

$$\hat{m}(t) = (1/16)[(1/2)y(t - 8) + y(t - 7) + \dots + y(t + 7) + (1/2)y(t + 8)]$$

и вычесть его значение из значения солнечной радиации: $\tilde{s}(t) = y(t) - \hat{m}(t)$ (измерения собираются ежечасно). Оценку суточного сезонного эффекта $\hat{s}(t)$ можно получить по формуле усреднения $\tilde{s}(t)$:

$$\hat{s}(t) = \frac{1}{N + 1} \sum_{k=0}^N \tilde{s}(t + kT)$$

где T – сезонный период. Рассмотрим модель, содержащую аддитивную сезонную индикаторную переменную согласно [7]:

$$A(L)(1 - L)^d y(t) = s(t) + \alpha + C(L)v(t),$$

где $s(t) = s(t - kT)$ – сезонный член, повторяющийся каждый цикл, k является целым числом, а α является константой. Если $y(t)$ имеет период T , то член $s(t)$ можно убрать операцией $y(t) - y(t - T) = (1 - L^T) y(t)$. Так как ARIMA используется для прогнозирования солнечной радиации с учетом сезонности, применяется сезонная модель ARIMA

$$\tilde{A}(L^T)A(L)(1 - L^T)D(1 - L)^d y(t) = \tilde{C}(L^T)C(L)v(t), \quad (2)$$

где $\tilde{A}(L)$ и $\tilde{C}(L)$ являются сезонными авторегрессионными полиномами скользящего среднего:

$$\begin{aligned} \tilde{A}(L) &= I - (\tilde{a}_1 L^T + \dots + \tilde{a}_p L^T), \\ \tilde{C}(L) &= I + \tilde{c}_1 L^T + \dots + \tilde{c}_q L^T \end{aligned}$$

D – интегрированный сезонный порядок. Описание сезонной модели ARIMA обычно обозначается SARIMA(p, d, q)(P, D, Q)_T.

Оценка моделей SARIMA требует выбора правильного значения модельных порядков p, d, q, P, D, Q. В ходе эксперимента было установлено, что $T = 12$, из предварительных данных о солнечной радиации. Другие параметры выбираются таким образом, чтобы преобразованный временной ряд имел стационарную характеристику, которая может наблюдаться на графиках АКФ, и оценка модели была сведена к минимуму. Когда параметры для нестационарных частей (d, D, P, Q) определены, остается оценить часть модели ARMA. В качестве наглядного примера можно привести график АКФ необработанных данных о солнечной радиации. На рис. 1 (сверху) видно, что данные не являются стационарным процессом, т.к. АКФ периодична каждые 12 лагов (сдвигов во времени), а также медленно затухает после 1 лага. Чтобы найти подходящий параметр d, разность на y применяется так, чтобы АКФ разностного ряда вела себя аналогично АКФ стационарного процесса по следующим характеристикам [8].

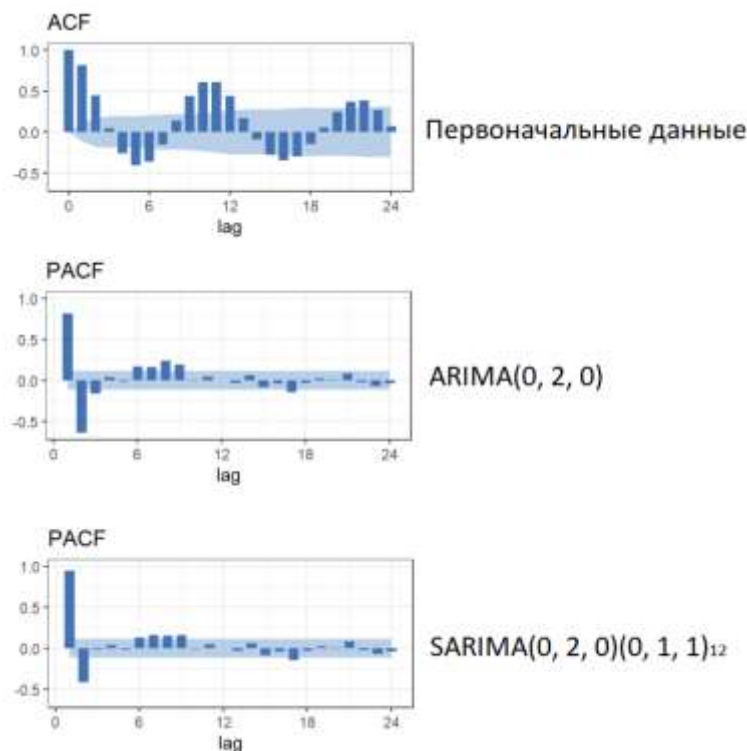


Рис. 1. Автокорреляционная функция (АКФ) первоначальных данных (вверху), АКФ остатка ARIMA (0,2,0) (в центре) и АКФ остатка SARIMA (0, 2, 0)(0, 1, 1)₁₂ (внизу)

Обнаружено, что $d = 2$ является подходящим выбором для этого. Набор данных и результаты АКФ, представленные на рис. 1 (в центре), показывают резкое снижение АКФ. Однако, некоторые лаги АКФ малы и лежат за пределами доверительного интервала. Далее задаем параметры $(P, D, Q)_T = (0, 1, 1)_{12}$ сезонных членов и это дает лишь несколько лагов АКФ, лежащих за пределами доверительного интервала, как показано на рис. 1 (внизу). Следовательно, соответствующий выбор p и q может быть найден путем минимизации информационного критерия Акаике (AIC) или байесовского информационного критерия (BIC) в качестве критериев выбора модели [8].

2. Импутация данных. После того, как определен метод прогнозирования временных рядов, встает вопрос о его работе с поврежденными данными. Модель SARIMA требует числовые входные значения, которые должны присутствовать для каждой строки и столбца в наборе данных. Таким образом, пропущенные значения могут вызвать проблемы во время получения прогноза.

Таким образом, обычно выявляют отсутствующие значения в наборе данных и заменяют их подходящим числовым значением. Это называется импутацией данных. В российских источниках термин «импутация» встречается в социологических исследованиях [9]. Вместо импутации можно также встретить термин «восстановление пропущенных значений», например [10].

Простой и популярный подход импутации данных включает использование статистических методов для оценки значения столбца на основе имеющихся значений, а затем замену всех отсутствующих значений в столбце рассчитанными статистическими. Он прост, потому что статистические данные подсчитываются быстро, и популярен, потому что часто оказывается очень эффективным [11].

Рассчитываемые общие статистические данные включают: среднее арифметическое значение столбца; медиану столбца; моду столбца; константу.

Необходимо было провести эксперимент для проверки каждой статистической стратегии и выяснить, что лучше всего работает для этого набора данных, сравнивая среднее значение, медиану, моду (наиболее частую) и постоянную (0) стратегии. Затем можно сравнить среднюю точность каждого подхода. Перед экспериментом было удалено 20% первоначальных погодных данных с различными интервалами, от одного часа (одна запись в базе данных) до суток (двадцать четыре записей подряд в базе данных). В качестве константы было выбрано значение «0» (отсутствие солнечной радиации, ночное время суток).

В результатах на рис. 2 видны смещения во всех четырех классических стратегиях. Самой лучшей стратегией модель выбрала вариант с константой, однако, в самих данных наблюдается неправдоподобность. Несмотря на хорошие результаты, провал эксперимента обусловлен потерей контекста данных – пропадает характер сезонности, а именно, нулевая радиация на протяжении нескольких суток.

Потому было принято решение использовать любой из двух специализированных методов: максимальное правдоподобие или множественную импутацию.

В статистике метод максимального правдоподобия – это статистическая оценка интересующего распределения, обладающего тем свойством, что он максимизирует функцию правдоподобия этих данных [12].

Статистическая оценка берет случайную выборку данных и пытается выяснить что-то об общем распределении, обобщая эту выборку. Например, $\sum(y)/len(y)$ является оценкой среднего значения набора данных y . Это метод максимального правдоподобия, потому что он не имеет смещения: он сходится к истинному среднему значению распределения (при достаточно большом количестве выборок). Для большинства задач метод максимального правдоподобия является простейшим оценщиком, но иногда оценка с его помощью невозможна, а в других случаях полезна некоторая степень смещения в оценке.

Импутация максимального правдоподобия – это оценка максимального правдоподобия, применяемая к отсутствующим данным.

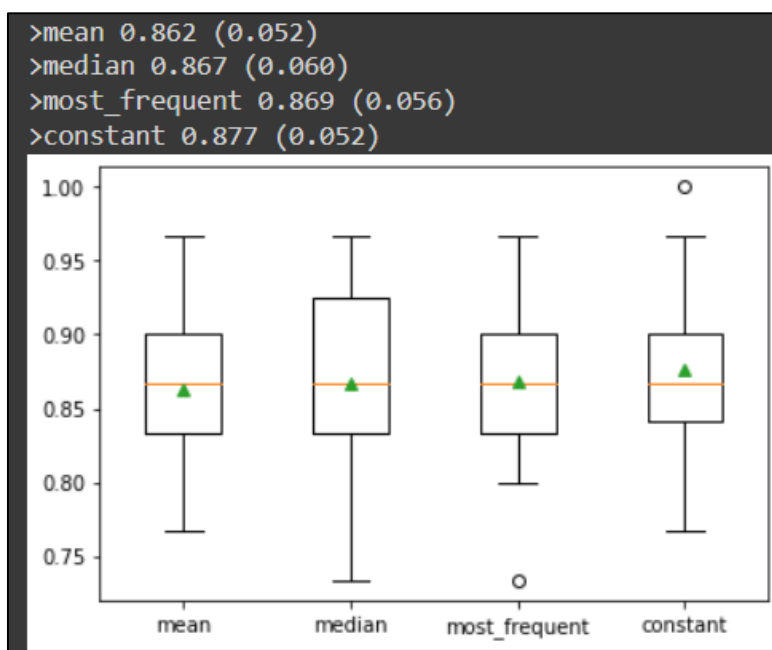


Рис. 2. Результаты импутации с использованием классических методов (среднее по столбцу, медиана, мода, константа)

Все методы, которые обсуждались до сих пор, можно назвать «одионой импутацией»: каждое значение в наборе данных заполняется ровно один раз. В общем, ограничение такой импутации заключается в том, что, поскольку эти методы находят максимально вероятные значения, они не генерируют записи, которые точно отражают распределение базовых данных.

Возьмем, к примеру, крайний случай замены отсутствующих значений в данных средним значением. Если бы мы могли наблюдать недостающие данные, мы, естественно, ожидали бы увидеть в них некоторую изменчивость: экстремальные значения, выбросы и записи, которые не полностью соответствуют «шаблону» данных. Этот шум присущ набору данных, но замена среднего значения не пытается представить его в своем результате. Это приводит к смещению в любых последующих моделях, которые подвержены тенденции (наличию среднего значения в наборе данных), которой нет в базовых данных. Это, в свою очередь, снижает точность.

В литературе [13, 14, 16], вероятно, наиболее продвинутой методологией выполнения отсутствующих данных является множественная импутация. При множественной импутации мы много раз генерируем пропущенные значения из набора данных. Затем отдельные наборы данных объединяются в окончательный импутированный набор данных, при этом значения, выбранные для замены отсутствующих данных, каким-то образом извлекаются из объединенных результатов. Другими словами, множественная импутация разбивается на три этапа: импутация (несколько циклов), анализ (постановка того, как результаты должны быть объединены) и объединение (интеграция результатов в окончательную матрицу импутации).

Любой метод, который следует этой общей схеме, является методом множественной импутации. Таким образом, существует множество доступных алгоритмов и реализаций импутации. Самый популярный алгоритм называется Multiple Imputation by Chained Equation – MICE (рус. – Множественная импутация с помощью связанного уравнения). Алгоритм импутации следующий [13, 15, 16]:

1. Простая импутация, например, среднего арифметического, выполняется для каждого отсутствующего значения в наборе данных. Эти средние значения можно рассматривать как «заполнители».

2. «Заполнитель» означает, что импутация для одной переменной («var») снова устанавливается как отсутствующее значение.
3. Наблюдаемые значения переменной «var» на шаге 2 регрессируют по другим переменным в модели импутации, которая может состоять или не включать все переменные в наборе данных. Другими словами, «var» является зависимой переменной в регрессионной модели, а все остальные переменные являются независимыми переменными в регрессионной модели.
4. Отсутствующие значения для «var» затем заменяются прогнозами из регрессионной модели. Когда «var» впоследствии используется в качестве независимой переменной в регрессионных моделях для других переменных, будут использоваться как наблюдаемые, так и эти заполненные значения.
5. Затем шаги 2–4 повторяются для каждой переменной, для которой отсутствуют данные. Перебор каждой из переменных составляет одну итерацию или «цикл». В конце одного цикла все пропущенные значения были заменены прогнозами регрессий, отражающими взаимосвязи, наблюдаемые в данных.
6. Шаги со 2 по 4 повторяются в течение нескольких циклов, при этом импутации обновляются в каждом цикле. В конце этих циклов окончательные импутации сохраняются, в результате чего получается один импутированный набор данных. Обычно выполняется десять циклов; однако необходимы исследования для определения оптимального количества циклов при импутации данных в различных условиях. Идея состоит в том, что к концу циклов распределение параметров, определяющих импутацию (например, коэффициентов в регрессионных моделях), должно сходиться в том смысле, что становится стабильным.

Применяя алгоритм MICE, на рис. 3 мы уже можем наблюдать график, отображающий качество модели. Условия для эксперимента остаются прежними. Точность алгоритма MICE составляет 0,881. Тем не менее, местами также наблюдается потеря сезонности. В случае применения константы отличие составляет 78,8%, MICE отличается от оригинала на 36%.

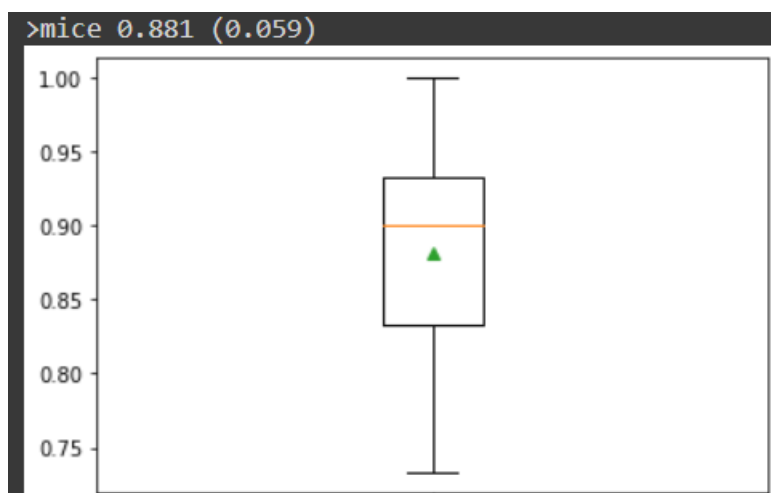


Рис. 3. Результаты импутации с использованием MICE

3. Результаты. Для тестирования результатов в прогнозировании, полученных в результате эксперимента по выбору стратегии импутации, был выбран четырехлетний отрезок, где будет применен MICE. Результаты применения MICE на четырехлетнем отрезке представлены на рис. 4, точность составляет 0,891. График выборки с примененной импутацией продемонстрирован на рис. 5.2.

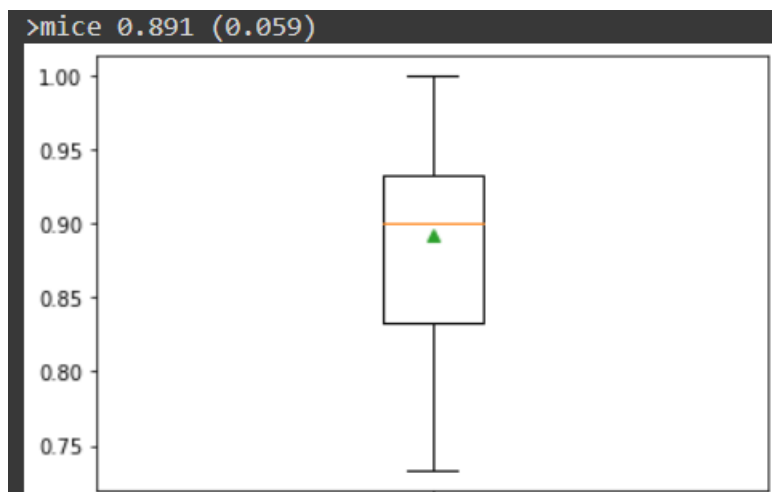


Рис. 4. Результаты импутации с использованием MICE на четырехлетнем отрезке

В результате исследования был испытан алгоритм SARIMA для прогнозирования погодных данных. На рис. 5.3 представлены результаты применения SARIMA: график слева – результат модели, обученной на четырехлетнем отрезке, где был применен MICE для импутации данных; график справа – результат модели, обученной на оригинальных данных, где не было потерь записи и не применялась импутация. Критерий RMSE (среднеквадратическое отклонение) для графика слева на рис. 5.3 составил 0,309, для графика справа RMSE составил 0,201. RMSE в обоих случаях не превышает 0,5, что говорит о способности модели точно предсказывать данные. На рис. 5.1 представлен график солнечной радиации, где синим цветом выделены оригинальные данные за семь лет.

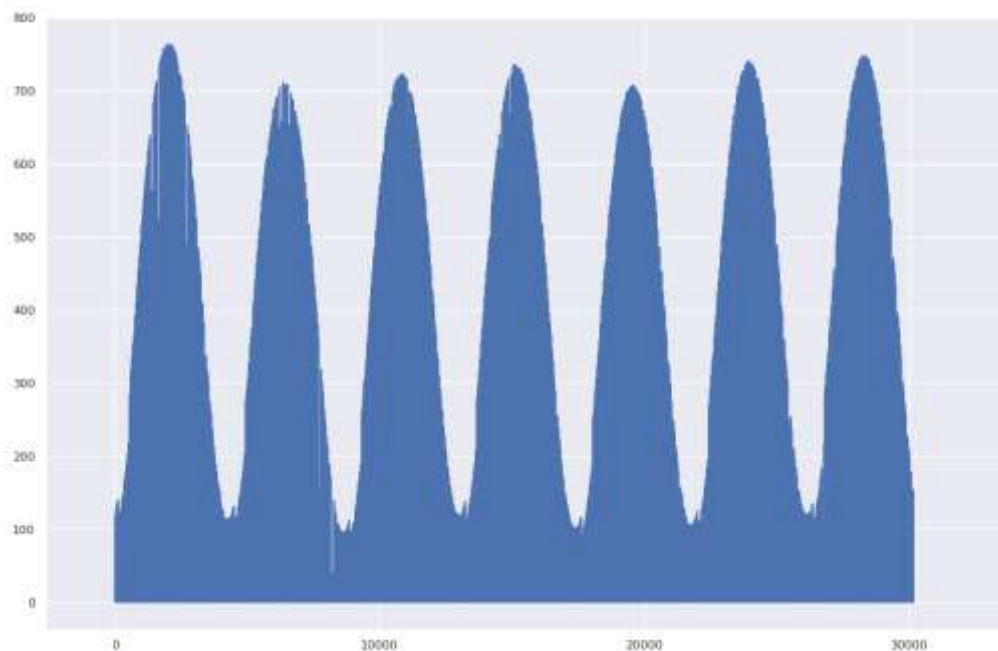


Рис. 5.1. График солнечной радиации, первоначальные данные

Вывод. Опытным путем были получены правдоподобные результаты и обосновано применение алгоритма SARIMA для прогнозирования солнечной радиации. Также было установлено, что в данном случае MICE является наиболее подходящим алгоритмом импутации. Исходя из обнаруженной проблемы потери сезонности данных после применения.



Рис. 5.2. График солнечной радиации, выборка с применением импутации

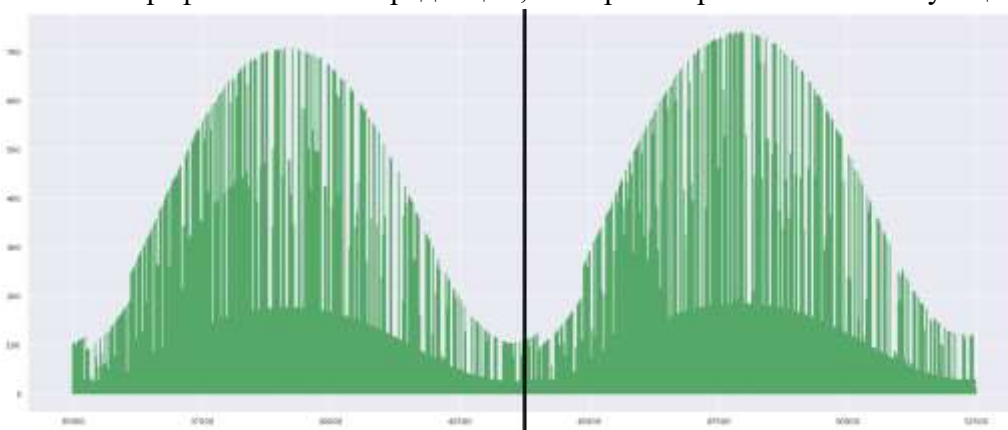


Рис. 5.3. Графики солнечной радиации, спрогнозированные моделью SARIMA на основе данных с применением импутации (слева) и на основе первоначальных данных (справа)

Комплексное рассмотрение параметров, таких, как влажность, скорость ветра и температура воздуха, позволит не только кардинально пересмотреть контекст нулевых значений, но и разработать более совершенную модель прогноза данных.

