

БИЗНЕС- ИНФОРМАТИКА

Научный журнал НИУ ВШЭ

СОДЕРЖАНИЕ

А.С. Каукин, П.Н. Павлов, В.С. Косарев

Краткосрочное прогнозирование цен на электроэнергию с использованием генеративных нейронных сетей..... 7

Р.Д. Гутгарц

Влияние алгоритмизации и интерфейса на подготовку управленческих решений..... 24

Л.А. Родионова, Е.Д. Копнова

Применение мер «тяжеловестности» в задачах анализа финансовых временных рядов..... 38

Г.С. Завалин, О.В. Недолужко, К.С. Солодухин

Формирование каузального поля показателей развития интеллектуального капитала организации: концепция и нечеткая экономико-математическая модель..... 53

А.Л. Бекларян, Л.А. Бекларян, А.С. Акопов

Имитационная модель интеллектуальной транспортной системы «умного города» с адаптивным управлением светофорами на основе нечеткой кластеризации 70

А.В. Зиненко

Прогнозирование финансовых временных рядов с использованием сингулярного спектрального анализа..... 87



Издатель:
Национальный
исследовательский университет
«Высшая школа экономики»

Подписной индекс
Объединенного каталога
«Пресса России» – E79128

Выпускается ежеквартально

Журнал включен в Перечень
российских рецензируемых
научных журналов,
в которых должны быть
опубликованы основные научные
результаты диссертаций
на соискание ученых степеней
доктора и кандидата наук

Главный редактор
Е.П. Зараменских

Заместитель главного редактора
Э.А. Бабкин

Компьютерная верстка
О.А. Богданович

Дизайн обложки
О.А. Богданович
(с использованием изображения,
сгенерированного нейросетью Midjourney)

Администратор веб-сайта
И.И. Хрусталева

Адрес редакции:
119049, г. Москва,
ул. Шаболовка, д. 26-28
Тел./факс: +7 (495) 772-9590 *28509
<http://bijournal.hse.ru>
E-mail: bijournal@hse.ru

За точность приведенных сведений
и содержание данных,
не подлежащих открытой публикации,
несут ответственность авторы

**При перепечатке ссылка на журнал
«Бизнес-информатика» обязательна**

Тираж:
русскоязычная версия – 100 экз.,
англоязычная версия – 100 экз.,
онлайн-версии на русском и английском –
свободный доступ

Отпечатано в типографии НИУ ВШЭ
г. Москва, Измайловское шоссе, д. 44, стр. 2

© Национальный
исследовательский университет
«Высшая школа экономики»

О ЖУРНАЛЕ

«**Б**изнес-информатика» – рецензируемый междисциплинарный научный журнал, выпускаемый с 2007 года Национальным исследовательским университетом «Высшая школа экономики» (НИУ ВШЭ). Администрирование журнала осуществляется Высшей школой бизнеса НИУ ВШЭ. Журнал выпускается ежеквартально, на русском и английском языках.

Миссия журнала – развитие бизнес-информатики как новой области информационных технологий и менеджмента. Журнал осуществляет распространение последних разработок технологического и методологического характера, способствует развитию соответствующих компетенций, а также обеспечивает возможности для дискуссий в области применения современных информационно-технологических решений в бизнесе, менеджменте и экономике.

Журнал публикует статьи по следующей тематике: моделирование социальных и экономических систем, цифровая трансформация бизнеса, управление инновациями, информационные системы и цифровые технологии в бизнесе, анализ данных и системы бизнес-интеллекта, математические методы и алгоритмы бизнес-информатики, моделирование и анализ бизнес-процессов, поддержка принятия управленческих решений.

Журнал «Бизнес-информатика» включен в Перечень рецензируемых научных изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученых степеней кандидата и доктора наук (Перечень ВАК).

Журнал входит в базы Scopus, Web of Science Emerging Sources Citation Index (WoS ESCI), Russian Science Citation Index на платформе Web of Science (RSCI), EBSCO.

Журнал распространяется как в печатном виде, так и в электронной форме.

РЕДАКЦИОННАЯ КОЛЛЕГИЯ

ГЛАВНЫЙ РЕДАКТОР

Зараменских Евгений Петрович

Национальный исследовательский университет
«Высшая школа экономики», Москва, Россия

ЗАМЕСТИТЕЛЬ ГЛАВНОГО РЕДАКТОРА

Бабкин Эдуард Александрович

Национальный исследовательский университет
«Высшая школа экономики», Нижний Новгород, Россия

ЧЛЕНЫ РЕДКОЛЛЕГИИ

Авдошин Сергей Михайлович

Национальный исследовательский университет
«Высшая школа экономики», Москва, Россия

Акопов Андраник Сумбатович

Центральный экономико-математический институт РАН,
Москва, Россия

Алескеров Фуад Тагиевич

Национальный исследовательский университет
«Высшая школа экономики», Москва, Россия

Афанасьев Александр Петрович

Институт проблем передачи информации им. А.А. Харкевича
РАН, Москва, Россия

Афанасьев Антон Александрович

Центральный экономико-математический институт РАН,
Москва, Россия

Баранов Александр Павлович

Главный научно-исследовательский вычислительный центр
Федеральной налоговой службы, Москва, Россия

Барахнин Владимир Борисович

Федеральный исследовательский центр информационных
и вычислительных технологий, Новосибирск, Россия

Беккер Йорг

Университет Мюнстера, Мюнстер, Германия

Вестнер Маркус

Технический университет прикладных наук,
Регенсбург, Германия

Гаврилова Татьяна Альбертовна

Санкт-Петербургский государственный университет,
Санкт-Петербург, Россия

Глотен Эрве

Тулонский университет, Ла-Гард, Франция

Гурвич Владимир Александрович

Ратгерский университет (Университет Нью-Джерси),
Ратгерс, США

Джейкобс Лоренц

Университет Цюриха, Цюрих, Швейцария

Дискин Иосиф Евгеньевич

Всероссийский центр изучения общественного мнения,
Москва, Россия

Зандкуль Курт

Университет Ростока, Росток, Германия

Иванников Александр Дмитриевич

Институт проблем проектирования в микроэлектронике РАН,
Москва, Россия

Исаев Дмитрий Валентинович

Национальный исследовательский университет
«Высшая школа экономики», Москва, Россия

Калягин Валерий Александрович

Национальный исследовательский университет
«Высшая школа экономики», Нижний Новгород, Россия

Кравченко Татьяна Константиновна

Национальный исследовательский университет
«Высшая школа экономики», Москва, Россия

Кузнецов Сергей Олегович

Национальный исследовательский университет
«Высшая школа экономики», Москва, Россия

Лугачев Михаил Иванович

Московский государственный университет
им. М.В. Ломоносова, Москва, Россия

Лин Квей-Жей

Технологический институт Нагои, Нагоя, Япония

Мальцева Светлана Валентиновна

Национальный исследовательский университет
«Высшая школа экономики», Москва, Россия

Мейор Питер

Комиссия ООН по науке и технологиям, Женева,