Благодарности. Результаты получены в рамках выполнения базового проекта ИСЭМ СО РАН АААА-А21-121012090007-7 по госзаданию FWEU-2021-0007.

СПИСОК ИСТОЧНИКОВ

1. Массель Л.В., Массель А.Г., Ворожцова Т.Н. Использование научной ИТ-инфраструктуры исследований энергетики для построения цифровых теней и цифровых двойников объектов и систем энергетики /Л.В. Массель, А.Г. Массель, Т.Н. Ворожцова // Применение технологий виртуальной реальности и смежных информационных систем в междисциплинарных задачах // Труды Международной научной конференции. – М.: Знание, 2020. – С. 163-169. – ISBN 978-5-907345-75-1. – DOI 10.38006/907345-75-1.2020.1.310
2. Huang R., Huang T., Gadh R., Li N. Solar generation prediction using the ARMA model in a laboratory-level micro-grid. 2012 IEEE Third International Conference on Smart Grid Communications (SmartGridComm). Tainan, Taiwan: IEEE, 11/5/2012 - 11/8/2012, pp. 528–533.
3. Badescu V. Modeling solar radiation at the earth's surface: Recent advances. Viorel Badescu, (ed.), Berlin, Springer, 2008, URL: <http://www.loc.gov/catdir/enhancements/fy0904/2007942168-d.html>.
4. Folch-Fortuny A., Arteaga F., Ferrer A. Missing Data Imputation Toolbox for MATLAB. Chemometrics and Intelligent Laboratory Systems. 2016, 154, pp. 93–100.

5. Fara L., Diaconu A., Craciunescu D., Fara S. Forecasting of Energy Production for Photovoltaic Systems Based on ARIMA and ANN Advanced Models. International Journal of Photoenergy. 2021, pp. 1–19, URL: <https://doi.org/10.1155/2021/6777488>
6. Huang, J., Liu, H. A hybrid decomposition-boosting model for short-term multi-step solar radiation forecasting with NARX neural network. Journal of Central South University, 28(2), pp. 507-526, URL: <http://doi.org/10.1007/s11771-021-4618-9>.
7. Metcalfe A. V., Cowpertwait P. S. Introductory Time Series with R. NY: Springer New York, 2009, 262p.
8. Shumway R. H., Stoffer D. S. Time Series Analysis and Its Applications. NY: Springer New York, 2011.
9. Фомина Е.Е. Обзор методов и программного обеспечения для восстановления пропущенных значений в массивах социологических данных // Гуманитарный вестник /Е.Е. Фомина. 2019. – №4 (78). – URL: <https://cyberleninka.ru/article/n/obzor-metodov-i-programmnogo-obespecheniya-dlya-vosstanovleniya-propuschnnyh-znacheniy-v-massivah-sotsiologicheskikh-dannyh>
10. Канищев И. С. Восстановление пропущенных значений в задачах классификации при несбалансированности данных / И. С. Канищев // Естественные и Технические Науки, 2021. – № 5. – С. 63–66.
11. Jamshidian M., Mata M. Advances in Analysis of Mean and Covariance Structure when Data are Incomplete. Handbook of Latent Variable and Related Models. Elsevier, 2007, pp. 21–44.
12. Hippel P. T. v., Bartlett J. Maximum likelihood multiple imputation: Faster imputations and consistent standard errors without posterior draws, 10/2/2012, 38p., URL: <http://arxiv.org/pdf/1210.0870v10>
13. Van Buuren S. Flexible imputation of missing data. Boca Raton: CRC Press, 2012, URL: <https://stefvanbuuren.name/fimd/>
14. Hayati Rezvan P., Lee K. J., Simpson J. A. The rise of multiple imputation: a review of the reporting and implementation of the method in medical research. BMC medical research methodology. 2015, no.15, 30p., DOI.org/10.1186/s12874-015-0022-1
15. Azur M. J., Stuart E. A., Frangakis C., Leaf P. J. Multiple imputation by chained equations: what is it and how does it work? International journal of methods in psychiatric research. 2011, no. 20 (1), pp. 40–49.
16. Van Buuren S. Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. Journal of Statistical Software. 2011, no. 45 (3).

Лосев Алексей Сергеевич, инженер исследователь отдела систем искусственного интеллекта в энергетике, Институт систем энергетики им. Л.А. Мелентьева СО РАН, lo-sev.as@isem.irk.ru, Россия, Иркутск, Лермонтова 130

Массель Алексей Геннадьевич, кандидат технических наук, старший научный сотрудник отдела систем искусственного интеллекта в энергетике, Институт систем энергетики им. Л. А. Мелентьева СО РАН, amassel@gmail.com, AuthorID: 173012, SPIN-код: 2940-2510, Россия, Иркутск, Лермонтова 130

Time series forecasting and data imputation for the digital twin of a solar power plant

Aleksei S. Losev, Alexey G. Massel

Melentiev Energy Systems Institute Siberian Branch of the Russian Academy of Sciences, Russia, Irkutsk, losev.as@isem.irk.ru

Abstract. The article presents the implementation of an autoregressive model for predicting solar radiation data. An experiment was carried out with some approaches to data imputation and the choice of the MICE (Multiple Imputation by Chained Equation) approach was justified. A comparative experiment was also carried out for predicting solar radiation based on initial data and based on data in which imputation was applied.

Keywords: Digital twin, imputatuion, solar radiation, autoregression, forecast, ARIMA, SARIMA, MICE

Acknowledgements: The results were obtained as part of the implementation of the basic project ISEM SB RAS AAAA-A21-121012090007-7 under the state order FWEU-2021-0007.

REFERENCES

1. Massel L.V., Massel A.G., Vorozhtsova T.N. Ispol'zovaniye nauchnoy IT-infrastruktury isledovaniy energetiki dlya postroyeniya tsifrovyykh teney i tsifrovyykh dvoynikov ob'yektov i sistem energetik [Using the scientific IT infrastructure of energy research to build digital shadows and digital twins of energy objects and systems]. *Primeneniye tekhnologiy virtual'noy real'nosti i smezhnykh informatsionnykh sistem v mezhdistsiplinarnykh zadachakh* [Application of virtual reality technologies and related information systems in interdisciplinary tasks]. Proceedings of the International Scientific Conference. M., Znaniye,, 2020, pp. 163-169, ISBN 978-5-907345-75-1, DOI 10.38006/907345-75-1.2020.1.310
2. Huang R., Huang T., Gadh R., Li N. Solar generation prediction using the ARMA model in a laboratory-level micro-grid. 2012 IEEE Third International Conference on Smart Grid Communications (SmartGridComm). Tainan, Taiwan: IEEE, 11/5/2012 - 11/8/2012, pp. 528–533.
3. Badescu V. Modeling solar radiation at the earth's surface: Recent advances. Viorel Badescu, (ed.), Berlin, Springer, 2008, URL: <http://www.loc.gov/catdir/enhancements/fy0904/2007942168-d.html>.
4. Folch-Fortuny A., Arteaga F., Ferrer A. Missing Data Imputation Toolbox for MATLAB. *Chemometrics and Intelligent Laboratory Systems*. 2016, 154, pp. 93–100.
5. Fara L., Diaconu A., Craciunescu D., Fara S. Forecasting of Energy Production for Photovoltaic Systems Based on ARIMA and ANN Advanced Models. *International Journal of Photoenergy*. 2021, pp. 1–19, URL: <https://doi.org/10.1155/2021/6777488>
6. Huang, J., Liu, H. A hybrid decomposition-boosting model for short-term multi-step solar radiation forecasting with NARX neural network. *Journal of Central South University*, 28(2), pp. 507-526, URL: <http://.doi.org/10.1007/s11771-021-4618-9>.
7. Metcalfe A. V., Cowpertwait P. S. *Introductory Time Series with R*. NY: Springer New York, 2009, 262p.
8. Shumway R. H., Stoffer D. S. *Time Series Analysis and Its Applications*. NY: Springer New York, 2011.

9. Fomina E.E. Obzor metodov i programmno obespecheniya dlya vosstanovleniya propuschnykh znacheniy v massivakh sotsiologicheskikh dannykh [Review of methods and software for recovering missing values in arrays of sociological data]. Gumanitarnyy vestnik [Humanities Bulletin]. 2019, no. 4 (78), URL: <https://cyberleninka.ru/article/n/obzor-metodov-i-programmnogo-obespecheniya-dlya-vosstanovleniya-propuschnykh-znacheniy-v-massivakh-sotsiologicheskikh-dannykh>
10. Kanishchev I. S. Vosstanovleniye propuschnykh znacheniy v zadachakh klassifikatsii pri nesbalansirovannosti dannykh [Recovery of missing values in classification problems with data imbalance]. Yestestvennyye i Tekhnicheskiye Nauki [Natural and Technical Sciences]. 2021, no. 5, pp. 63–66.
11. Jamshidian M., Mata M. Advances in Analysis of Mean and Covariance Structure when Data are Incomplete. Handbook of Latent Variable and Related Models. Elsevier, 2007, pp. 21–44.
12. Hippel P. T. v., Bartlett J. Maximum likelihood multiple imputation: Faster imputations and consistent standard errors without posterior draws, 10/2/2012, 38p., URL: <http://arxiv.org/pdf/1210.0870v10>
13. Van Buuren S. Flexible imputation of missing data. Boca Raton: CRC Press, 2012, URL: <https://stefvanbuuren.name/fimd/>
14. Hayati Rezvan P., Lee K. J., Simpson J. A. The rise of multiple imputation: a review of the reporting and implementation of the method in medical research. BMC medical research methodology. 2015, no. 15, 30p., DOI.org/10.1186/s12874-015-0022-1
15. Azur M. J., Stuart E. A., Frangakis C., Leaf P. J. Multiple imputation by chained equations: what is it and how does it work? International journal of methods in psychiatric research. 2011, no. 20 (1), pp. 40–49.
16. Van Buuren S. Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. Journal of Statistical Software. 2011, no. 45(3).

Aleksei S. Losev, Research Engineer of the Department of Artificial Intelligence Systems in the Energy Industry of the Institute of Energy Systems named after. L. A. Melentiev SB RAS, losev.as@isem.irk.ru, Russia, Irkutsk, st. Lermontova d.130

Alexey G. Massel, Ph. D, Senior Research Fellow of the Department of Artificial Intelligence Systems in the Energy Industry of the Institute of Energy Systems named after. L. A. Melentiev SB RAS, amassel@gmail.com, AuthorID: 173012, SPIN: 2940-2510, Russia, Irkutsk, st. Lermontova d.130

Статья поступила в редакцию 05.03.2022; одобрена после рецензирования 25.03.2022; принята к публикации 29.03.2022.

The article was submitted 03.05.2022; approved after reviewing 03.25.2022; accepted for publication 03.29.2022.

Моделирование и программирование

УДК 004.43 (042.4)

doi:10.38028/ESI.2022.25.1.009

Место функционального программирования в организации параллельных вычислений

Городняя Лидия Васильевна

Институт систем информатики СО РАН,

Россия, Новосибирск, gorod@iis.nsk.su

Аннотация. Статья посвящена результатам анализа современных тенденций функционального программирования, рассматриваемого как метапарадигма решения проблем организации параллельных вычислений и многопоточных программ для многопроцессорных комплексов и распределённых систем. Учитывая мультипарадигмальность языков параллельного программирования, здесь использован парадигмальный анализ, позволивший выявить ряд особенностей и прогнозировать ход процессов применения программ, а также их изучения и разработки. Есть основания рассчитывать, что функциональное программирование способно повышать производительность программ с помощью предварительной подготовки их прототипов. Приведено описание семантических (универсальность, самоприменимость, равноправие параметров) и прагматических (гибкость ограничений, неизменяемость данных, строгость результата) принципов функционального программирования, а также следствий из этих принципов, нацеленных на снижение трудоёмкости и повышение технологичности создания прототипов, автономно развиваемых модулей и отладки программ (конструктивность, верификация, факторизация, непрерывность процессов, обратимость действий, унарные функции). Отмечена роль парадигмальной декомпозиции программ в технологии разработки долгоживущих программ, какими нередко бывают программы параллельных вычислений. Особо подчёркнута перспектива функционального программирования (ФП) как вспомогательной универсальной метапарадигмы решения сложных и важных задач, обременённых трудно удостоверяемыми и слабо совместимыми требованиями эффективности, надёжности, безопасности и улучшаемости. Показано разнообразие парадигмальных характеристик, присущих подготовке и отладке долгоживущих программ параллельных вычислений.

Ключевые слова: функциональное программирование, параллельные вычисления, языки программирования, парадигмы программирования, мультипарадигмальность

Цитирование: Городняя Л. В. Место функционального программирования в организации параллельных вычислений // Информационные и математические технологии в науке и управлении. – 2022. – № 1 (25). – С. 102-119. – DOI:10.38028/ESI.2022.25.1.009.

Введение. Функциональное программирование - одна из первых парадигм, нацеленных не просто на получение эффективной реализации заранее подготовленных алгоритмов, а на решение новых задач информационной обработки, обладающих исследовательским компонентом [1, 2]. Если множество языков программирования стремительно растёт, то число парадигм не столь велико. Оно допускает достаточно полное сравнение и выбор характеристик для чёткого разделения парадигм, уяснения причин их разнообразия и сферы

их успешного применения [3, 4]. Рассматривая результаты анализа проблем, средств и методов организации параллельных вычислений, а также подготовки многопоточных программ для многопроцессорных комплексов и распределённых систем, можно обратить внимание на явно широкий спектр постановок задач и соответствующих им приоритетов в принятии решений на разных этапах разработки и отладки программ. В результате методика парадигмального анализа позволяет сводить постановки решаемых задач к комплектам автономно развиваемых подзадач, оценивать их сходство и различия, учёт которых необходим при прогнозировании сложности процессов применения программируемых решений, начиная с планирования, изучения и организации разработки программ решения таких подзадач для долгоживущих программ, какими часто оказываются программы параллельных вычислений [5].

Изложение начинается с рассмотрения ряда парадигм параллельных вычислений, поддержанных в известных языках программирования (ЯП). Затем идёт обсуждение основных принципов и ограничений, обнаруживающихся при переносе техники ФП на параллельные вычисления. В заключении описаны требования к языковой поддержке параллельного программирования.

1. Общее представление. Рассматривая идеи функционального программирования и практику их применения при анализе проблем, средств и методов организации параллельных вычислений, следует обратить внимание, что уже созданы значительное число языков и систем параллельного программирования. Используя систематизацию парадигм программирования на основе различий в приоритетах принятия программируемых решений, можно сделать вывод, что одной из причин сложности разработки программ параллельных вычислений является их скрытая мультипарадигмальность [6]. Для разработки параллельной программы многое требуется продумывать по-разному одновременно в различных парадигмах, удобных для решения отдельных подзадач без традиции и навыков решения полного комплекса подзадач в единой обстановке. Наиболее очевидно парадигмальные различия видны при решении задач масштабирования вычислений на различные многопроцессорные комплексы, синхронизации взаимодействий над локальной и общей памятью в многопоточных программах, представлении естественного асинхронного параллелизма уровня постановок задач и достижения высокой производительности программ с учётом критерия полноты загрузки доступных много-процессорных комплексов или распределённых систем. Отдельно имеет место образовательная проблема – ознакомление с феноменами параллелизма, дающее специалистам интуитивное понимание методов решения особо трудных задач. Новые работы по ФП существенно нацелены на поиск более производительных решений проблем параллельного программирования [7]. Созданы заметное число языков и систем программирования (ЯиСП), позволяющих решать отдельные из этих задач в рамках конкретных парадигм, поддержка которых представлена в разных языках программирования (табл. 1).

Таблица 1. Проблемы ПВ, поддержанные в разных языках программирования и инструментах

№	Проблема	Парадигма	ЯП и API
1	Масштабирование	Многопроцессорное программирование	VHDL, XC, СИГМА, bash, Occam, mpC, Эль-76, Limbo, Kotlin, MPI
2	Синхронизация потоков	Синхронное (синхронизирующее) программирование	APL, VAL, Sisal, Alef, E, X10, LuNA, Charm, Go, Java, Scala, Rust, Пифагор, OpenMP
3	Постановки задач	Асинхронное программирование	БАРС, Haskell, Erlang, JavaScript, Python, C#

4	Производитель-ность программ	Высокопроизводитель-ное программирование	Setl, HPF, G, Sparkel, mpC, Sanscript, D, Rest, F#
5	Ознакомление с параллелизмом	Учебное программирование	Logo, Робик, Karel, OZ, A++, СИНХРО, Lsl

Таким образом, для каждой из трудно решаемых задач параллельных вычислений уже сформирована отдельная удобная парадигма её решения и создан ряд языков программирования, поддерживающих такую парадигму, что подтверждает её статус как парадигмы. Различие между парадигмами проявляется в упорядочении важности средств и методов, используемых при решении отдельных задач, другие задачи требуют иного упорядочения. В каждый момент разработки программы обычно используется одна парадигма. Соответственно и в разных языках программирования выделяется одна ведущая парадигма. Требования к решению достаточно сложных задач параллельных вычислений связано с целым рядом трудных задач, что влечёт необходимость использования разных парадигм на разных этапах их создания и фазах их жизни. При переходе к технологии параллельного программирования важна гарантия получения практического результата, что требует поддержки полного спектра парадигм, используемых на разных этапах разработки программ. Трудоемкость использования разных парадигм при решении одной задачи обычно минимизируется созданием много-языковых систем, допускающих по мере необходимости возможность перехода от одной парадигмы к другой без затрат на освоение разных интерфейсов. Это приводит к целесообразности создания мультипарадигмального языка параллельных вычислений, поддерживающего одновременно все основные парадигмы параллелизма и разные уровни абстрагирования – от более низкого, чем принято в языках высокого уровня, до языков сверхвысокого уровня [8].

Как показывает опыт языков БАРС и Haskell, в таких случаях удобно декомпозировать определение ЯП на отдельные подязыки, поддерживающие основные парадигмы или монады, нацеленные на конкретные модели подготовки и представления программ. Важно на каждом этапе разработки программы локализовать использование одной парадигмы, характеризующейся сравнительно небольшим набором средств и методов в рамках одного способа мышления. Каждая парадигма имеет своё наполнение категорий семантических систем и своё упорядочение их роли в процессе программирования [9, 10].

Средства многопроцессорного программирования обычно опираются на данные и характеристики доступной архитектуры, включая основной многопроцессорный комплекс и взаимосвязи между его элементами. Оперирование комплексом позволяет инициировать процессы функционирования отдельных процессоров, их блокировку, возобновление и отмену процессов. По ходу процессов возможны обмены данными по определённым протоколам, результаты которых можно рассматривать как цель программы. Принятие решений начинается с определения пространства возможных многопроцессорных комплексов, что можно рассматривать как особую разновидность памяти со своей дисциплиной функционирования и взаимодействия элементов. Далее происходит выбор подходящих конфигураций и структур пространства итерирования процессов, предназначенных для выполнения на отдельных процессорах. Затем полученная схема управления процессами наполняется собственно действиями, выполняющими вычисления. Обычно предпочитают процедуры без рекурсии, освобожденные от ряда сложностей управления вычислениями и побочных эффектов над общей памятью. Строится

многопроцессорная программа, допускающая в динамике реконфигурацию многопроцессорного комплекса¹.

При синхронном (синхронизирующем) многопоточном программировании выделяются достаточно чёткие схемы управления вычислениями и разделяются регулярные участки и типичные модели программы, удобные для распараллеливания. Главное отличие от многопроцессорного программирования сводится к абстрагированию от реальных структур конкретных комплексов. Потоки формируются ближе к уровню постановки задачи. Обычно выделяются фрагменты, свободные от побочных эффектов в памяти, и допускается не императивное управление временем исполнения потоков программы с учётом иерархии схемы управления программой и некоторых временных отношений. Принятие решений начинается с выбора стандартных схем управления, используется понятие «пространство итерирования», которое можно структурировать в зависимости от распределения данных и методов их хранения, что может влиять на эффективность и дисциплину обработки многоуровневой памяти. Управление итерированием может использовать произвольные предикаты. Схема управления над пространством итерирования потоков наполняется сравнительно простыми для отладки фрагментами, возможно отлаженными заранее или программируемыми автономно. Для вывода результатов программы выделяются специальные средства формирования результатов вычислений, полученных на равноправных потоках многопоточной программы. Таким образом, получается многопоточная программа с динамически изменяемым пространством потоков над локальной памятью с возможностью эпизодической синхронизации их отдельных фрагментов².

Асинхронное программирование нацелено на предельное выражение независимых элементов программы, обусловленных природой решаемой задачи, что может быть базой для максимального распараллеливания при условии представления специальных схем организации вычислений, возможно с учётом специфики доступного оборудования. Начинается принятие решений с выбора общих схем управления действиями и представления условий их срабатывания. Действия могут использовать ту или иную дисциплину обработки памяти. Поддерживается неявное и/или программируемое разнообразие дисциплин доступа к памяти, включая иерархию неоднородной памяти, и схем вычислений – фрагментов, наполняющих общую схему программы процедурами или библиотечными модулями. Возможна схема программируемых синхросетей, асинхронно управляющая выполнением действий в зависимости от условий их готовности к выполнению.

Для высокопроизводительного программирования необходим переход от отдельного прогона программы к учёту перспектив её многократного применения и улучшения. Появляется возможность использовать недогруженные мощности многопроцессорных комплексов выполнением фрагментов программы в расчёте на предстоящие в будущем прогоны с данными, возможно востребованными в дальнейших прецедентах её отладки и применения. Примерно так организуют программы, предназначенные для сверхбыстрого реагирования на опасные события, например, при прогнозировании цунами. Принятие решений начинается со схем и моделей вычислений, возможно, над общей памятью, но с приоритетом локальной памяти. Управление вычислениями учитывает особенности многократного выполнения программы при её отладке и применении, включая возможность наследования результатов между сеансами и сравнения измеримых характеристик производительности версий программы. Получа-

¹ <https://mooc.tsu.ru/mooc-openedu/mpi/> «Курс Параллельное программирование с использованием OpenMP и MPI»

² <https://habr.com/ru/post/121925/> Курс «Основы MPI»

ется ряд улучшаемых версий программы, выбор одной из которых может учитывать особенности текущих условий применения, включая конфигурацию многопроцессорного комплекса и требования к измеримым характеристикам производительности программ.

Учебное программирование призвано компенсировать недостатки общепринятых образовательных ориентиров на безошибочность, единственность, последовательность, императивность и однозначность решений при решении любых задач. Такие ориентиры затрудняют обучение методам программирования вообще, а параллельных вычислений тем более. Учебное программирование может давать старт ко всем необходимым парадигмам параллельных вычислений на базе опыта применения и конструирования игровых программ типа визуализации роботов. Это достаточная причина для мультипарадигмальности учебных языков программирования, обычно поддерживающих некоторые средства представления параллельных вычислений.

Долгоживущие ЯП, как и новые ЯП нашего века, обычно мультипарадигмальны. Есть основания для вывода, что успешная практика параллельного программирования требует мультипарадигмальной поддержки полного спектра парадигм параллельных вычислений. Необходимо допускать их развитие, пополнение и применение по мере необходимости с возможностью перехода к очередной парадигме без изменения общезыковой и системной обстановки.

Жаргон современного практического программирования использует понятие «язык программирования» как «входной язык или расширенное подмножество ЯП типовой системой программирования (СП), функционирующей на базе определённой конфигурации оборудования». Различие заключается в том, что СП обычно сопровождает реализацию ЯП расширяемым комплектом библиотечных модулей и может не поддерживать отдельные сложности семантики ЯП. В результате происходит сглаживание видимых на практике различий между разными ЯиСП. Кроме того, прямые измерения трудоёмкости программирования и производительности программ почти не отражают зависимости результата от принятых программистом решений и выбора конструкций ЯП. Хотя программируемые решения представляются в терминах ЯП, их влияние растворяется в весьма сложном комплексе, наследующем производительность СП и оборудования с большим доминированием характеристик элементной базы и аппаратуры. Таким образом, существует проблема создания методики, позволяющей выявлять такие зависимости совмещением прямых измерений с результатами экспертных оценок особенностей ЯиСП, возможно, отличающихся от оценок ЯП.

Понятийная сложность решаемых задач ПВ выходит за пределы обычного образования и сами понятия программирования обретают более широкое толкование, отчасти противоречащее привычному интуитивному пониманию. Представление результатов оценки понятийной сложности, позволяющее структурировать пространство таких параметров, рассмотрено в статье [11].

Успех функционального программирования на практике существенно зависит от выбора постановок задач, решения которых представляются системами функций, сравнительно простых, не слишком трудоёмких и удобных при отладке [12, 13]. По степени изученности существенно различаются следующие категории постановок задач, влияющие на выбор методов решения задач и трудоёмкость их программирования:

- новые;
- исследовательские;
- практические;
- точные.

Новые постановки задач характеризуются отсутствием доступного прецедента практического решения задачи, новизной используемых средств или недостатком опыта исполнителей. Задачи параллельных вычислений поставлены ещё в докомпьютерную эпоху и постановки многих таких задач обладают математической точностью [5]. Тем не менее часть задач параллельного программирования их решений приходится рассматривать как новые из-за стремительного обновления информационных технологий, элементной базы и нерешённых образовательных проблем программирования в целом. Любая проблема, не получившая хорошего решения, остаётся в статусе новой задачи независимо от времени её постановки.

Исследовательские постановки задач параллельного программирования в настоящее время следуют тенденции функционального подхода и изучения схем структурирования пространства итерирования процессов, позволяющего оптимизировать обработку памяти. Функциональный подход можно рассматривать как методику сведения решений сложных задач к композициям из планарных проекций, достаточно удобных для понимания, анализа и обработки.

Практичные постановки математических задач параллельных вычислений, нацеленные на актуальность и удобство применения, преимущественно используют мощностные доступных ИТ для воспроизведения математических моделей, ранее ограниченных низкой эффективностью оборудования, а теперь получивших перспективу стать новой информационной революцией.

Точные постановки большинства задач параллельного программирования сложились на базе последовательных алгоритмов и включают в себя испытание возможностей используемых средств, связанных с мерой организованности ранее созданной императивной программы. Некоторые сложности связаны с использованием однопроцессорных конфигураций как исходной модели параллельных многопоточных программ. Не исключено, что для предельного случая удобнее рассматривать двухпроцессорные конфигурации. Появление собственно параллельных алгоритмов встречается не так уж часто, хотя в середине 1990-ых годов проводились конференции такого направления.

Обучение параллельным вычислениям входит в образовательные программы многих университетов, что достаточно для понимания их сложности и постановки задач их исследования, нередко на младших курсах и сразу вслед за функциональным программированием до императивного. Проблемой является переход от теории параллельных вычислений к практике реализации высокопроизводительных программ, удовлетворяющих особо сложным, трудно удостоверяемым критериям надёжности и безопасности. Разница в требованиях, нацеленных на понимание смысла решаемых задач, развитие области приложения решений, определение границы применимости полученных результатов и достижение системности при обобщении созданных инструментальных средств важных для решения задач современных ИТ.

Характерной чертой системного подхода как ведущего метода программирования является переход к классам задач при содержательном анализе постановок задач. Границы класса устанавливаются выбором процесса решения задач. Переход к экспериментам на суперкомпьютерах показал, что именно системные решения могут дать весомый вклад в производительность параллельных вычислений, причём такой вклад может превышать теоретические прогнозы. Это можно рассматривать как обоснование необходимости более фундаментального подхода к программированию, особенно к системному программированию и его математическим основам, дающим путь к созданию высокопроизводительных программ. При создании, формировании и исследовании математических моделей как фундаментальной базы для решения особо трудных проблем эффективности, надёжности и безопасности программного обеспечения важную роль играет развитие моделей, связанных со временем и ре-

курсами, слабо представленными в курсах по классической математике [14]. Строго говоря, программирование, в отличие от математики, вообще не имеет дела ни со значениями, ни с функциями. Программирование работает с данными, которые могут представлять значения функций.

Квалификация системного программиста включает в себя как формирование способностей самостоятельно изобретать решения новых задач, так и навыки ответственно повышать качество готовых решений, наряду с глубоким владением неклассической и фундаментальной математикой и ознакомлением с современной физикой, психологией и лингвистикой, что слишком неудобно для образовательной системы, нацеленной на формальный контроль усвоения устойчивых стереотипов, представленных в учебно-методической литературе и противоречит учебно-методическим традициям.

Экстенсивное развитие ИТ заметно опережает способности человека оперативно осваивать новые возможности аппаратуры и системных средств ИТ, выходящие за пределы пользовательского уровня, поддерживаемого поставщиками инструментария и программных продуктов. Миссия системного программирования – создание инструментария, позволяющего повышать качество информационных систем, включая поиск новых решений по обеспечению надежности и безопасности информационных технологий [5, 15].

2. Принципы функционального программирования. Обычно ФП подразумевает поддержку ряда семантических и прагматических принципов, способствующих созданию функциональных моделей на этапе исследовательского компьютерного эксперимента, естественного при решении новых и сложных задач. В их числе ФП поддерживает *внешние* семантические принципы представления алгоритмов, такие, как универсальность, самоприменимость, равноправие параметров, и системно реализационные *внутренние* прагматические принципы поддержки информационной обработки, такие, как гибкость ограничений, неизменяемость данных, строгость результата. Семантическим принципам программист следует при подготовке программы. Прагматические принципы обеспечивает система программирования, освобождая программиста от технически сложных непринципиальных решений, не зависящих от природы задачи. Многие из этих принципов были заложены Дж. Маккарти в первых реализациях языка Lisp [1].

Универсальность или всеобщность. Понятия «функция» и «значение» представляются теми же средствами, как и любые данные для компьютерной обработки. Исторически близкое понятие – «принцип хранимой программы».

Этот принцип позволяет строить представления функций из их частей и вычислять части по мере поступления и обработки данных. Нет принципиальных ограничений на манипулирование средствами языка, функциями из определения семантики языка, конструкциями реализации языка в системе программирования и выражениями в программе. Всё, что понадобилось при реализации языка программирования, может пригодиться при его применении. Поскольку не существует препятствий для обработки представлений функций теми же средствами, какими обрабатываются данные, постольку представления функций можно строить из их частей – символов. Их даже можно формировать по ходу процесса вычислений, поступления и обработки информации о них.

Самоприменимость. Представления функций прямо или косвенно могут использовать сами себя, что позволяет строить ясные лаконичные рекурсивные символьные формы.

Примеры самоприменимости дают многие математические функции, особенно рекурсивные, такие, как факториал, числа Фибоначчи, суммирование рядов и многие другие, определение которых использует математическую индукцию. В технологии программирования некоторым сходством обладают методы пошаговой или непрерывной разработки программ (рас-

крутки), а также экстремального программирования. Эти методы сводят организацию процесса программирования к ряду шагов, каждый из которых даёт или работоспособную часть программы, или инструмент для выполнения очередных шагов раскрутки.

Равноправие параметров. Порядок и метод вычисления параметров не имеют значения. Можно обратить внимание, что параметры при вызове функции вычисляются на одном и том же уровне иерархии, в одном и том же контексте. Поэтому порядок вычисления параметров не имеет значения, может быть произвольным. Все параметры функции определены в контексте, не изменяемом при их вычислении, они не должны быть взаимосвязаны, не могут воздействовать друг на друга. Часть параметров вычисляются до вызова функции, другие могут быть вычислены позднее, но в том же самом контексте. Поэтому представление любой выделенной формулы из определения функции можно представлять и как параметр этой функции. Любая символьная форма в определении функции может быть выделена из него как параметр и, наоборот, подставлена в него. Функции-переменные допустимы равноправно с обычными функциями-константами и могут быть значениями аргументов или выработаны в качестве результатов других функций.

Обеспечение многократного использования данных выполняется с помощью именования. Параметры имеют имена, часто называемые переменными, хотя в функциональном программировании они не меняют значений в рамках одного контекста. Чисто на уровне понятий переменная – именованная часть памяти, предназначенная для многократного доступа к изменяющимся данным, а константа – к неизменным данным. Изменение связи между именем и значением в функциональном программировании возможно лишь при переходе в другой контекст, что эквивалентно изменению имени.

Прагматические принципы поддерживает система программирования, точнее, её разработчики. Если семантические принципы достаточно провозгласить в определении языка или парадигмы, то прагматические принципы влекут заметные трудозатраты на уровне создания системы программирования.

Гибкость ограничений. Оперативный пересмотр распределения памяти или её освобождения поддержан для профилактики необоснованных простоев памяти.

Гибкость ограничений на пространственные характеристики позволяет исключать необоснованные простои памяти. Бывает, что не хватает памяти принципиально не на всю задачу, а лишь на отдельные блоки данных, возможно, малосущественные для её решения. Такая проблема в системах функционального программирования решается принципом гибкости ограничений на пространственные характеристики. Эту проблему решает специальная функция – «мусорщик» (garbage collector), пытающаяся при недостатке любой области памяти автоматизировать перераспределение или освобождение памяти. Новые реализации механизма «мусорщик» рационально учитывают преимущества восходящих процессов на больших объемах памяти.

Неизменяемость данных. Представление каждого результата применения функции размещается в новой части свободной памяти без искажения аргументов этой функции, которые могут быть полезны для других функций.

Таким образом, существенно упрощается отладка программ и обеспечивается повторный прогон любых действий. Можно быть уверенным, что все промежуточные результаты сохранены, их можно анализировать и снова использовать в любой момент. Если определение функции – это статическая конструкция, то процесс можно рассматривать как композицию из функций, разворачиваемую по этой конструкции в динамике.

Отдельный, не вполне очевидный, аспект связан с переходом от целых чисел к вещественным, возможно, допускающим изменение точности представления по ходу вычислений. Логически они остаются константами, а система обрабатывает их как переменные.

Строгость результата. Любое число результатов функции может быть представлено как одна символьная форма или структура данных или специальный код, из которых при необходимости можно выбрать нужный результат.

Нередко этот принцип трактуется как требование однозначности функций, что приводит к сомнениям в правомерности функций целочисленного деления, извлечения корня, обратных тригонометрических функций и многих других категорий математических функций. Выражения такому пониманию функций можно видеть ещё в первом издании Фихтенгольца [16].

Представление алгоритмов в виде функциональных программ имеет *практически значимые следствия*. Конструктивность, верификация и факторизация вытекают из семантических принципов. Из прагматических принципов следуют скрытые модели непрерывности процессов, обратимости действий и унарных функций, дающие основу для интуитивного выстраивания функциональных моделей, допускающих удобный компьютерный эксперимент. Кроме того, из ряда семантических и прагматических принципов вытекает поддержка подготовки функциональных моделей программ организации параллельных вычислений, сводимых к комплексам недетерминированных потоков.

Метапрограммирование является следствием принципа универсальности, позволяющего представления программ обрабатывать так же, как любые данные. Это даёт поддержку мета-компиляции, включая синтаксически управляемые методы генерации и анализа программ, а также однородно-гомогенные представления программ, внешне сохраняющих аналогичию или подобие обрабатываемым данным или прототипам, в том числе смешанные и частичные вычисления, оптимизирующие преобразования, макрогенерацию и многое другое, необходимое для создания операционных систем и систем программирования.

Верификация основана на связи принципа самоприменения с методами рекурсии, математической индукции, дедуктивного вывода и логики.

Большинство систем верификации программного обеспечения создаются в рамках функционального программирования. Такие системы делают возможным логически выводить отдельные свойства программ и, благодаря этому, обнаруживать некоторые тонкие ошибки. Если представление объекта имеет подобие некоторой логике вывода, то его свойства могут быть выведены с использованием этой логики. В результате увеличиваются надёжность и безопасность программ, хотя это и не позволяет решить проблему корректности в полной мере. Трудности связаны с недостаточностью классической логики по отношению к неклассическим логикам программирования, а также с некоторыми интуитивными ошибками при оценке истинности логических и вероятностных формул.

Факторизация непосредственно вытекает из принципа равноправия параметров. Любой помеченный фрагмент программы может быть вынесен из её представления и ассоциирован с определённым именем, допуская возможность восстановления исходного представления. Это позволяет декомпозировать программу на автономно развиваемые модули и накапливать правильность программ не только на уровне библиотечных функций, но и на уровне произвольных фрагментов и схем, а также представлять параллельные потоки, ленивые (отложенные) или опережающие вычисления. Можно сказать, что программа приводится в факторизованную форму по тем или иным параметрам в зависимости от цели её преобразования [17]. Благодаря обратимости действий, вытекающей из прагматического принципа неизменяемости данных, процесс отладки обретает сходимость.

Следствия принципов системно-реализационной прагматической поддержки информационной обработки в системах функционального программирования выражаются в использовании интуитивных моделей, таких, как непрерывность процессов (бесконечность), обратимость действий и унарные функции.

Непрерывность процессов интуитивно следует из прагматической поддержки принципа гибкости ограничений.

После выполнения любой функции можно выполнить ещё одну другую функции (команда STOP не является функцией, у нее нет ни аргументов, ни результатов, это просто сигнал процессору перестать работать). Одновременно с выполнением любой функции можно выполнять другие функции, также и перед выполнением функции можно выполнять другие функции. Это позволяет значительную часть работы программы основывать на модели неограниченной памяти без особого беспокойства о ее границах и разнообразии характеристик скорости доступа к различным структурам данных. Во многих языках функционального программирования поддерживается имитация работы с бесконечными структурами данных [9, 18].

Обратимость действий базируется на иллюзии неизменяемости данных, механизмы которой скрыты в СП, их применение почти не требует заботы при подготовке программы и основного объёма отладки. Это позволяет поддерживать механизм мемоизации³ функций на ранее обработанных аргументах. Фактически необходимые изменения данных, такие, как повторное использование памяти, просто автоматизированы. Программист может позволить себе не вмешиваться в реализацию таких средств до тех пор, пока не возникнут проблемы с производительностью.

Унарные функции естественно выводятся из принципа строгого результата, позволяющего при необходимости любое число представленных результатов рассматривать как общую структуру из них. Поскольку результаты часто являются аргументами объемлющих функций, логично возникает и сопутствующий принцип унарных функций. Функция любого числа аргументов может быть преобразована в функцию одного аргумента.

3. Приложения для параллельных вычислений. Параллелизм опирается на целый комплекс семантических и прагматических принципов функционального программирования, что позволяет при необходимости рассматривать любой объем представленных данных и при необходимости реорганизовывать пространство потоков. Такой комплекс делает функциональное программирование удобным для работы с программами, направленными на организацию независимых параллельных процессов. Прежде всего, это принцип равноправия параметров, который гарантирует одинаковый контекст при вычислении параметров функции. Становится возможным представлять независимые потоки и объединять их в многопоточные или многопроцессорные программы, в общий программный комплекс. Кроме того, в параллелизме используются принципы строгого результата и универсальности, что позволяет при необходимости учитывать любое количество представленных потоков и реорганизовывать пространство потоков. Их конкретизация сводится на уровне семантики к методам оттеснения мало вероятных вычислений, балансировки нагрузки, выстраивания пространства итераций.

Чистое функциональное программирование не совсем удобно для моделирования взаимодействующих и императивно синхронизированных процессов. Для организации таких процессов требуется дополнительное уточнение принципов, упрощающее подготовку высокопроизводительных программ параллельных вычислений.

Оттеснение мало вероятных вычислений направлено на предотвращение затрат на формирование и выполнение чрезмерного количества потоков, соответствующих слишком редким ситуациям.

Принцип универсальности имеет два аспекта – равноправие программ с данными и полноту определения функции. При решении задач параллельных вычислений сохраняется аспект равноправия программ с данными, традиционно востребованный в задачах операционных систем и управления процессами. Полнота функции, удобная для построения

³ Мемоизация – сохранение результатов выполнения функций для предотвращения повторных вычислений функций.

программ из универсально отлаженных модулей, может создавать проблемы в многопроцессорных программах из-за увеличения количества потоков, связанных с различными редкими диагностическими ситуациями.

Любой фрагмент, выполнение которого маловероятно или невозможно, может быть представлен как отложенное действие. Ветви для практически неактуальных ситуаций можно перенести в отладочную версию. Иногда эту проблему преодолевают путем выбора выражений, не требующих ветвления, чаще – путем проверки типов данных. Объем необходимой диагностики можно частично сократить за счет статического анализа типов данных.

Балансировка нагрузки сокращает реальное время выполнения многопоточной программы, позволяет выравнивать нагрузку процессоров и избегать их необоснованных простоев, что можно рассматривать как распространение принципа гибкости ограничений на временные границы.

Ленивые или опережающие вычисления дают возможность быстро и оперативно перераспределять нагрузку. Сложное определение функции иногда можно свести к двум функциям, первая из которых выполняет часть определения, откладывая выполнение остальных, а вторая возобновляет выполнение отложенных частей. Возможно, выполнение второй функции произойдет одновременно со второй или даже раньше. В языке mrc предлагается объем вычислений оперативно перераспределять при обнаружении неравномерной загрузки процессоров [19].

Само-применение в виде рекурсивных функций часто рассматривается как осложнение, приводящее к опасному росту стека. Здесь следует отметить, что многие системы функционального программирования предлагают ряд практичных решений. Кроме того, в функциональном программировании существуют отложенные действия, мемоизация, восходящая рекурсия, методы динамического программирования и оптимизация рекурсий за счет сведения к циклам, что во многих случаях позволяет практически устранить чрезмерное разбухание стека. Поддержку работы со стеком в рамках принципа гибкости ограничений можно поддерживать более эффективно, чем в большинстве языков и систем программирования. Факторизация программ на схемы и фрагменты позволяет разделять компоненты по уровню сложности отладки и наследовать корректность ранее отлаженных составляющих. Используются функции, не требующие предварительного вычисления параметров, например макротехника.

Пространство итераций используется в качестве основного параметра для управления распараллеливанием циклов.

Любое конечное множество может работать, как пространство итераций для определенной на нем функции. Если есть набор данных, на котором вычисление функции на одном элементе не требует ее результатов на других, то использование такого набора в качестве пространства итераций удобно для одновременного выполнения этой функции на всех элементах набора. Подобная техника распространения операций и функций имеется в языках APL, Альфа, BAPC и Sisal. Роль равноправия параметров растет, она обеспечивает решение проблем реорганизации потоков при настройке на различные конфигурации многопроцессорных систем, требующих декомпозиции программы на схемы и фрагменты. Многие работы в сфере суперкомпьютерных вычислений по существу являются поиском удобного пространства итераций. Техника и концепция пространств итераций убедительно поддерживаются в языке Sisal, в котором пространства итераций строятся над перечислимыми множествами с использованием операций скалярного и декартового произведения.

Несколько сложнее с прагматическими принципами, требующими пересмотра системных решений на уровне разработки СП. Здесь важную роль играют автоматическое

распараллеливание, идентичность повторных прогонов и представление много-контактных фрагментов.

Автоматическое распараллеливание заключается в извлечении из программы автономных частей, допускающих независимое выполнение. Если существует функция с известным временем выполнения, которую можно разложить на ряд независимых функций с меньшим временем счёта, то они могут быть выполнены в произвольном порядке, и время выполнения исходной функции, как правило, будет строго меньше.

Идентичность повторных прогонов для целей отладки программ и измерения их производительности. Переход к суперкомпьютерам показал, что при слишком большом числе процессоров исчезает нужная для отладки и измерений возможность пронаблюдать повторное исполнение программы. При очередном прогоне могут быть сбои на других процессорах. Функциональное программирование здесь может допускать специальные интерпретации программы, учитывающие протоколы и результаты ранее выполненных прогонов с отслеживанием идентичности исполнения. Обеспечение идентичности прогонов требует дальнейшего развития средств отладки программ.

Многоконтактные фрагменты, такие, как схемы управления или операции, имеющие несколько операндов и дающие более одного результата, на первый взгляд, противоречат принципу строгого результата. Однако возможность перехода от строгого результата позволяет строить многопоточные функции, которые принимают параметры от ряда одноранговых потоков и генерируют ряд результатов в терминах числа потоков. Таким образом, можно, как и в функциональном языке параллельного программирования Sisal, переключиться на операции, которые отображают один ряд операндов в другой ряд результатов. Это может соответствовать структуре некоторых особо эффективных аппаратных узлов и, таким образом, позволяет представлять более эффективные решения [20].

4. Повышение производительности. При переходе к производству многократно применяемых программ и реализации особо важных параллельных вычислений резко меняются предпочтения. *Успех приложения и производительность программ становятся важнее, чем их формальная корректность и эффективность.* Чистое функциональное программирование можно рассматривать как метод функционального моделирования при создании прототипов программ для решения сложных проблем. Более широкая парадигма производственного функционального программирования позволяет переходить от таких функциональных моделей и прототипов к более эффективным структурам данных, наследуя ранее достигнутую правильность и принимая практические решения и компромиссы по обработке данных в зависимости от реальных условий, если это действительно необходимо.

В дополнение к принципам и следствиям в реальных языках программирования и системах производственное функциональное программирование обычно включает компромиссные механизмы, внешне в языке программирования выглядящие как специальные функции. Например, Lisp 1.5, Clisp, Cmucl и другие члены семейства Lisp обычно предоставляют следующие функции практического компромисса:

- **контроль типов данных** смягчает принцип *универсальности* функциями статического и динамического анализа типов данных и рангов функций;
- **схемы циклов**, моделирующие несколько расширенное множество знакомых механизмов управления вычислениями, позволяют преодолеть типичные опасения по поводу сложности реализации принципа *само-применения* на базе рекурсии;
- **возможность восстановления данных** при использовании деструктивных функций сопровождается безопасными аналогами, позволяющими исключить чрезмерное потребление памяти, частично противодействуя принципу *неизменности данных*;

- **программируемые прогнозы** объема памяти и скорости выполнения позволяют, по мере необходимости, императивно выполнять в программе распределение памяти для нейтрализации грубоватых механизмов принципа *гибкости ограничений*;
- **псевдо-функции**, помимо генерации результата функции, выполняют нагрузку в виде воздействия на внешнюю память или взаимодействия с устройствами, включая операции ввода-вывода и файловые операции, что несколько влияет на принципы *равноправия параметров и неизменяемости данных*;
- **мемоизация** позволяет радикально снижать сложность многократно повторяемых вычислений, делая понятную уступку успешному опыту против принципов *строгого результата и верификации* [21].

Общее взаимодействие принципов, следствий, приложений и практических компромиссов в системах функционального программирования представлено в табл. 2.

Таблица 2. Ключевые взаимосвязи принципов и механизмов ФП

	Семантика	Прагматика
Принципы	универсальность, само-применение, равноправие параметров	гибкость ограничений, неизменяемость данных, строгость результата
Следствия	мета-программирование, верификация, факторизация	непрерывность процессов, обратимость действий, унарные функции
Приложения к параллельным вычислениям	оттеснение мало вероятных вычислений, балансировка нагрузки, пространства итераций	автоматическое распараллеливание, идентичность повторных прогонов, многоконтактные фрагменты
Практичный компромисс	контроль типа данных, схемы циклов, восстановление данных	программируемый учет прогнозов, псевдо-функции доступа к устройствам, мемоизация

В рамках функционального программирования возможен подход, позволяющий учитывать особенности решаемых задач и парадигм программирования, необходимых для решения проблем параллельных вычислений, влияющих на выбор методов их решения, в зависимости от приоритетов в выборе языковых средств и реализационных конструкций. Можно наметить линию, позволяющую сравнивать языки, выделяя сопоставимые примеры программ, и особенности базовых средств и реализационных решений в системах программирования, нацеленных на улучшения создаваемых программных продуктов. Программирование давно уже стало массовой профессией, требующей объективных метрик для оценки качества программируемых решений. Парадигмальные ошибки, обнаруживаемые при эксплуатации привычных систем программирования на современном многопроцессорном и сетевом оборудовании, показывают, что часть из них были просто незаметны до появления сетей, мобильных устройств и суперкомпьютеров.

В целом следует отметить, что следствия из семантических и прагматических принципов функционального программирования и высокая моделирующая сила аппарата функций, расширенные специальными функциями практических компромиссов, позволяют полезно дополнять основные парадигмы параллельных вычислений и поддерживать работы по повышению производительности программ. В этом плане существенный вклад дают мемоизация, отложенные и опережающие вычисления, возможность синтаксического

конструирования и декомпозиции программ на автономно развиваемые модули, а на более общем уровне – метапрограммирование, позволяющее строить специализированные проблемно ориентированные DSL-языки. Синтаксическое конструирование – механизм надёжного использования прототипов и структур обрабатываемых данных. Декомпозиция программ приводит к выводу автономно развиваемых модулей, изменение которых не требует повторного программирования смежных модулей. Метапрограммирование может быть применено для создания подязыков, учитывающих индивидуальные особенности процессоров, а также для обустройства переходов от одного языка к другому.

Заключение. В феврале 2021 года состоялась 22-я конференция, посвященная современным тенденциям функционального программирования [7]. Представленные доклады убедительно показали нацеленность функционального программирования на решение основных проблем организации параллельных вычислений.

Многие вопросы пока не получили практического ответа. Появление новых парадигм можно связать с проявлением круга новых задач, решение которых вызывает трудности в рамках привычных парадигм. Не ясно, насколько целесообразно взаимодействие парадигм. Остаются в стороне образовательные проблемы овладения новыми парадигмами. Кроме того, особенности парадигм лишь отчасти выражаются на уровне представления программы, часть являются требованиями к прагматике системно-реализационной поддержки в ЯиСП. Граница между семантикой программируемых решений и прагматикой их системной поддержки у каждой парадигмы своя.

Похоже, путь к решению проблем параллельного программирования лежит через создание многоуровневого мультипарадигмального языка для поддержки основных парадигм параллельного программирования, включая средства низкого и сверхвысокого уровней, дополненных метапарадигмой функционального программирования. Собственно функциональное программирование выполняет роль проектно-конструкторского бюро для производственного программирования в рамках любой парадигмы и любого языка программирования.

СПИСОК ИСТОЧНИКОВ

1. McCarthy J. LISP 1.5 Programming Manual. The MIT Press, Cambridge, 1963, 106 p.
2. Backus J. Can programming be liberated from the von Neumann style? A functional stile and its algebra of programs. Commun. ACM 21, 8, 1978, pp. 613-641.
3. Gorodnyaya L. Method of paradigmatic analysis of programming languages and systems. CEUR Workshop Proceedings 2543 (2020), pp. 149-158.
4. Gorodnyaya L. On the presentation of the results of the analysis of programming languages and systems. CEUR Workshop Proceedings 2260 (2018), pp. 152–166.
5. Воеводин В.В., Воеводин Вл.В. Параллельные вычисления / В.В., Воеводин, Вл.В Воеводин. – СПб.: БХВ-Петербург, 2002. – 608 с.
6. Городняя Л.В. О неявной мультипарадигмальности параллельного программирования / Л.В. Городняя //Научный сервис в сети Интернет: труды XXIII Всероссийской научной конференции (20-23 сентября 2021 г., онлайн). – М.: ИПМ им. М.В.Келдыша, 2021. – С.104-116.
7. Pieter Koopman, Steffen Michels, and Rinus Plasmeijer Dynamic Editors for Well-Typed Expressions. Trends in Functional programming/ 22nd internationa Symposium, TFP 2021, February 17-19, 2021 Springer, LNCS 12834, pp. 44-66.

8. Schwartz Jacob T. "Set Theory as a Language for Program Specification and Programming". Courant Institute of Mathematical Sciences, New York University, 1970⁴.
9. Душкин Р. Функциональное программирование на языке Haskell / Р. Душкин. – ДМК-Пресс, 2016. – 606 с. – URL: <https://www.labirint.ru/books/497084/>
10. Котов В.Е. МАРС: архитектуры и языки для реализации параллелизма / Котов В.Е. // Системная информатика. Вып 1. Проблемы современного программирования. – Новосибирск: Наука. Сиб. отд-ние, 1991. – С.174-194.
11. Городняя Л.В. Систематизации парадигм программирования по приоритетам принятия решений / Л.В. Городняя // Электронные библиотеки, 2020. – Т. 23 ,ч. 2. – № 4. – С. 666-696. – DOI:10.26907/1562-5419-2020-23-4-666-696
12. Лавров С.С. Функциональное программирование /С.С. Лавров// Компьютерные инструменты в образовании, 2002. – №2-4.
13. Лавров С.С., Городняя Л.В. Функциональное программирование. Интерпретатор языка Лисп / С.С.Лавров, Л.В. Городняя //Компьютерные инструменты в образовании. – СПб, 2002. – № 5.
14. Мальцев А.И. Алгоритмы и рекурсивные функции / Мальцев А.И. – М.: Наука, 1965 г. – 392 с.
15. Бурдонов И.Б., Косачев. А.С. Семантики взаимодействия с отказами, дивергенцией и разрушением. Часть 2. Условия конечного полного тестирования / И.Б. Бурдонов, А.С. Косачев //Вестник Томского Государственного Университета, 2011. – № 2(15).
16. Фихтенгольц Г.М. Основы математического анализа. Том 1 / Г.М. Фихтенгольц. – М: «Наука», 1968. – 440 с.
17. Городняя Л.В., Кирпотина И.А. Методика декомпозиции сложных информационных систем / Л.В. Городняя, И.А Кирпотина И.А // 30-я Международная конференция "Информационные технологии в образовании 2019" (ИТО-Троицк-Москва-2019), 2019 г.
18. Сошников Д. В. Функциональное программирование на языке F# / Д. В. Сошников. –М.: ДМК Пресс, 2011.
19. Ластовецкий А.Л. Предварительное сообщение о языке программирования mPC / А.Л. Ластовецкий // Вопросы кибернетики: Приложения системного программирования. Научный совет по комплексной проблеме "Кибернетика" РАН. Москва, 1995. – С. 20-39.
20. Касьянов В.Н., Гордеев Д.С., Золотухин Т.А. Система облачного параллельного программирования CPPS: визуализация и верификация Cloud Sisal программ / В.Н. Касьянов, Д.С. Гордеев, Т.А. Золотухин // Конструирование и оптимизация программ. – Новосибирск: НГУ, 2020.
21. Файфель Б.Л. Мемоизация в Лиспе / Б.Л. Файфель. – URL: <https://habr.com/ru/post/576278/>

Городняя Лидия Васильевна, к.ф.-м.н., доцент, с.н.с., Институт систем информатики СО РАН, gorod@iis.nsk.su, Россия, г. Новосибирск, проспект Академика Лаврентьева, 6

⁴ Данфорд Н., Шварц Дж. Линейные операторы. Том 2. Спектральная теория. Самосопряженные операторы в гильбертовом пространстве. М.: Мир, 1966

On functional programming in the organization of parallel computing

Lidia V. Gorodnyaya

A.P. Ershov Institute of Informatics Systems, Siberian Branch of the Russian Academy of Sciences, Russia, Novosibirsk, lidvas@gmail.com

Abstract. The article is devoted to the results of the analysis of modern trends in functional programming, considered as a meta-paradigm for solving the problems of organizing parallel computing and multi-threaded programs for multi-processor complexes and distributed systems. Taking into account the multi-paradigm nature of parallel programming languages, here we used the results of paradigm analysis, which made it possible to find out a number of features and predict the course of the processes of using programs, as well as their study and development. There is reason to believe that functional programming can improve the performance of programs by preparing their prototypes in advance. A description of the semantic (universality, self-applicability, equality of parameters) and pragmatic (flexibility of constraints, immutability of data, rigor of the result) principles of functional programming is given, as well as the consequences of these principles aimed at reducing labor intensity and improving the technology of prototyping and debugging programs (constructiveness, verification, factorization, continuity of processes, reversibility of actions). The role of paradigmatic decomposition of programs in the technology of developing long-lived programs, which are often parallel computation programs, is noted. The perspective of functional programming as an auxiliary universal meta-paradigm for solving complex and important problems, burdened with difficult to verify and poorly compatible requirements of efficiency, reliability, safety and improvement opportunity, is especially emphasized. A variety of paradigmatic characteristics inherent in the preparation and debugging of long-lived parallel computing programs are shown.

Keywords: functional programming, parallel computing, programming languages, programming paradigms, multiparadigm

REFERENCES

1. McCarthy J. LISP 1.5 Programming Manual. The MIT Press, Cambridge, 1963, 106 p.
2. Backus J. Can programming be liberated from the von Neumann style? A functional style and its algebra of programs. *Commun. ACM* 21, 8, 1978, pp. 613-641.
3. Gorodnyaya L. Method of paradigmatic analysis of programming languages and systems. *CEUR Workshop Proceedings* 2543 (2020), pp. 149-158.
4. Gorodnyaya L. On the presentation of the results of the analysis of programming languages and systems. *CEUR Workshop Proceedings* 2260 (2018), pp. 152–166.
5. Voyevodin V.V., Voyevodin V.I. *Parallel'nyye vychisleniya [Parallel computing]* / V.V. Voyevodin, V.I. Voyevodin – SPb.: BKHV-Peterburg = St. Petersburg: BHV-Petersburg, 2002. – 608 p.
6. Gorodnyaya L.V. O neyavnoy mul'tiparadigmal'nosti parallel'nogo programmirovaniya [On the implicit multi-paradigm of parallel programming] / L.V. Gorodnyaya // *Nauchnyy servis v seti Internet: trudy XXIII Vserossiyskoy nauchnoy konferentsii (20-23 sentyabrya 2021 g., onlayn)* = Scientific service on the Internet: Proceedings of the XXIII All-Russian Scientific

- Conference (September 20-23, 2021, online). – Moscow: IPM im. M.V. Keldysh, 2021, pp.104-116 p.
7. Pieter Koopman, Steffen Michels, and Rinus Plasmeijer Dynamic Editors for Well-Typed Expressions. Trends in Functional programming/ 22nd international Symposium, TFP 2021, February 17-19, 2021/ Springer, LNCS 12834, pp. 44-66.
 8. Schwartz Jacob T. "Set Theory as a Language for Program Specification and Programming". Courant Institute of Mathematical Sciences, New York University, 1970⁵.
 9. Dushkin R. Funktsional'noye programmirovaniye na yazyke Haske [Functional programming in Haskell] / R. Dushkin. – DMK-Press, 2016, – 606 p.
 10. Kotov V.Ye. MARS: arkhitektury i yazyki dlya realizatsii parallelizma [MARS: architectures and languages for implementing parallelism] / V.Ye. Kotov // Sistemnaya informatika. Vyp 1. Problemy sovremennogo programmirovaniya = System informatics. Issue 1. Problems of modern programming. – Novosibirsk: Nauka, Sib. otd-niye= Science. Sib. Dept, 1991, – pp.174-194.
 11. Gorodnyaya L.V. Sistematizatsii paradigm programmirovaniya po prioritetam prinyatiya resheniy [Systematization of programming paradigms according to decision-making priorities] / L.V. Gorodnyaya /Elektronnyye biblioteki = Digital Libraries, Volume 23, № 4 (2020), Part 2, – Pp. 666-696. DOI:10.26907/1562-5419-2020-23-4-666-696.
 12. Lavrov S.S. Funktsional'noye programmirovaniye [Functional programming] / S.S. Lavrov // Komp'yuternyye instrumenty v obrazovanii = Computer tools in education, 2002, № 2-4.
 13. Lavrov S.S., Gorodnyaya L.V. Funktsional'noye programmirovaniye. Interpretator yazyka Lisp [Functional programming. Lisp language interpreter] / S.S. Lavrov, L.V. Gorodnyaya // Komp'yuternyye instrumenty v obrazovanii = Computer tools in education, – S-Pb, 2002, № 5.
 14. Mal'tsev A.I. Algoritmy i rekursivnyye funktsii [Algorithms and recursive functions] / A.I. Mal'tsev. – M.: Nauka = Science, 1965, 392 p.
 15. Burdonov I.B., Kosachev A.S. Semantiki vzaimodeystviya s otkazami, divergentsiyey i razrusheniyem. Chast' 2. Usloviya konechnogo polnogo testirovaniya [Semantics of interaction with failures, divergence and destruction. Part 2. Conditions for final full testing] / I.B. Burdonov, A.S. Kosachev. Vestnik Tomskogo Gosudarstvennogo Universiteta = Bulletin of Tomsk State University, № 2(15), 2011.
 16. Fikhtengol'ts G.M. Osnovy matematicheskogo analiza [Fundamentals of mathematical analysis] / G.M. Fikhtengol'ts. – M: «Nauka»= "Science", Tom 1, 1968, – 440 p.
 17. Gorodnyaya L.V., Kirpotina I.A. Metodika dekompozitsii slozhnykh informatsionnykh system [Decomposition methodology for complex information systems] / L.V. Gorodnyaya, I.A. I.A. // 30-ya Mezhdunarodnaya konferentsiya "Informatsionnyye tekhnologii v obrazovanii 2019" (ITO-Troitsk-Moskva-2019). 25 iyunya 2019 g. = 30th International Conference "Information Technologies in Education 2019" (ITO-Troitsk-Moscow-2019), June 25, 2019.
 18. Soshnikov D.V. Funktsional'noye programmirovaniye na yazyke F# [Functional programming in F#] / D.V. Soshnikov. – Moscow: DMK Press, 2011.
 19. Lastovetskiy A.L. Predvaritel'noye soobshcheniye o yazyke programmirovaniya mpC [Preliminary report on the mpC programming language] / A.L. Lastovetskiy // Voprosy kibernetiki: Prilozheniya sistemnogo programmirovaniya. Nauchnyy sovet po kompleksnoy probleme "Kibernetika" RAN = Issues of Cybernetics: Applications of System Programming.

⁵ Dunford N., Schwartz J. Linear Operators. Volume 2. Spectral theory. Self-adjoint operators in a Hilbert space. Moscow: Mir, 1966 (in Russian).

- Scientific Council on the complex problem "Cybernetics" RAS. – Moscow, 1995, – pp. 20-39.
20. Kas'yanov V.N., Gordeyev D.S., Zolotukhin T.A. Sistema oblachnogo parallel'nogo programirovaniya CPPS: vizualizatsiya i verifikatsiya Cloud Sisal program [Cloud parallel programming system CPPS: visualization and verification of Cloud Sisal programs] / V.N.Kas'yanov, D.S. Gordeyev, T.A. Zolotukhin // Konstruirovaniye i optimizatsiya program = Design and optimization of programs. Novosibirsk: NGU = – Novosibirsk State University, 2020.
21. Fayfel' Memoizatsiya v Lispe [Memoization in Lisp] / B.L. Fayfel' – URL: <https://habr.com/ru/post/576278/>

Lidia V. Gorodnyaya, candidate of Physical and Mathematical Sciences, Associate Professor, Senior Researcher, A.P. Ershov Institute of Informatics Systems, Siberian Branch of the Russian Academy of Sciences, lidvas@gmail.com, Russia, Novosibirsk, Akademika Lavrentiev Avenue, 6

Статья поступила в редакцию 21.12.2021; одобрена после рецензирования 20.01.2022; принята к публикации 25.01.2022.

The article was submitted 12.21.2021; approved after reviewing 01.20.2022; accepted for publication 01.25.2022.

О моделировании железнодорожных узлов на основе теории массового обслуживания

Жарков Максим Леонидович¹, Супруновский Антон Викторович²

¹ Институт динамики систем и теории управления имени В.М. Матросова СО РАН,
Россия, Иркутск, zharkm@mail.ru

² Иркутский государственный университет путей сообщения,
Россия, Иркутск, as.irgups@gmail.com

Аннотация. В работе предложена методика математического моделирования движения поездов по железнодорожным узлам, основанная на применении сетей массового обслуживания специального вида. Моделирование железнодорожных узлов является актуальной задачей из-за их важности для обеспечения грузо- и пассажироперевозок, а также межрегионального экономического взаимодействия. Применение сетей массового обслуживания позволяет учесть влияние случайных факторов на работу узлов и отобразить в модели структурные особенности конкретного объекта. Для апробации методики выбран один из крупнейших узлов Транссибирской магистрали. В статье построена модель его работы в виде сети массового обслуживания с двумя входящими потоками заявок, выполнено ее численное исследование, по результатам которого определены максимальная допустимая нагрузка и «узкие места» в структуре транспортно-портной системы, и выработаны рекомендации по увеличению ее пропускной способности.

Ключевые слова: железнодорожный узел, транспортный поток, математическая модель, теория массового обслуживания, вычислительный эксперимент

Цитирование: Жарков М.Л., Супруновский А.В. О моделировании железнодорожных узлов на основе теории массового обслуживания // Информационные и математические технологии в науке и управлении. – 2022. – № 1 (25). – С. 120-132. – DOI:10.38028/ESI.2022.25.1.010.

Введение. Под железнодорожным узлом понимается совокупность нескольких взаимодействующих между собой железнодорожных станций на относительно небольшой территории [1]. При их исследовании обычно применяют методы математического моделирования, так как а) они обладают высокой сложностью из-за большого числа составных элементов и связей между ними, которые необходимо учитывать; б) проведение натурных экспериментов связано с большими трудностями, либо попросту невозможно.

Наиболее распространенными математическими моделями на железнодорожном транспорте и в транспортной логистике являются оптимизационные [2, 3], которые применяются как на грузовом, так и на пассажирском [4] транспорте. Используются также и стохастические модели [5], которые позволяют учесть влияние случайных факторов. Особенно существенное влияние они оказывают на входящий транспортный поток [6]. Обычно из-за сложности железнодорожных транспортных систем их математические модели приходится исследовать численно, в частности, с применением методов имитационного моделирования [7–9].

В настоящее время для исследования железнодорожных узлов чаще всего применяются система имитационного моделирования ИМЕТРА [7,8] и различные библиотеки AnyLogic, в частности, «железнодорожные перевозки» или Enterprise Library [9]. При очевидном удобстве, использование данных инструментов сопряжено и с определенными сложностями. Во-

первых, они не обладают достаточной гибкостью. В частности, в структуре железнодорожного узла возможно присутствие нетиповых элементов: специализированных станций или станций с особенностями структуры. Во-вторых, нам не удалось найти исчерпывающего описания математического аппарата, поэтому остается открытым вопрос о точности моделирования. Так, в [9], по-видимому, используются модели теории массового обслуживания, однако адекватность описания транспортных потоков, на наш взгляд, недостаточна.

Ранее с участием авторов был разработан новый подход к моделированию работы сортировочных и пассажирских железнодорожных станций с использованием сетей массового обслуживания (СеМО) с входящим ВМАР-поток [10, 11]. Для анализа получаемых моделей используется имитационное моделирование. Это позволило избежать проблем, описанных выше. В данной работе указанный подход адаптирован для более крупных систем – железнодорожных узлов, что потребовало его значительного усовершенствования. Во-первых, мы используем несколько независимых МАР-потоков (частный случай ВМАР-потока) для описания поступления поездов с различных направлений. Это позволяет учесть параметры поездопотоков, зависящие от категории поездов и направления их движения. Во-вторых, работа станций и участков железной дороги между ними моделируется одним или более узлами СеМО. Благодаря этому в модели удастся отобразить при минимальных трудозатратах: а) сложную иерархическую структуру системы, в которой отдельные категории поездов имеют разные маршруты; б) различные параметры и особенности работы станций, а также пропускную способность участков железной дороги.

Объектом для апробации подхода выбран Иркутский железнодорожный узел (далее – ИЖУ) – один из крупнейших на Транссибирской магистрали.

1. Объект исследования. ИЖУ имеет протяженность порядка 60 км и включает восемь станций: Суховская, Мегет, Батарейная, Иркутск-Сортировочный, Военный городок, Иркутск-Пассажирский, Кая и Гончарово. Помимо главного хода Транссибирской магистрали, имеется и обходной путь от станции Иркутск-Сортировочный до станции Гончарово, который предназначен для пропуска транзитных поездов в обход станции Иркутск-Пассажирский. Схема Иркутского узла представлена на рис. 1.

Рассмотрим станции ИЖУ:

- Суховская является крупной грузовой станцией, в составе которой имеется парк приема на 6 путей, сортировочный парк на 18 путей с сортировочной горкой, и грузовой двор, в котором могут загружаться до четырех поездов за сутки. Также в ее структуру включены два приемоотправочных парка (ПОП) по 7 и 13 путей для отправки обработанных на этой станции поездов и остановки транзитного сообщения без переработки.

Остальные грузовые станции имеют меньший размер. Они состоят из одного ПОП и нескольких путей необщего пользования, где ведется грузовая работа (далее – грузовой фронт). С учетом главного хода:

- станция Батарейная имеет – 9 путей;
- станция Гончарово также – 9 путей;
- станция Военный городок – 4 пути;

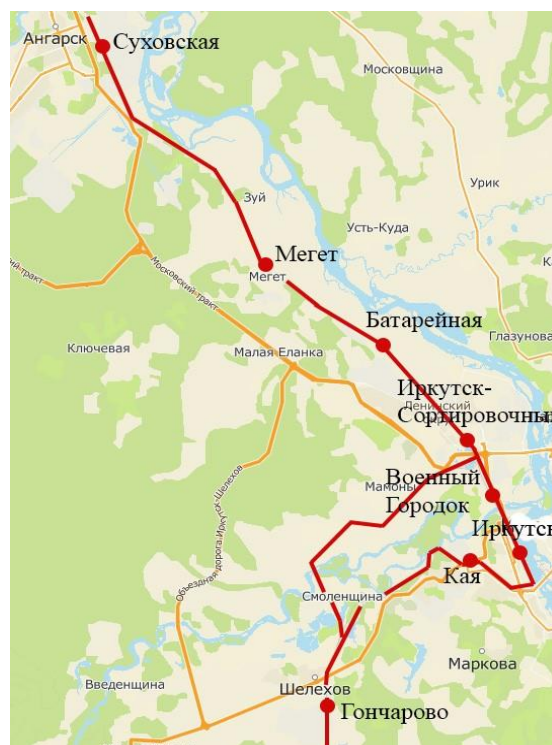


Рис. 1. Схема Иркутского узла

- станция Кая – 10 путей;
- станция Мегет имеет один парк на 9 путей, включая главный ход; она предназначена для пропуска или кратковременной остановки поездов;
- Иркутск-Сортировочный состоит из двух систем, каждая из которых включает три последовательно расположенных парка: парки приема и парки отправления имеют по 8 путей, четный сортировочный парк – 24 пути, а нечетный – 18; в этих системах работают по 4 бригады отправления, обрабатывающие как местные, так и транзитные поезда;
- станция Иркутск-Пассажирский включает две части: пассажирская состоит из 8 путей: 2 – главный ход, 5 путей с платформами, на которых ведутся работы с пассажирскими поездами, и один технический путь без платформы; грузовая часть – грузовой фронт, на котором можно разместить до четырех поездов; грузовой поездопоток в основной массе проходит транзитом по станции.

Участки железной дороги между станциями (далее – участки) имеют различное число путей: с запада до Суховской, станции Мегет и Батарейной – 3; Батарейная – Иркутск-Сортировочный – 4; Иркутск-Сортировочный – Гончарово – 2, от Гончарово и на восток – 3, на обходном пути – 2.

По плану среднесуточный поездопоток, который проходит через Иркутский железнодорожный узел, в восточном направлении составляет 76 транзитных, 25 местных и 18 пассажирских поездов, в западном – 71, 30 и 18 поездов соответственно. Также курсируют 14 электричек, которые имеют минимальные время стоянки на станциях и значение занятости пути. Поэтому далее данная категория поездов не рассматривается при моделировании. Среднесуточное количество формируемых грузовых поездов по станциям: Суховская – 4 поезда; Батарейная и Иркутск-Пассажирский – по 3; Военный городок и Кая – по 0,5; Гончарово – 1; Иркутск-Сортировочный – по 25 в западной и восточной системах.

2. Методика моделирования работы железнодорожных узлов. При построении модели мы выделяем две группы элементов, для описания которых необходимо использовать различный математический аппарат: входящие поездопоток и инфраструктурные элементы узла. Они моделируются в три этапа: 1) описание входящих поездопотоков; 2) моделирование работы инфраструктурных элементов; 3) учет в модели маршрутов движения поездов в системе и других ее особенностей.

Математическое описание входящий поездопотоков. Как правило, поезда поступают с двух и более направлений. При этом имеется несколько категорий поездов с различной интенсивностью поступления составов. Для моделирования входящего поездопотока с каждого направления используется отдельный МАР-поток [10–12]. Под заявкой понимается весь поезд целиком.

Математическое описание функционирования элементов железнодорожных узлов. К элементам железнодорожного узла относятся участки между станциями и, собственно, сами станции. В них выполняются, соответственно, два класса операций над поездами: а) пропуск по путям; б) обработка на станциях. Операции обоих классов являются однотипными и регулярно повторяются, а их продолжительность подвержена воздействию случайных факторов, что позволяет использовать модельный аппарат ТМО. Каждый элемент железнодорожных узлов, в зависимости от типа, описывается одной или более системами массового обслуживания (далее СМО или узлы). В каждой СМО каналы моделируют работу путей главного хода на участке или обслуживающих устройств на станции. Длина очереди определяется количеством поездов, которые могут остановиться на участке или станции в ожидании обслужи-

вания. Распределение времени обслуживания в каналах СМО подбирается в соответствии с пропускной способностью рассматриваемого участка или станции.

Описание маршрутов движения поездов и других его особенностей. Участки и станции образуют нелинейную иерархическую структуру, в которой может быть несколько маршрутов движения поездов. При этом поезд вынужден остановиться и ожидать освобождения участка на пути его следования или места на станции. Это означает, что между структурными элементами железнодорожного узла имеются обратные связи. Математическое описание маршрутов движения строится в виде маршрутной матрицы. Ее элементы – вероятности переходов между узлами СеМО [10] – определяются, как относительная частота (отношение числа поездов, поступающих в конкретный узел, к общему числу поездов для одного направления). Обратная связь задается в виде временной блокировки каналов предыдущей СМО до тех пор, пока в следующей не освободится достаточно места для принятия заявок. Построенные на предыдущих шагах МАР-поток и СМО, а также маршрутная матрица составляют СеМО.

Исследование СеМО из-за большого количества узлов и их взаимосвязей, как правило, выполняется с помощью численных методов, наиболее распространенным из которых является имитационное моделирование [10, 11]. С его помощью возможно определить показатели эффективности СеМО: вероятность отказа, абсолютную пропускную способность, время пребывания заявки в системе и отдельно в узле, число занятых каналов и длину очереди, время блокировки каналов в узлах. На их основе затем делается вывод об эффективности работы выбранного объекта и определяются «узкие места» в его структуре.

3. Построение модели работы Иркутского железнодорожного узла. *Первый этап* – построение модели общего поездопотока. По плану в сутки через ИЖУ проходит 119 поездов трех категорий в восточном и столько же в западном направлениях. Поездопоток каждого из них мы описываем отдельным МАР-поток: *МАР-В* – восточное и *МАР-З* – западное. Суммарная интенсивность поступления заявок составляет $\lambda_B = \lambda_3 = 4,96$ в час для каждого потока (направления). В качестве вероятностей поступления заявок из отдельного подпотока берутся относительные частоты. Тогда матрицы МАР-потоков будут иметь вид [11, 12]:

$$\begin{aligned} D_{B0} &= \begin{pmatrix} -4,96 & 0 & 0 \\ 0 & -4,96 & 0 \\ 0 & 0 & -4,96 \end{pmatrix}, \quad D_{B1} = \begin{pmatrix} 3,17 & 1,04 & 0,75 \\ 3,17 & 1,04 & 0,75 \\ 3,17 & 1,04 & 0,75 \end{pmatrix}, \\ D_{30} &= \begin{pmatrix} -4,96 & 0 & 0 \\ 0 & -4,96 & 0 \\ 0 & 0 & -4,96 \end{pmatrix}, \quad D_{31} = \begin{pmatrix} 2,96 & 1,25 & 0,75 \\ 2,96 & 1,25 & 0,75 \\ 2,96 & 1,25 & 0,75 \end{pmatrix}. \end{aligned} \quad (1)$$

Второй этап – описание структурных элементов ИЖУ в виде СМО. Сперва моделируется функционирование главного хода, затем – станций.

На ИЖУ мы выделяем восемь участков (по числу станций) и обходной путь. Последний имеет достаточную протяженность, чтобы одновременно пропускать три поезда в одном направлении. Поэтому условно разделим его на три участка. По этим 11 участкам поезда следуют в двух направлениях, их работу моделируем двумя СМО: узлы 1–11 описывают движение поездов в восточном направлении, узлы 12–22 – в западном. узлы 5–10, 13, 14, и 18 моделируют работу обходного пути.

При описании участка в виде СМО путь, по которому движутся поезда, считается каналом, а число путей в приемоотправочных парках на станции этого участка определяет максимальную вместимость очереди (узлы 1, 2, 11, 12 и 20–22). Если приемоотправочного парка нет на станции, или поезда могут произвести остановку в нем только для дальнейшей пере-

работки на этой же станции, то в СМО, которые описывают выбранный участок, не будет очереди (узлы 3–10 и 13–19). Каждому узлу отводится половина доступных путей главного хода. Время прохода поезда по участку было получено на основе натурных наблюдений и составляет 6–10 минут. Это время мы считаем случайной величиной, распределенной по нормальному закону [10]. Тогда время обслуживания поездов в узлах 3 и 4–17 описывается нормальным распределением $N(8; 0,67)$, где первый параметр – математическое ожидание – равно среднему времени прохода поезда по участку $6 + (10 - 6) / 2 = 8$ (в мин.), второй – ср. квадратическое отклонение, который подобран по «правилу трех сигм» и равен $(10 - 8) / 3 = 0,67$ мин. В случае если путей на участке нечетное количество, то один из них не учитывается, а его пропускная способность поровну добавляется к пропускной способности остальных путей (узлы 1, 2, 11 и 20–22). Следовательно, время обслуживания поездов в узлах 1, 2, 11 и 18–22 подчиняется $N(5,2; 0,67)$, т.е. математическое ожидание уменьшено на 35%.

Таким образом работу главного хода ИЖУ моделируем с помощью 22-х узлов. В терминах ТМО они будут иметь следующий вид [12]: узел 1 – *МАР-В*/G/1/12; узлы 2 и 21 – */G/1/3; узлы 3, 4 и 5 – */G/2/0; узлы 6–10 – */G/1/0; узел 11 – */G/1/3; узел 12 – *МАР-3*/G/2/4; узлы 13–17 – */G/1/0; узлы 18 и 19 – */G/2/0; узел 20 – */G/1/1; узел 22 – */G/1/6.

Работа станций, в зависимости от типа, описывается в виде одной или нескольких СМО. Грузовые станции Суховская, Батарейная, Военный городок, Кая, Гончарово и грузовой фронт на станции Иркутск-Пассажирский моделируются с помощью одноканальных СМО (узлы 23, 24, 27, и 29–31). Модель сортировочной станции Иркутск-Сортировочный строится в виде двух четырехканальных СМО, которые описывают работу ее восточной и западной систем (узлы 25 и 26). Максимальная длина очереди в СМО определяется суммарным числом путей на соответствующих станциях (см. раздел «Объект исследования»). Пассажирская станция Иркутск-Пассажирский имеет 8 путей, из которых мы не учитываем два пути главного хода и технический путь. Тогда модель ее работы будет иметь вид пятиканальной СМО без очереди (узел 28). В терминах ТМО модели описанных станций имеют следующий вид: узел 23 – */G/1/24; узел 24 – */G/1/7; узел 25 – */G/4/38; узел 26 – */G/4/40; узел 27 – */G/1/6; узел 28 – */G/5/0; узел 29 – */G/1/3; узел 30 – */G/1/8; узел 31 – */G/1/6. Распределения времени обслуживания заявок в каналах следующие, все параметры указаны в минутах: узел 23 – $N(288; 20)$; узел 24 – $N(360; 24)$; узлы 25 и 26 – $N(91; 15)$; узлы 27 и 30 – $N(720; 60)$; узел 28 – $N(40; 6)$; узел 29 – $N(206; 15)$; узел 31 – $N(480; 30)$.

Выше был описан 31 узел. К ним необходимо добавить еще два фиктивных, которые выступают в качестве источников потоков заявок [10, 11]: узел 0 для *МАР-В* и узел 32 для *МАР-3*. Таким образом, модель работы ИЖУ включает 33 узла и два *МАР*-потока. Схема полученной СеМО представлена на рис. 2. На нем круги обозначают узлы, которые описывают работу участков, а эллипсы – работу станций, жирными стрелками отмечен главный ход и обходной путь.

Третий этап – описание маршрутов движения поездов строится в виде маршрутной матрицы. Она будет разреженной, из-за того, что поезда могут передвигаться между станциями только по главному ходу, и, следовательно, неинформативной. Поэтому маршрутную матрицу мы не приводим, а вероятности переходов между узлами представлены на рис. 2 в виде весов.

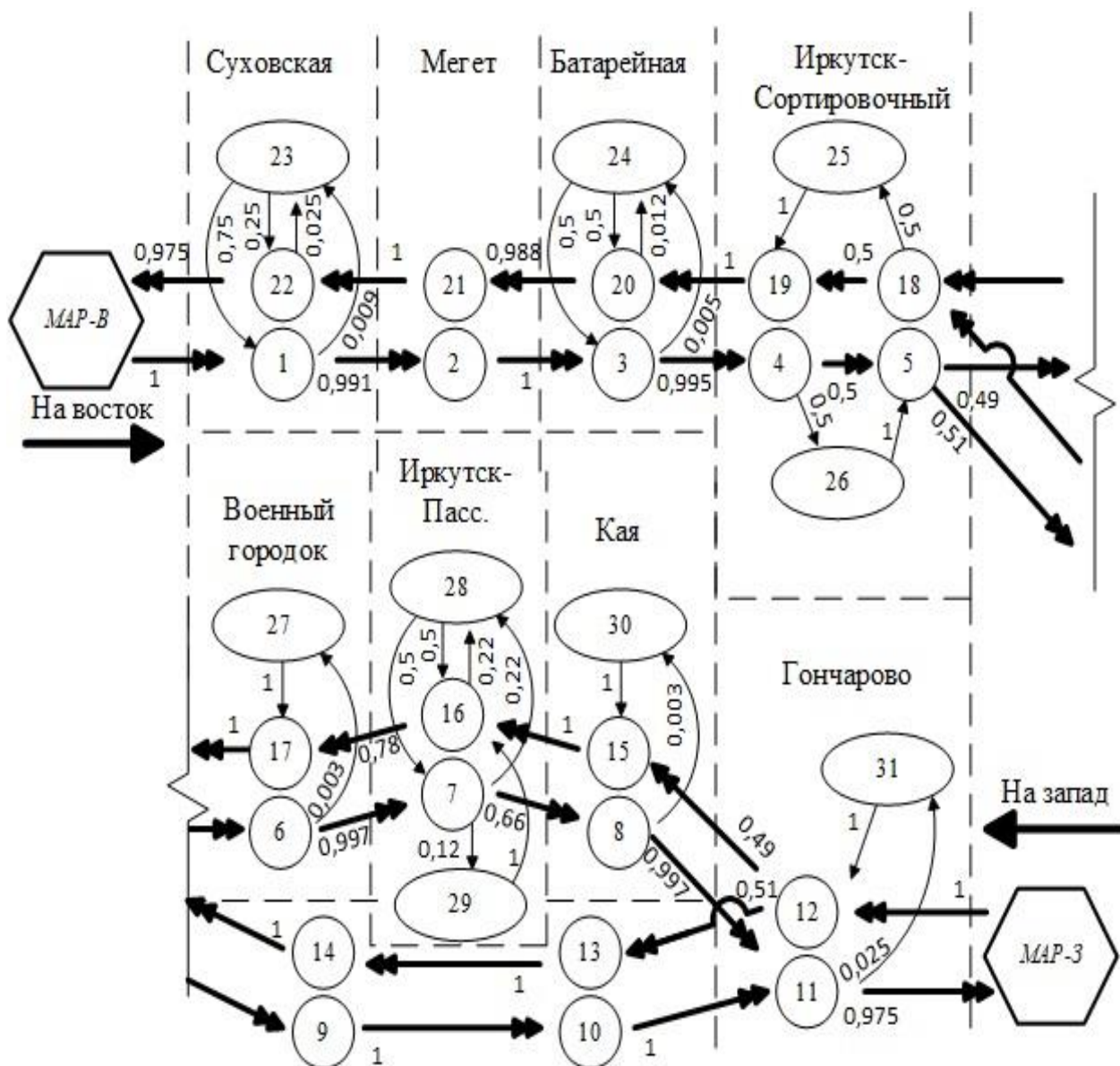


Рис. 2. Схема СеМО, описывающей работу Иркутского железнодорожного узла

Далее приведено исследование полученной СеМО с помощью имитационной модели, которая позволяет найти показатели эффективности. Она входит в состав интеллектуальной системы [13], краткое описание которой представлено в конце статьи.

4. Вычислительный эксперимент. Далее в табл. 1 и 2 представлены результаты двух экспериментов, в которых моделируется работа вышеописанной СеМО при разных параметрах. В каждой таблице приведены средние результаты за десять пусков имитационной модели. Виртуальное время моделирования для каждого пуска составило 30 дней. Это минимальное значение, за которое имитационная модель может получить показатели эффективности данной СеМО.

Эксперимент 1. В табл. 1 представлены результаты моделирования при текущем поездопотоке, т.е. 119 поездов в сутки или $\lambda_{\text{в}} = \lambda_{\text{з}} = 4,96$ поездов в час для каждого направления. Они описываются матрицами (1).

Таблица 1. Результаты эксперимента 1

Общие показатели эффективности СеМО								
Поступило заявок	7101,0	P_{loss}	0,0008	T_{B1}	2,21	T_{B2}	2,05	
Отклонено заявок	5,3	T_{lock}	8,35	T_{31}	1,66	T_{32}	1,49	
Показатели эффективности отдельных узлов (время измеряется в минутах)								
	Узел 1	Узел 2	Узел 3	Узел 4	Узел 5	Узел 6	Узел 7	Узел 8
K	0,47	0,47	0,68	0,55	0,65	0,39	0,46	0,31
L	0,21	0,02	-	-	-	-	-	-
T_y	7,91	5,77	21,28	40,86	23,66	9,47	8,59	8,52
T_l	0	32,4	84,8	941,4	2746,4	1669,6	186,6	16,5
	Узел 9	Узел 10	Узел 11	Узел 12	Узел 13	Узел 14	Узел 15	Узел 16
K	0,37	0,37	0,44	0,75	0,38	0,38	0,40	0,49
L	-	-	0,07	0,08	-	-	-	-
T_y	8,55	8,51	6,39	9,98	8,57	8,54	9,85	8,58
T_l	66,6	33,0	9,0	3539,8	107,6	64,3	2360,8	166,0
	Узел 17	Узел 18	Узел 19	Узел 20	Узел 21	Узел 22	Узел 23	Узел 24
K	0,38	0,58	0,61	0,48	0,49	0,48	0,88	0,43
L	-	-	-	0,08	0,03	0,02	3,08	0,14
T_y	8,55	22,45	21,83	6,46	6,02	5,62	1295,00	498,74
T_l	83,5	869,8	234,2	0,0	1817,5	0	21,0	6,0
	Узел 25	Узел 26	Узел 27	Узел 28	Узел 29	Узел 30	Узел 31	
K	3,99	3,99	0,21	1,16	0,97	0,09	0,88	
L	32,73	34,46	0,04	-	2,00	-	1,97	
T_y	920,28	961,76	917,68	3270,82	664,82	1047,79	1538,68	
T_l	450,0	495,2	13,8	1768,8	497,8	23,0	9,0	

Здесь и далее: K – среднее число работающих каналов; L – средняя длина очереди; T_y – среднее время пребывания заявки в узле; T_l – продолжительность (в мин.) блокировки работы каналов за все время моделирования; P_{loss} – вероятность отказа; T_{lock} – среднее суммарное время (в часах) блокировки в течение суток каналов узлов с 1-го по 22-й; T_{B1} и T_{31} – среднее время (в часах) прохождения поезда по главному ходу ИЖУ в восточном (узлы 1–8 и 11) и западном (узлы 12, 15–22) направлениях; T_{B2} и T_{32} – среднее время (в часах) прохождения поезда по ИЖУ через обходной путь в восточном (узлы 1–5 и 9–11) и западном (узлы 12–14 и 18–22) направлениях.

Наиболее загруженными являются узлы 25 и 26 (станция Иркутск-Сортировочный), среднее число работающих каналов которых близко к максимальному значению (4 шт.), а средняя длина очереди превышает 85% от максимальной ее вместимости. Из-за этого наблюдается продолжительная блокировка работы каналов в узлах 4 и 18. Также и в узлах 4 и 15 имеется значительная блокировка, но уже из-за загрузки узлов 6–8 и 15–17 (главный ход). Это подтверждается параметрами T_{B1} и T_{31} , которые в среднем на 36% больше, чем T_{B2} и T_{32} . Переноса результаты на объект исследования получаем, что ИЖУ имеет достаточную пропускную способность. Среднее суммарное время «простоя» (T_{lock}) в сутки для всех поездов (238 шт.) составляет менее 9 часов, т.е. в среднем 2 минуты на поезд. Этот «простой» образуется в основном из-за большой загрузки станции Иркутск-Сортировочный, а также значительного местного поездопотока и пассажирского сообщения на участке при станции Иркутск-Пассажирский (главный ход).

Пропускную способность Иркутского железнодорожного узла теоретически можно увеличить. В настоящее время имеются планы ВСЖД филиала ОАО «РЖД» по увеличению транзитного поездопотока по ИЖУ на 15 %, что потребует перенаправить на 20% больше поездов этой категории на обходной путь.

Эксперимент 2. В табл. 2 представлены результаты моделирования при увеличении объемов транзитных поездопотоков на 15%, т.е. 87 и 82 поездов в сутки для восточного и западного направлений соответственно. В этом случае интенсивность поступления поездов с каждого направления составит $\lambda_B = \lambda_3 = 5,42$ в час. Также увеличим поездопоток, который проходит через обходной путь на 20%. Для этого в модели изменим вероятности перехода заявок между узлами (см. рис. 2): из узла 5 в узел 6 $p_{5,6} = 0,42$; из узла 5 в узел 9 $p_{5,9} = 0,58$; из узла 12 в узел 13 $p_{12,13} = 0,58$; из узла 12 в узел 15 $p_{12,15} = 0,42$.

Таблица 2. Результаты эксперимента 2

Общие показатели эффективности СеМО								
Поступило заявок	7758,5	P_{loss}	0,0003	T_{B1}	1,89	T_{B2}	1,73	
Отклонено заявок	2,3	T_{lock}	9,00	T_{31}	1,69	T_{32}	1,52	
Показатели эффективности отдельных узлов (время измеряется в минутах)								
	Узел 1	Узел 2	Узел 3	Узел 4	Узел 5	Узел 6	Узел 7	Узел 8
K	0,50	0,50	0,73	0,76	0,78	0,39	0,47	0,31
L	0,26	0,03	-	-	-	-	-	-
T_y	8,28	5,81	20,88	21,63	23,89	9,54	8,60	8,52
T_l	0,0	50,6	258,8	1738,4	2881,8	1840,2	202,2	0
	Узел 9	Узел 10	Узел 11	Узел 12	Узел 13	Узел 14	Узел 15	Узел 16
K	0,42	0,42	0,47	0,84	0,44	0,43	0,41	0,50
L	-	-	0,08	0,11	-	-	-	-
T_y	8,56	8,51	6,43	10,37	8,60	8,56	9,88	8,58
T_l	103,2	0	9,2	4301,0	175,6	100,0	2442,4	185,4
	Узел 17	Узел 18	Узел 19	Узел 20	Узел 21	Узел 22	Узел 23	Узел 24
K	0,39	0,79	0,75	0,52	0,52	0,52	0,95	0,45
L	-	-	-	0,10	0,02	0,02	4,18	0,13
T_y	8,61	23,89	22,11	6,51	5,70	5,72	1556,64	490,35
T_l	180,4	1584,2	143,0	0	0	0	0	8,8
	Узел 25	Узел 26	Узел 27	Узел 28	Узел 29	Узел 30	Узел 31	
K	3,98	3,98	0,20	1,19	0,98	0,08	0,93	
L	34,90	35,92	0,03	-	2,08	-	2,92	
T_y	976,13	1000,48	945,55	3397,68	677,15	1337,77	1986,14	
T_l	724,4	768,4	16,3	1858,6	529,2	16,7	0	

Перенаправление дополнительно 20% транзитного поездопотока на обходной путь привело к снижению времени движения поезда по главному ходу на 17,5% в среднем (T_{B1} , T_{31}), при этом загрузка обходного пути увеличилась незначительно, в среднем на 2% (T_{B2} и T_{32}). Благодаря этому стало возможно увеличить объем транзитных поездов на 15%. Однако появились и негативные эффекты: загрузка станции Иркутск-Сортировочный приблизилась к максимально-возможной; общее время «простоя» в сутки увеличилось на 7,8% в целом.

В обоих экспериментах вероятность отказа больше нуля, что недопустимо для такого типа систем. Однако она незначительна (0,001), и может быть объяснена тем, что в модели не учитывается наличие диспетчерского управления, которое стремится равномерно распре-

делить нагрузку на систему в течение суток и тем самым избежать большого скопления поездов на одном участке.

Таким образом, Иркутский железнодорожный узел справляется с текущей плановой нагрузкой и имеет запас производительности, благодаря которому возможно увеличить объем транзитных поездов на 15%. Однако потребуется перенаправить большую часть этого поездопотока на обходной путь и оптимизировать технологические процессы, проводимые с поездами на станции Иркутск-Сортировочный, в частности, увеличить число бригад технического осмотра. Дальнейшее повышение объемов транзитного поездопотока потребует изменений не только технологического характера, но и технического. Так, необходимо будет провести модернизацию станции Иркутск-Сортировочный и увеличить пропускную способность обходного пути.

5. Интеллектуальная система моделирования и анализа работы СеМО. При участии авторов разработан прототип интеллектуальной системы [13], которая предназначена для автоматизации построения имитационных моделей различных по сложности СеМО, с последующим их исследованием и формированием рекомендаций по улучшению параметров работы. Второстепенной целью ее создания является отчуждение модельно-алгоритмического аппарата от разработчика и предоставление возможности его использования неспециалисту в области теории массового обслуживания.

Опишем кратко архитектуру и основные функции интеллектуальной системы, которая за два года, прошедшие с публикации работы [13], претерпела определенные изменения. Итак, в ней имеется два программных модуля: модуль имитационного моделирования и экспертная система. Первый реализует имитационную модель работы СеМО, в которой допустимо наличие до 100 узлов и нескольких МАР- и ВМАР-потоков. Число каналов и длина очереди различны для каждого узла. Каналы работают независимо друг от друга, возможны разные законы распределения времени обслуживания и размеров обслуживаемых групп заявок в каналах. Данный модуль позволяет приближенно находить вероятности состояний и показатели эффективности СеМО, а также проводить многовариантные сценарные расчеты.

Экспертная система предназначена для автоматического анализа результатов имитационного моделирования, в частности, для определения «узких мест» в структуре СеМО, формирования рекомендаций по их устранению и увеличения пропускной способности системы в целом. В этот модуль также включена база знаний экспертов в виде продукционных правил. Экспертная система проводит поиск «узких мест» и выработку рекомендаций по следующему алгоритму:

Шаг 1. Загрузка параметров моделируемой СеМО и требований пользователя к ее пропускной способности. Пользователь может задать необходимую вероятность отказа или абсолютную пропускную способность.

Шаг 2. Модуль имитационного моделирования на основе полученных данных строит соответствующую модель и проводит ее численное исследование.

Шаг 3. Экспертная система с помощью продукционных правил из базы знаний проверяет характеристики СеМО, полученные в результате моделирования. Затем на их основе формирует список «узких мест». Если характеристики соответствуют требованиям пользователя, то выполняется Шаг 5, иначе переходим к Шагу 4.

Шаг 4. Из базы знаний выбираются решения для устранения найденных «узких мест» и вносятся соответствующие изменения в параметры моделируемой СеМО. Далее выполняется Шаг 2.

Шаг 5. Полученные результаты передаются пользователю.

Дополнительно интеллектуальная система также может отображать процесс работы СеМО в графическом виде и сохранять результаты в таблицах MS Excel.

Заключение. В статье представлена методика моделирования работы железнодорожных узлов, позволяющая учесть в модели: поступление различных категорий поездопотоков с двух и более направлений, нелинейную иерархическую структуру системы с несколькими возможными маршрутами движения поездов, а также местные поездопотоки, которые зарождаются на станциях. На основе предложенной методики была построена и исследована модель ИЖУ. В результате установлено, что он имеет достаточную пропускную способность для повышения транзитного поездопотока на 15%, что согласуется с планами ВСЖД филиала ОАО «РЖД» по развитию Транссибирской магистрали.

В работе также был затронут вопрос об отчуждаемости предложенного модельно-алгоритмического аппарата от разработчика. Его решение видится в создании интеллектуальной системы, которая позволит неспециалисту в области теории массового обслуживания использовать предложенный аппарат.

Дальнейшие исследования могут быть связаны с повышением точности и адекватности получаемых моделей. Этого можно достичь за счет учета движения отдельных вагонов по железнодорожному узлу и/или с построением детального описания станций, в частности, рассмотрением ее парков как отдельных элементов с различной производительностью.

Благодарности. Работа выполнена при финансовой поддержке РФФИ в рамках научного проекта № 20-010-00724; РФФИ и Правительства Иркутской области в рамках научного проекта № 20-47-383002.

СПИСОК ИСТОЧНИКОВ

1. Правдин Н.В., Вакуленко С.П., Головнич А.К. Проектирование инфраструктуры железнодорожного транспорта (станции, железнодорожные и транспортные узлы) / Н.В. Правдин, С.П. Вакуленко, А.К. Головнич М. [и др.]. Учебно-методический центр по образованию на железнодорожном транспорте, 2012. – 1086 с. – ISBN 978-5-89035-619-2
2. Бухаров Д.С., Казаков А.Л. Программная система "Виголт" для решения задач оптимизации, возникающих в транспортной логистике / Д.С. Бухаров, А.Л. Казаков // Вычислительные методы и программирование, 2012. – Т. 13. – №2. – С. 65–74. – eISSN 1726-3522.
3. Казаков А.Л., Лемперт А.А., Бухаров Д.С. Об одном численном методе решения некоторых задач оптимизации, возникающих в транспортной логистике/ А.Л. Казаков, А.А. Лемперт, Д.С. Бухаров // Вестник Иркутского государственного технического университета, 2011. – №6 (53). – С. 6–12. –ISSN 1814-3520.
4. Журавская М.А., Казаков А.Л., Парсюроева П.А. О размещении остановочных пунктов при осуществлении мультимодальных пассажирских перевозок / М.А. Журавская, А.Л. Казаков, П.А. Парсюроева // Транспорт Урала, 2012. –№4 (35). –С. 50–53. – ISSN 1815-9400.
5. Ковалев В.И., Верховых Г.В., Осьминин А.Т. Стохастические сети обслуживания в задачах моделирования работы железнодорожных сортировочных станций /В.И. Ковалев, Г.В. Верховых, А.Т. Осьминин [и др.]. – СПб: Информационный центр "Выбор", 2005. – 136 с. – ISBN 5-93518-036-7.
6. Казаков А.Л., Маслов А.М. Построение модели неравномерного транспортного потока на примере железнодорожной грузовой станции / А.Л. Казаков, А.М. Маслов // Современные технологии. Системный анализ. Моделирование, 2009. – №3 (23). – С. 27–32. – ISSN 1813-9108.

7. Козлов П.А., Тушин Н. А., Слободянюк И.Г. Макромоделирование железнодорожных станций и узлов / П.А.Козлов, Н. А. Тушин, И.Г.Слободянюк // Наука и техника транспорта, 2015. – №2. – С. 82–88. – ISSN: 2074-9325.
8. Козлов П.А., Колокольников В.С., Сорокин В.Н. Совместное использование аналитических методов и имитационных моделей /П.А. Козлов, В.С. Колокольников, В.Н. Сорокин // Транспорт Урала, 2016. – №3 (50). – С. 3–8. – ISSN 1815-9400.
9. Любченко А.А., Бартош С.В., Смирнов В.А., Castillo P.A. Дискретно-событийная модель железнодорожного узла в среде AnyLogic / А.А. Любченко, С.В. Бартош, В.А. Смирнов, [и др.] // Динамика систем, механизмов и машин, 2016. – №3. – С. 87–92. – ISSN 2310-9793.
10. Жарков М.Л., Казаков А.Л., Супруновский А.В., Павидис М.М. Моделирование крупнейшей в мире железнодорожной сортировочной станции с использованием теории массового обслуживания / М.Л. Жарков, А.Л. Казаков, А.В. Супруновский [и др.] // Вестник УрГУПС, 2021. – №3 (51). – С. 4–14. – DOI 10.20291/2079-0392-2021-3-4-14. –ISSN 2079-0392.
11. Bychkov I., Kazakov A., Lempert A., Zharkov M. Modeling of Railway Stations Based on Queuing Networks. Applied Sciences, 2021, vol. 11(5), pp. 2425, ISSN: 2076-3417.
12. Dudin A., Klimenok V., Vishnevsky V. The Theory of Queuing Systems with Correlated Flows. Springer Nature: Switzerland, 2019, 530 p, ISBN 978-3-030-32071-3.
13. Интеллектуальная технология моделирования железнодорожных станций на основе теории массового обслуживания /И.В. Бычков, А.Л. Казаков, М.Л. Жарков // Управление товарными потоками и перевозочным процессом на железнодорожном транспорте на основе клиентоориентированности и логистических принципов: коллективная монография членов и научных партнеров Объединенного ученого совета ОАО «РЖД»; под ред. Б.М. Лapidуса, А.Т. Осьминина. – СПб.: ЛЕМА, 2019. – С. 185–193. – ISBN 978-5-00105-520-4.

Жарков Максим Леонидович, к.т.н., научный сотрудник, Институт динамики систем и теории управления имени В.М. Матросова СО РАН, zharkm@mail.ru, Россия, г. Иркутск, ул. Лермонтова, 134

Супруновский Антон Викторович, старший преподаватель, Иркутский государственный университет путей сообщения, as.irgups@gmail.com, Россия, г. Иркутск, ул. Чернышевского, 15

On modeling sections of the railway network based on queuing theory

Maksim L. Zharkov¹, Anton V. Suprunovsky²

¹Matrosov Institute for System Dynamics and Control Theory of Siberian Branch of Russian Academy of Sciences, Irkutsk, Russia, zharkm@mail.ru

²Irkutsk State Transport University, Irkutsk, Russia, as.irgups@gmail.com

Abstract. In this paper, we proposed a methodology for the mathematical modeling of the movement of trains at railway junctions. It is based on the use of queueing networks of a special type. Sections of the railway network are important for cargo and passenger transportation, as well as international economic interaction. Therefore, their study is an urgent task. We use queueing networks, which allow us to take into account the influence of random factors on the operation of nodes and display the structural features of a particular object in the model. One of the large sections of the Trans-Siberian Railway sections was selected for testing the methodology. In the article, we built the model of its operation in the form of a queueing network with two incoming flows of applications and perform its numerical study. Then we used the results to determine the maximum allowable load and "bottlenecks" in the structure of the transport system and develop recommendations for increasing its throughput.

Keywords: sections of the railway network, traffic flow, mathematical model, queueing theory, computational experiment

Acknowledgements: The work was supported by the Russian Foundation for Basic Research within the framework of scientific project No. 20-010-00724; RFBR and the Government of the Irkutsk region within the framework of the scientific project no. 20-47-383002.

REFERENCES

1. Pravdin N.V., Vakulenko S.P., Golovnich A.K. [et al.]. *Proyektirovaniye infrastruktury zheleznodorozhnogo transporta (vokzaly, zheleznodorozhnye i transportnye uzly)* [Designing the infrastructure of railway transport (stations, railway and transport nodes)]. Moscow, Uchebno-metodicheskiy tsentr po obrazovaniyu na zheleznodorozhnom transporte, 2012, 1086 p, ISBN 978-5-89035-619-2.
2. Bukharov D.S., Kazakov A.L. *Programmnaya sistema "Vigolt" dlya resheniya zadach optimizatsii, vznikayushchikh v transportnoy logistike* [Vigolt system for solving transport logistics optimization problems]. *Vychislitel'nyye metody i programmirovaniye* [Computational methods and programming], 2012, vol. 13, no. 2. pp. 65–74, eISSN 1726-3522.
3. Kazakov A.L., Lempert A.A., Bukharov D.S. *Ob odnom chislennom metode resheniya nekotorykh zadach optimizatsii, vznikayushchikh v transportnoy logistike* [On a numerical method to solve some optimization problems occurring in transport logistics]. *Vestnik Irkutskogo gosudarstvennogo tekhnicheskogo universiteta* [Bulletin of the Irkutsk State Technical University], 2011, no. 6 (53), pp. 6–12, ISSN 1814-3520.
4. Zhuravskaya M. A., Kazakov A. L., Parsiurova P. A. *Parsyurova P.A. O razmeshchenii ostanovochnykh punktov pri osushchestvlenii mul'timodal'nykh passazhirskikh perevozok* [On placement of stopping points during multimodal passenger traffic]. *Transport Urala* [Transport of the Urals], 2012, no. 4 (35), pp. 50–53, ISSN 1815-9400.

5. Kovalev V.I., Verkhoviykh G.V., Osminin A.T. [et al.]. Stokhasticheskiye seti obsluzhivaniya v zadachakh modelirovaniya raboty zheleznodorozhnykh sortirovochnykh stantsiy [Stochastic service networks in the problems of modeling the work of railway sorting stations]. St. Petersburg, Information Center "Choice", 2005, 136 p, ISBN 5-93518-036-7.
6. Kazakov A. L., Maslov A. M. Postroenie modeli neravnomernogo transportnogo potoka na primere zheleznodorozhnoy gruzovoy stantsii [Constructing a model of uneven transport flow on the example of a railway freight station]. Modern technologies. System analysis. Modeling, 2009, no. 3 (23), pp. 27–32, ISSN 1813-9108.
7. Kozlov P.A., Tushin N.A., Slobodyanyuk I.G. Makromodelirovaniye zheleznodorozhnykh stantsiy i uzlov [Macro-modeling of railway stations and nodes]. Science and technology in transport, 2015, no. 2, pp. 82–88, ISSN 2074-9325.
8. Kozlov P.A., Kolokolnikov V.S., Sorokin V.N. Sovmestnoye ispol'zovaniye analiticheskikh metodov i imitatsionnykh modeley [Shared use of analytical methods and simulation models]. Transport Urala [Transport of the Urals], 2016, no. 3 (50), pp. 3–8, ISSN 1815-9400.
9. Lyubchenko A.A., Bartosh S.V., Smirnov V.A., Castillo P.Á. Diskretno-sobytiynaya model' zheleznodorozhnogo uzla v srede AnyLogic [Discrete-event model of a railway junction in the AnyLogic environment]. Dynamics of Systems, Mechanisms and Machines (Dynamics), 2016, no. 3, pp. 87–92, ISSN 2310-9793.
10. Zharkov M.L., Kazakov A.L., Suprunovsky A.V., Pavidis M.M. Modelirovaniye krupneyshey v mire zheleznodorozhnoy sortirovochnoy stantsii s ispol'zovaniyem teorii massovogo obsluzhivaniya [Simulation of the world's largest railway marshalling yard using queuing theory]. Vestnik UrGUPS [Bulletin of the UrGUPS], 2021, no. 3(51), pp. 4–14, ISSN 2079-0392.
11. Bychkov I., Kazakov A., Lempert A., Zharkov M. Modeling of Railway Stations Based on Queuing Networks. Applied Sciences, 2021, vol. 11(5), pp. 2425. ISSN 2076-3417.
12. Dudin A., Klimenok V., Vishnevsky V. The Theory of Queuing Systems with Correlated Flows. Springer Nature, Switzerland, 2019, 530 p, ISBN 978-3-030-32071-3.
13. Bychkov I.V., Kazakov A.L., Zharkov M.L. Intel'lectual'naya tekhnologiya modelirovaniya zheleznodorozhnykh stantsiy na osnove teorii massovogo obsluzhivaniya [Intelligent technology for modeling railway stations based on the theory of mass service]. Upravleniye tovarnymi potokami i perevozochnym protsessom na zheleznodorozhnom transporte na osnove kliyantooriyentirovannosti i logisticheskikh printsipov [Management of commodity flows and the transportation process in railway transport based on customer focus and logistics principles], St. Petersburg, LEMA, 2019, pp. 185–193, ISBN 978-5-00105-520-4.

Maksim. L. Zharkov, PhD, Researcher, Matrosov Institute for System Dynamics and Control, zharkm@mail.ru, Irkutsk, st. Lermontov, 134

Anton V. Suprunovsky, S. Lecturer, Irkutsk State Transport University, as.irgups@gmail.com, Irkutsk, Russia, Chernyshevsky st. 15

Статья поступила в редакцию 18.02.2022; одобрена после рецензирования 25.03.2022; принята к публикации 28.03.2022.

The article was submitted 02.18.2022; approved after reviewing 03.25.2022; accepted for publication 03.28.2022.

УДК 519.86

doi:10.38028/ESI.2022.25.1.011

Статистическая и динамическая задачи минимизации инновационного риска в системе управления инновационным потенциалом предприятия

Мансурова Юлия Талгатовна¹, Туктарова Полина Андреевна¹, Мансуров Тагир Валерьевич²

¹Уфимский государственный авиационный технический университет,

Россия, Уфа, mansurova.j@mail.ru, p.korobova@mail.ru

²ООО «СМК-10», Россия, Уфа, spacetag@gmail.com

Аннотация. В статье предлагается рассматривать процесс управления инновационным развитием предприятия как оптимизационную векторную задачу. При этом в качестве целевых функций предлагается рассматривать одновременно максимизацию прироста инновационного потенциала и минимизацию коэффициента риска. В качестве первого критерия оптимизации задана максимизация прироста инновационного потенциала, в качестве второго критерия – минимизация сопутствующего инновационного риска, который отражается в виде коэффициента риска через функцию полезности.

Ключевые слова: Риск, инновационный потенциал, конкурентоспособность, оптимизация

Цитирование: Мансурова Ю.Т., Туктарова П.А., Мансуров Т.В. Статистическая и динамическая задачи минимизации инновационного риска в системе управления инновационным потенциалом предприятия // Информационные и математические технологии в науке и управлении. – 2022. – № 1 (25). – С. 133-142. – DOI:10.38028/ESI.2022.25.1.011.

Введение. Инновационный процесс, охватывающий разработку инноваций, их внедрение в хозяйственную деятельность и дальнейшую коммерциализацию, является основой конкурентоспособности, развития и обеспечения экономической безопасности социально-экономической системы. Для определения оптимального размера инновационного потенциала, который должна иметь эффективно функционирующая социально-экономическая система, а также для получения корректных оценок влияния инновационного потенциала на обеспечение ее экономической безопасности, необходимо четко определить сущность инновационного потенциала и его структуру.

Постановка проблемы. В состав инновационного потенциала экономистами обычно включались следующие составные элементы: производственные фонды и работники; производственный аппарат и уровень технологии, природные ресурсы и материально-сырьевой баланс, система коммуникаций, техника, организация и система получения, хранения и переработки информации и т.д. Разнообразных трактовок и предложений достаточно много, но большинство исследователей [1, 2, 3] подчеркивают наличие в составе инновационного потенциала ресурсной составляющей.

Обобщая многолетний опыт отечественных исследований [4, 5], можно считать, что в состав экономического потенциала организации входят следующие основные элементы: финансовый (ФП), природно-ресурсный (ПРП), научный (НП), технический (ТП), кадровый (КП), информационный (ИП), организационный (ОП), потребительский (ПП) и творческий потенциалы (ТвП). Экономический и инновационный потенциалы по-разному связаны с

научным, техническим, научно-техническим, кадровым, интеллектуальным, промышленным, агропромышленным и другими разновидностями (компонентами) экономического потенциала. Экономический потенциал полностью вбирает в себя все другие потенциалы, а инновационный – только в той их части, в которой они могут быть реализованы в производстве товаров и услуг (как материальных, так и нематериальных). Однако инновационная сфера деятельности имеет существенные особенности, связанные с непрерывным риском, который необходимо учесть, как в текущем планировании, так и в перспективном. Учет рисков, связанных с нововведениями, должен проводиться при анализе составляющих инновационного потенциала [6].

Важнейшим резервом увеличения инновационного потенциала является обеспечение совместимости каждой из его составляющих со всеми другими. Работоспособность любой системы определяется качеством функционирования каждого ее элемента, а также эффективностью их взаимодействия. Изменение каждого элемента осуществляется в соответствии с тенденциями, логикой собственного развития, а также исходя из требований эффективного взаимодействия со всеми другими составляющими. Эти взаимодействия бывают очень сложными, и определить их можно лишь с позиций рассмотрения всей системы. Интегрирование взаимодействий, целенаправленное изменение составляющих является одним из путей совершенствования инновационного потенциала. Упрощенная предлагаемая матрица формирования требований к величине и структуре инновационного потенциала, исходя из взаимодействия его составляющих, показана в таблице 1.

Показанное формирование требований совместимости элементов является упрощенным отражением реального положения, но вместе с тем позволяет структурировать и упорядочивать, формализовать и количественно оценивать изменения, происходящие в инновационном потенциале, моделировать, анализировать и прогнозировать процессы его формирования и использования.

Таблица 1 – Формирование требований совместимости элементов инновационного потенциала (ИП)

№ п/п	Наименование элементов	Наименование элементов								
		1	2	3	4	5	6	7	8	9
		ФП	ПРП	НП	ТП	КП	ИП	ОП	ПП	ТвП
1	ФП	$\frac{y_{11}}{c_{11}(t)}$	$\frac{y_{12}}{c_{12}(t)}$	$\frac{y_{13}}{c_{13}(t)}$	$\frac{y_{14}}{c_{14}(t)}$	$\frac{y_{15}}{c_{15}(t)}$	$\frac{y_{16}}{c_{16}(t)}$	$\frac{y_{17}}{c_{17}(t)}$	$\frac{y_{18}}{c_{18}(t)}$	$\frac{y_{19}}{c_{19}(t)}$
2	ПРП		$\frac{y_{22}}{c_{22}(t)}$	$\frac{y_{23}}{c_{23}(t)}$	$\frac{y_{24}}{c_{24}(t)}$	$\frac{y_{25}}{c_{25}(t)}$	$\frac{y_{26}}{c_{26}(t)}$	$\frac{y_{27}}{c_{27}(t)}$	$\frac{y_{28}}{c_{28}(t)}$	$\frac{y_{29}}{c_{29}(t)}$
3	НП			$\frac{y_{33}}{c_{33}(t)}$	$\frac{y_{34}}{c_{34}(t)}$	$\frac{y_{35}}{c_{35}(t)}$	$\frac{y_{36}}{c_{36}(t)}$	$\frac{y_{37}}{c_{37}(t)}$	$\frac{y_{38}}{c_{38}(t)}$	$\frac{y_{39}}{c_{39}(t)}$
4	ТП				$\frac{y_{44}}{c_{44}(t)}$	$\frac{y_{45}}{c_{45}(t)}$	$\frac{y_{46}}{c_{46}(t)}$	$\frac{y_{47}}{c_{47}(t)}$	$\frac{y_{48}}{c_{48}(t)}$	$\frac{y_{49}}{c_{49}(t)}$
5	КП					$\frac{y_{55}}{c_{55}(t)}$	$\frac{y_{56}}{c_{56}(t)}$	$\frac{y_{57}}{c_{57}(t)}$	$\frac{y_{58}}{c_{58}(t)}$	$\frac{y_{59}}{c_{59}(t)}$
6	ИП						$\frac{y_{66}}{c_{66}(t)}$	$\frac{y_{67}}{c_{67}(t)}$	$\frac{y_{68}}{c_{68}(t)}$	$\frac{y_{69}}{c_{69}(t)}$
7	ОП							$\frac{y_{77}}{c_{77}(t)}$	$\frac{y_{78}}{c_{78}(t)}$	$\frac{y_{79}}{c_{79}(t)}$
8	ПП								$\frac{y_{88}}{c_{88}(t)}$	$\frac{y_{89}}{c_{89}(t)}$
9	ТвП									$\frac{y_{99}}{c_{99}(t)}$

В таблице 1 в числителях на пересечении строки i и столбца j стоят величины y_{ij} - приращения инновационного потенциала, вызываемые взаимодействием его элементов i и j с точки зрения их совместимости и взаимодействия.

Величина y_{ij} может быть отрицательной или равной нулю. В знаменателях стоят функции $c_{ij}(t)$, характеризующие вложения, затраты в расчете на единицу по годам для достижения данного y_{ij} . При построении математической модели инновационного потенциала в нее могут быть включены не только парные взаимодействия y_{ij} , но и тройственные взаимодействия y_{ijk} , а также взаимодействия более высоких порядков (в зависимости от особенностей системы).

Следует стремиться к оптимизации величин y_{ij} , что можно сформулировать в виде следующей задачи условной оптимизации функционала Φ : найти такие оптимальные значения y_{ij}^* , чтобы

$$\begin{aligned} \Phi(y_{ij}) &= \sum_{i=1}^n \sum_{j=1}^m y_{ij} \rightarrow \max; \\ \sum_{i=1}^n \sum_{j=1}^m c_{ij} \cdot y_{ij} &\leq A; \\ y_{ij} &\geq 0. \end{aligned} \quad (1)$$

где A – количество возможных ресурсов, выделенных на приращение инновационного потенциала.

Среди инновационных процессов различают следующие виды неопределенности по степени их приближения к детерминированным процессам [7].

1) детерминированный случай, когда можно пренебречь влиянием случайных и неопределенных параметров. Здесь могут быть использованы методы математического моделирования детерминированных систем, в частности методики бизнес-планов;

2) зашумленный детерминированный случай. При этом влияние вектора случайных и неопределенных параметров незначительно. В этом случае также могут использоваться детерминированные методы, дополненные специальными методами оценки различных характеристик шумов и их элиминации;

3) стохастический случай неопределенности, когда в модели системы есть случайные величины, для которых известна или может быть оценена по наблюдениям функция распределения. Здесь используется математический аппарат классической теории вероятностей. Примером служит вычисление различных мер риска;

4) случай полной неопределенности, когда часть компонентов вектора случайных и неопределенных параметров, являются величинами неопределенными. Для них лишь можно указать границы изменения, используя метод экспертных оценок [8].

Большинство инновационных процессов обычно можно отнести к третьему классу. Таким образом, применение стандартных методов оценки риска становится некорректным и возникает необходимость введения новых критериев оценки. Обычно под неопределенной ситуацией принято понимать такую ситуацию, когда имеется целый набор возможных последствий принимаемого решения. Часто этот набор последствий характеризуется всевозможными значениями некоторой случайной величины X . Эту случайную величину принято называть функцией отдачи [10]. Если последствия описываются как значения функции отдачи, то функционалом полезности будет функция $U(x)$, обладающая следующим свойством: значение x функции отдачи тогда и только тогда предпочтительней значения y , когда $U(x) > U(y)$.

В качестве количественного показателя риска обычно рассматривается коэффициент риска [11]. Пусть x – функция отдачи, являющаяся дискретной случайной величиной, $\psi(x)$ – монотонно возрастающая функция полезности. При этом события $x < z$ и $x \geq z$ являются соответственно неблагоприятным и благоприятным. Тогда коэффициент риска вычисляется по формуле:

$$K_z = - \frac{\sum_{x_i < z} (\psi(z) - \psi(x_i)) p_i}{\sum_{x_i \geq z} (\psi(z) - \psi(x_i)) p_i} \quad (2)$$

С учетом того, что коэффициент риска вводился для получения наиболее адекватного метода оценки эффективности инновационного процесса, возникает необходимость постановки математической задачи оптимизации коэффициента риска. Основные предпосылки этой задачи, следующие: должен быть набор управляющих воздействий – это вектор $U = (u_1, u_2, \dots, u_n)$, измеримый количественно. Индивид, принимающий решение, должен иметь возможность по своему усмотрению менять значение U в допустимой области $U \in U_{\text{доп}}$ и тем самым влиять на числовое значение функции цели, например на K_z , зависящей от U так, чтобы достигался экстремум этой функции:

$$K_z \rightarrow \min \text{ или } K_z = - \frac{\sum_{x_i < z} (\psi(z) - \psi(x_i)) p_i}{\sum_{x_i \geq z} (\psi(z) - \psi(x_i)) p_i} \rightarrow \min \quad (3)$$

При этом, кроме управлений, накладываются ограничения и на функцию цели

$$K_z \leq K_{\text{доп}} \text{ или } K_z \geq K_{\text{доп}}.$$

Многокритериальная оптимизация. Принимая во внимание, что ранее была выполнена постановка задачи максимизации приращения инновационного потенциала за счет основных ресурсов, можно перейти к задаче многокритериальной оптимизации. При этом вектор управления U будет заменен вектором отдачи от использования различных видов потенциала. Таким образом, возникает необходимость решения двухкритериальной задачи. Первым критерием выступает необходимость достижения максимума приращения инновационного потенциала, а вторым критерием – минимизация риска (коэффициента риска), который сопутствует этому приращению [12].

Однако для определения коэффициента риска возникает необходимость построения функции полезности. Пусть на срок t лет задан плановый уровень приращения инновационного потенциала организации, который равен z . В случае, если реальный прирост будет меньше планового уровня, то фирма потеряет часть своей конкурентоспособности на рынке. Следовательно, эти действия приведут к дополнительным затратам, которые равны β долей. Причем данный исход является неблагоприятным для организации, а значит, в функции полезности он отражается со знаком минус. С другой стороны, если реальный прирост инновационного потенциала превысит запланированную величину, фирма достигнет необходимого уровня конкурентоспособности на рынке. При этом в случае, если реальный прирост инновационного потенциала будет полностью совпадать с намеченным уровнем, то функция полезности примет нулевое значение, то есть будет характеризовать собой точку перехода от неблагоприятных событий к благоприятным. Тогда функция полезности имеет вид:

$$\psi(x) = \begin{cases} -\frac{(z-x)}{z} \cdot (1+\beta) \sum_{i=1}^n \sum_{j=1}^m c_{ij} y_{ij}, & \text{если } x < z, \\ (x-z), & \text{если } x \geq z \end{cases} \quad (4)$$

где $\frac{(z-x)}{z}$ – доля недостающего прироста инновационного потенциала по отношению к достигнутому уровню; $x = \sum_{i=1}^n \sum_{j=1}^m y_{ij}$ – совокупный прирост инновационного потенциала; β – доля удорожания затрат в случае привлечения сторонних сил. Имея представленную функцию полезности, можно составить коэффициент риска:

$$K_z = - \frac{-\frac{(z-x)}{z} \cdot (1+\beta) \sum_{i=1}^n \sum_{j=1}^m c_{ij} y_{ij}}{(x-z)} = \frac{\frac{(z-x)}{z} \cdot (1+\beta) \sum_{i=1}^n \sum_{j=1}^m c_{ij} y_{ij}}{(x-z)} \quad (5)$$

Данный коэффициент риска представляет собой второй критерий, который должен достигать своего минимума. Кроме того, этот критерий дополняется ограничением по допустимому уровню значения коэффициента риска.

Рассмотренная выше задача относится к классу статических моделей, то есть исследуемая система и ее параметры анализируются в фиксированный единственный момент времени. Данная постановка при моделировании процесса управления инновационным потенциалом организации возможна, но представляет больше аналитический интерес, чем практический. Поскольку с практической точки зрения понятие «процесс» всегда связано с динамической характеристикой. Таким образом, наиболее адекватной и применимой является динамическая постановка задачи управления что для инновационного развития предприятия является инновационным потенциалом. В этой связи требуется дальнейшее усложнение рассмотренной ранее модели.

Постановка задачи в динамической форме возможна в следующих вариантах [13].

1) состояние системы в момент времени t зависит от состояний системы в предыдущие периоды:

а) время дискретно. Тогда возможны два типа зависимостей переменных модели от времени:

– рекуррентная, $y^{(t)} = \Phi(y^{(t-1)}, y^{(t-2)}, \dots)$. В частном случае, когда рекуррентная зависимость задана в виде $y^{(t)} = \Phi(y^{(t-1)})$ (текущее состояние зависит только от предыдущего), то получаем задачу динамического программирования. Каждому переходу соответствует эффект $\Phi_{ks}^t(y) = \Phi(y_k^t, y_s^{t+1})$ – эффект от перехода от узла k на слое t к узлу s на слое $t + 1$. Недостатком такой постановки является заранее ограниченное и фиксированное число состояний на каждом временном слое, что ограничивает пути развития системы в целом. Увеличение числа состояний делает задачу более громоздкой, резко увеличивая ее размерность, что приводит к излишнему усложнению модели, не повышая ее точность. Однако в случае заранее определенных возможных путей развития системы данный подход является предпочтительным, так как связан с определением оптимального пути развития системы из множества возможных;

– зависимость через сеточные функции $\Phi(t, y)$. В отличие от случая независимых шагов по времени в этом случае накладываются дополнительные зависимости на связь переменных по времени. Недостатком этого подхода является эмпирическое задание этих зависимостей, фактически, устанавливающих приоритеты развития системы, которые следует отыскивать на каждом временном шаге;

б) время непрерывно. Это задача оптимального управления, связанная с нахождением функций управления $y_{ij}(t)$, максимизирующих функционал интегрального эффекта $\int_0^T (\sum_i \sum_j y_{ij}(t)) dt \rightarrow \max$. Так как функции управления являются непрерывными, то они задают траектории развития системы при наложенных ограничениях. Наиболее простым способом решения задачи оптимального управления является ее сведение путем дискретизации к задаче нелинейного программирования [14]. Недостатком этого способа является очень большая размерность полученной задачи нелинейного программирования, даже при небольшом количестве шагов дискретизации;

2) время дискретно и шаги по времени независимы друг от друга. В этом случае получается обобщенная модель задачи в статической постановке, так как ко всем ограничениям и к целевой функции добавляется дискретная сумма по времени. Это самый простой случай, но в то же время наименее адекватный, так как на каждом шаге по времени не учитываются состояния системы в прежние моменты времени, таким образом, пренебрежение «историей» процесса не позволяет получить достоверные результаты относительно ее состояния в будущем.

Проведя сравнительный анализ возможных постановок задач оптимизации управления инновационным потенциалом в динамической постановке, можно заключить, что наиболее

адекватным методом является решение задачи в динамической постановке с определенным множеством состояний на каждом временном шаге. Это множество состояний определяется, исходя из приоритета критериев, и может представлять собой различные плановые состояния системы потенциалов в каждый момент времени, например, следующие:

- максимально возможный эффект от взаимодействия потенциалов на слое t ;
- наиболее вероятный эффект на слое t ;
- минимальный эффект на слое t .

Каждому состоянию k на слое t соответствует своя матрица взаимодействий. Стоимость перехода от элемента k на слое t к элементу s на слое $t + 1$ можно определить по формуле:

$$C_t(k, s) = \sum_i \sum_j c_{ij} \Delta y_{ij}^{(t)} \quad (6)$$

Эффект от перехода от элемента k на слое t к элементу s на слое $t + 1$ можно определить по формуле:

$$\Phi_t(k, s) = \alpha F_t^H(k, s) - \beta K_Z^H(k, s), \alpha + \beta = 1, \quad (7)$$

где $F_t^H(k, s)$, $K_Z^H(k, s)$ - нормированные значения критериев суммарного эффекта от взаимодействия потенциалов и риска.

В формулах (6) и (7) значения функций $F_t^H(k, s)$, $K_Z^H(k, s)$ можно рассчитать по формулам

$$F_t(k, s) = \sum_{j=2}^n (k_j \sum_{i=1}^{j-1} \Delta y_{ij}^{(t)}) \quad (8)$$

$$K_Z(k, s) = \frac{\sum_{X_i \leq z} \frac{z - X_i}{z} (1 + \beta) \sum_{j=2}^n \sum_{i=1}^{j-1} c_{ij} \Delta y_{ij}^{(t)}}{\sum_{X_i > z} (X_i - z)} \quad (9)$$

Изменение эффекта от взаимодействия потенциалов i -го и j -го видов в формулах (8) и (9) вычисляется по формуле:

$$\Delta y_{ij}^{(t)} = \begin{cases} y_{ij}^{(k,t+1)} - y_{ij}^{(s,t)}, & \text{если } y_{ij}^{(k,t+1)} > y_{ij}^{(s,t)} \\ 0, & \text{если } y_{ij}^{(k,t+1)} \leq y_{ij}^{(s,t)} \end{cases}, j = 2..n, i = 1..j-1, t = 1..T-1. \quad (10)$$

Необходимо определить оптимальное развитие системы, то есть такой путь, который бы доставлял максимум функции суммарного эффекта при ограничениях на затраты. Оптимальной траекторией является путь, максимизирующий функцию $\sum_{t=1}^T \Phi_t(k_t, s_t)$ – суммарный эффект от взаимодействия потенциалов за все время при заданных ограничениях.

Исходными данными являются траектории предполагаемого изменения эффектов от взаимодействия каждого из потенциалов при наилучшем, наиболее вероятном и наихудшем вариантах развития событий. Для каждого исхода задан вектор эффектов для каждого типа потенциалов, исходя из определенных приоритетов.

Метод динамического программирования позволяет значительно сократить количество вариантов, имеющих место при простом переборе, до размерности матрицы условно-оптимального управления $(3 \times T)$. Итак, задача состоит в выборе пути (из возможных), доставляющем оптимум целевой функции.

Над стрелками показано значение агрегированного критерия при переходе от одного элемента схемы к другому. Суммарное значение агрегированного критерия для оптимального пути достигает своего максимального значения, равного 2,12.

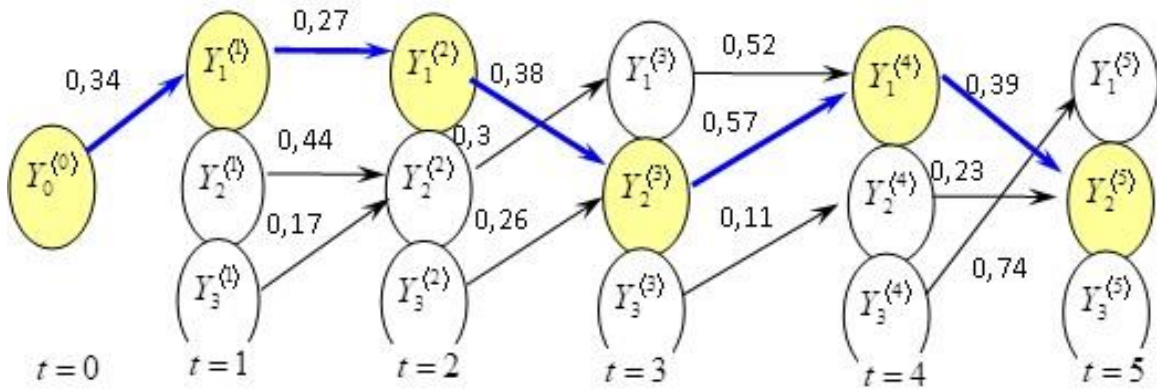


Рис. 1. Условно-оптимальное управление процессом развития системы взаимодействия потенциалов

На каждом шаге (кроме последнего) условно-оптимальное управление определяется с учетом значения условно-оптимального управления на предыдущем шаге. Такая рекуррентная зависимость позволяет учесть всю «предысторию» процесса, то есть строить оптимальное управление на каждом шаге таким образом, чтобы обеспечивалось дальнейшее развитие системы оптимальным образом.

Так как начальное состояние определено, то условно-оптимальное управление U_0^u для первого шага будет являться числом (а не вектором), указывающим на переход к траектории $i, i = \overline{1,3}$ на шаге $t = 1$. Это упрощает процедуру построения оптимального управления, поскольку нужно просто выписать условно-оптимальное управление, начиная с первого шага по времени и до последнего.

Вектор оптимального управления можно определить следующим образом: $U^T = (2 \ 1 \ 1 \ 2 \ 1 \ 2)$.

После определения траекторий развития можно перейти к нахождению динамики изменения эффекта от взаимодействия потенциалов.

Чтобы перейти от безразмерного вида к абсолютным значениям приращения эффекта от взаимодействия потенциалов за исследуемый период времени T , необходимо найти величину

$$E = \sum_{t=1}^T \sum_i \sum_j \Delta y_{ij}^{(t)}, \quad (11)$$

где суммирование по времени ведется по оптимальной траектории развития системы.

Рассчитанные таким образом значения суммарного эффекта от взаимодействия потенциалов показаны на рисунке 2 (за начальный уровень эффекта принято значение, полученное при решении стационарной задачи).

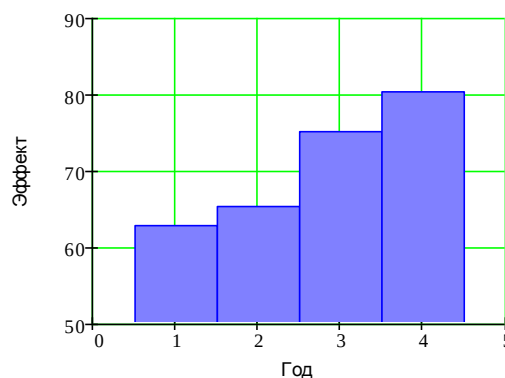


Рис. 2. Динамика суммарного эффекта от взаимодействия потенциалов

Суммарный прирост за время $T = 5$ лет составит $E=26,71$ условных единиц, а суммарный эффект за это же время составит 81,02 условных единиц. Необходимо отметить, что

данный результат является не максимальным решением, а компромиссным, поскольку при нахождении условно-оптимального управления использовался агрегированный критерий.

Заключение. Представленная выше модель управления инновационным потенциалом коммерческой организации, в отличие от существующих немногочисленных методов оценки инновационного потенциала, позволяет не столько оценивать уже существующий потенциал, сколько управлять его развитием, а также находить оптимальные параметры прироста инновационного потенциала. Кроме того, в отечественной и зарубежной литературе отмечается, что для инновационной деятельности риск является неотъемлемой и значительной частью. Однако существующие модели предполагают лишь оценку этого риска для тех или иных инновационных проектов. С помощью представленной модели появляется возможность управлять инновационным риском, сводя его до минимального значения, не теряя при этом максимального значения совокупного прироста инновационного потенциала. К тому же в модели учтены объективные ограничения, налагаемые на процесс данного управления.

В динамическом случае была поставлена и решена задача по определению оптимальной траектории развития системы взаимодействующих потенциалов, которая соответствует максимальному значению агрегированной функции с ограничениями по ресурсам и эффекту от взаимодействия. В агрегированную функцию входят два критерия – суммарного эффекта и риска, поэтому найденная траектория развития оптимальна по двум этим параметрам.

СПИСОК ИСТОЧНИКОВ

1. Ресурсы инноваций: организационный, финансовый, административный: Учеб. пособие для вузов/ Под ред. И.П. Николаевой. – М.: Юнити-Дана, 2003. – 102 с.
2. Зинов В. Г. Управление интеллектуальной собственностью. АНХ, Центр коммерциализации технологий. М.: Монолит, 2002. – 154 с.
3. Родионова Л.Н., Пашин С.Т. и др. Управление инновациями: экономические аспекты. – Уфа: ГУП РБ УПК, 2009. – 248 с.
4. Бублик Н.Д. Проблемы инновационного развития современной экономики. – Уфа, 2002. – 98 с.
5. Бублик Н.Д., Силантьев В.Б. Риск-ресурс: Проблемы венчурно-стохастической деятельности. – Уфа, 1999, – 84 с.
6. Инновационный менеджмент в России: вопросы стратегического управления и научно-технологической безопасности / Рук. авт. колл. В.Л. Макаров и А.Е. Варшавский. М.: Наука, 2004. – 880 с
7. Фролов И.Э. Научоемкий сектор промышленности РФ: экономико-технологический механизм ускоренного развития. М.: Макс Пресс, 2004. – 320 с.
8. Яшин С.Н., Солдатова Ю.С. Совершенствование и практическая апробация методики оценки экономического состояния и уровня инновационного развития предприятия // Финансы и кредит, 2013. – № 12. – С. 39-47
9. Евтушенко Е., Юсупова Э. Оценка инновационного потенциала предприятия // Инвестиции и инновации, 2006. – № 11. – С. 63-68
10. Матузова И.В. Методика оценки инновационного потенциала промышленного предприятия // Вестник Ленинградского государственного университета им. А.С. Пушкина, 2012– Вып. № 4. – Т. 6. – С. 87-97.
11. Туктарова П.А., Мансурова Ю.Т. Формирование системы показателей деятельности промышленного предприятия в процессе идентификации кризисных явлений// Вопросы экономики и права. – 2018. – № 118. – 97-101 с.
12. Мансурова Ю.Т. Управление инновационным потенциалом предприятия // экономические науки, –2009. – № 60. – 364-368 с.

13. Анисимов Ю.П., Пешкова И.В., Солнцева Е.В. Методика оценки инновационной деятельности предприятия // Инновации, 2006. – № 11. – С. 49-55
14. Митякова О.И. Оценка инновационного потенциала промышленного предприятия // Финансы и кредит, 2004. – № 13. – С. 69-74

Мансурова Юлия Талгатовна. 1980 г. рождения. Окончила ФГБОУ ВО УГНТУ (2002). Канд. экон. наук (2005). Доцент кафедры Экономики предпринимательства УГАТУ, mansurova.j@mail.ru, (РИНЦ): 1005445, SPIN-код: 4596-9811, ORCID: 0000-0001-7373-0344, Россия, Уфа

Туктарова Полина Андреевна. 1985 г. рождения. Окончила ФГБОУ ВО УГНТУ (2007). Магистр экономики (УГАТУ, 2012). Канд. экон. наук (2005). Доцент кафедры Экономики предпринимательства, p.korobova@mail.ru, AuthorID (РИНЦ): 944221, SPIN-код: 7889-8580, ORCID: 0000-0003-0773-3138, Россия, Уфа

Мансуров Тагир Валерьевич. 1980 г. рождения. Окончил ФГБОУ ВО УГНТУ (2002). Канд. техн. наук (2004). Директор ООО «Строительно-монтажная компания 10», spacetag@gmail.com, AuthorID (РИНЦ): 377489, ORCID: 0000-0002-6439-6126, Россия, Уфа

UDC 519.86

doi:10.38028/ESI.2022.25.1.011

Statistical and dynamic problems of minimizing innovation risk in the system of managing the innovative potential of an enterprise

Yulia T. Mansurova¹, Polina A. Tuktarova¹, Tagir V. Mansurov²

¹Ufa State Aviation Technical University, Russia, Ufa, mansurova.j@mail.ru, p.korobova@mail.ru.

²SMK-10 LLC, Russia, Ufa, spacetag@gmail.com

Abstract. In article it is offered to consider process of management of innovative development of the enterprise as an optimizing vector task. At the same time as criterion functions it is offered to consider at the same time maximizing gain of innovative potential and minimization of coefficient of risk. Therefore, this aspect is key when developing strategy of innovative development of the enterprise. As the first criterion of optimization maximizing gain of innovative potential, as the second criterion – minimization of the accompanying innovative risk which is reflected in the form of risk coefficient through function of usefulness is set.

Keywords: risk, innovative potential, competitiveness, optimization

REFERENCES

1. Resursy innovatsii organizatsionnyi finansovyi administrativnyi: Ucheb posobie dlia vuzov [Innovation resources: organizational, financial, administrative: Proc. allowance for universities]. Ed. I.P. Nikolaeva. - M.: Unity-Dana, 2003, 102 p.
2. Zinov VG Upravlenie intellektualnoi sobstvennostiu [Management of intellectual property]. Academy of National Economy, Technology Commercialization Center. M.: Monolith, 2002, 154 p.
3. Rodionova L.N., Pashin S.T. Problemy innovatsionnogo razvitiia sovremennoi ekonomiki [Management of innovations: economic aspects]. Ufa: GUP RB UPC, 2009, 248 p.
4. Bublik N.D., Silantiev V.B. Risk-resurs Problemy venchurno-stokhasticheskoi deiatelnosti [Risk resource: Problems of venture-stochastic activity]. - Ufa, 1999. - 84 p.

5. Bublik N.D. Problemy innovatsionnogo razvitiia sovremennoi ekonomiki [Problems of innovative development of modern economy]. Ufa, 2002, 98 p.
6. Innovatsionnyi menedzhment v Rossii voprosy strategicheskogo upravleniia i nauchno-tekhnologicheskoi bezopasnosti [Innovation management in Russia: issues of strategic management and scientific and technological security]. Ruk. ed. coll. V.L. Makarov and A.E. Warsaw. M.: Nauka, 2004, 880 p.
7. Frolov I.E. Naukoemkii sektor promyshlennosti RF ekonomiko-tekhnologicheskii mekhanizm uskorenno razvitiia [Science-intensive industry sector of the Russian Federation: economic and technological mechanism of accelerated development]. M.: Maks Press, 2004, 320 p.
8. Yashin S.N., Soldatova Yu.S. Sovershenstvovanie i prakticheskaia aprobatsiia metodiki otsenki ekonomicheskogo sostoiianiia i urovnia innovatsionnogo razvitiia predpriiatiia [Improvement and practical approbation of the methodology for assessing the economic condition and the level of innovative development of an enterprise] // Finance and credit. 2013, no. 12, pp. 39-47
9. Evtushenko E., Yusupova E. Otsenka innovatsionnogo potentsiala predpriiatiia [Assessment of the innovative potential of the enterprise] // Investments and innovations. 2006, no. 11, pp. 63-68
10. Matuzova I.V. Metodika otsenki innovatsionnogo potentsiala promyshlennogo predpriiatiia [Methodology for assessing the innovative potential of an industrial enterprise]. Bulletin of the Leningrad State University. A.S. Pushkin, no. 4, v. 6, 2012, pp. 87-97.
11. Tuktarova P.A., Mansurova Yu.T. Formirovanie sistemy pokazatelei deiatelnosti promyshlennogo predpriiatiia v protsesse identifikatsii krizisnykh iavlenii [Formation of a system of indicators of the activity of an industrial enterprise in the process of identifying crisis phenomena] // Questions of Economics and Law, 2018, no. 118, 97-101p.
12. Mansurova Yu.T. Upravlenie innovatsionnym potentsialom predpriiatiia [Management of the innovative potential of the enterprise] // economic sciences, 2009, no. 60, 364-368p.
13. Anisimov Yu.P., Peshkova I.V., Solntseva E.V. Metodika otsenki innovatsionnoi deiatelnosti predpriiatiia [Methodology for assessing the innovative activity of an enterprise] // Innovations, 2006, no. 11, pp. 49-55
14. Mityakova O.I. Otsenka innovatsionnogo potentsiala promyshlennogo predpriiatiia [Assessment of the innovative potential of an industrial enterprise] // Finance and credit, 2004, no. 13, pp. 69-74

Yulia T. Mansurova, 1980 year of birth. Graduated from FGBOU VO UGNTU (2002). Cand. economy Sciences (2005). Associate Professor of the Department of Economics of Entrepreneurship, USATU, mansurova.j@mail.ru, AuthorID (RSCI): 1005445, SPIN: 4596-9811, ORCID: 0000-0001-7373-0344, Russia, Ufa

Polina A. Tuktarova, born in 1985. Graduated from FGBOU VO UGNTU (2007). Master of Economics (UGATU, 2012). Cand. economy Sciences (2005). Associate Professor of the Department of Economics of Entrepreneurship, USATU, p.korobova@mail.ru. AuthorID (RSCI): 944221, SPIN code: 7889-8580, ORCID: 0000-0003-0773-3138, Russia, Ufa

Tagir V. Mansurov, 1980 year of birth. Graduated from FGBOU VO UGNTU (2002), spacetag@gmail.com. Cand. tech. Sciences (2004). Director of Construction and Installation Company 10 LLC. AuthorID (RSCI): 377489, ORCID: 0000-0002-6439-6126, Russia, Ufa

Статья поступила в редакцию 05.03.2022; одобрена после рецензирования 22.03.2022; принята к публикации 28.03.2022.

The article was submitted 03.05.2022; approved after reviewing 03.22.2022; accepted for publication 03.28.2022.

ПАМЯТИ НИКОЛАЯ ИВАНОВИЧА ВОРОПАЯ



28 февраля 2022 г. ушел из жизни Николай Иванович Воропай, бывший директор и научный руководитель Института систем энергетики им. Л.А. Мелентьева СО РАН*.

Н.И. Воропай родился 1 ноября 1943 года. В 1966 году после окончания Ленинградского политехнического института (ЛПИ), приехал в Иркутск и пришел на работу в Сибирский энергетический институт СО АН СССР (позже СЭИ переименован в ИСЭМ СО РАН). С тех пор он прошел по всем ступенькам служебной лестницы: младший, затем старший научный сотрудник, заведующий лабораторией и отделом, заместитель директора и с 1997 года – директор института. На посту директора находился до 2016 года.

Н.И. Воропай – доктор технических наук (1990), профессор (1993), заслуженный деятель науки РФ (1999), почетный энергетик (2003). Награжден Орденом Дружбы (2004) и Орденом Почета (2013). Стал лауреатом Государственной премии СССР (1986) и двух премий Правительства РФ (1999, 2011), премий им. Г.М. Кржижановского (2005, 2021) и им. И.М. Губкина (2004).

Н.И. Воропай – известный специалист в области системных исследований энергетики, автор и соавтор более 700 опубликованных научных работ. Воропаем Н.И. выявлены и исследованы фундаментальные особенности сложных протяженных электроэнергетических систем, что позволило на основе понятных физических представлений объяснить многие специфические явления и процессы,

например, системные аварии, и разработать теоретические основы анализа и синтеза структурно неоднородных энергообъединений. Это направление является методической базой исследований в рамках действующей ведущей научной школы - под его руководством подготовлено более 20 кандидатов наук.

Под руководством Воропая Н.И. были разработаны методология, математические модели и методы обоснования развития электроэнергетики, электроэнергетических систем и компаний в условиях многих критериев и несовпадающих интересов субъектов отношений. Методология не имеет аналогов в электроэнергетической науке.

Воропай Н.И. был одним из инициаторов нового научного направления – обеспечение энергетической безопасности России и ее регионов. В последнее время активно занимался проблематикой интеллектуальных энергетических систем будущего. Много лет возглавлял Международный семинар им. Ю.Н. Руденко «Методические вопросы исследования надежности больших систем энергетики».

Н.И. Воропай был действительным членом Академии электротехнических наук РФ и Международной энергетической академии, членом Президиума Российского научно-технического общества энергетиков и электротехников, членом ряда научных советов РАН и СО РАН, член НТС ФСК ЕЭС и РАН, председателем диссертационного совета ИСЭМ СО РАН, членом технического комитета Международной федерации по автоматическому управлению (IFAC), членом Международной конференции по большим электрическим системам CIGRE, почетным членом IEEE и других организаций, главным редактором международного журнала «Energy Systems Research», состоял в редколлегиях известных российских и международных изданий.

Н.И. Воропай всегда принимал активное участие в научной и общественной жизни института, которое всегда было весомым, нужным и бескорыстным. Во многом благодаря ему в ИСЭМ СО РАН сложился коллектив единомышленников со здоровой и дружеской атмосферой, называемой «духом СЭИ». Его личные качества – искренность, человечность, терпеливость в сочетании с целеустремленностью, мудростью и интеллигентностью – это качества настоящего лидера и в институте, и за его стенами. Это большая утрата не только для российской и мировой академической науки.

Николай Иванович был членом редколлегии нашего журнала «Информационные и математические технологии в науке и управлении», членом программного комитета одноименной Байкальской Всероссийской конференции и Международного семинара «Критические инфраструктуры в цифровом мире». Мы очень благодарны ему за то, что он всегда поддерживал наши начинания, принимал активное участие в работе семинара. Последнее заседание семинара (16-21 марта), куда он собирался поехать, но не успел, мы посвятили его памяти.

Его уход из жизни – большая потеря для всех, кто знал его, работал и дружил с ним. Светлая память о нем навсегда останется в наших сердцах.

* Жизнеописание и фото – с сайта ИСЭМ СО РАН

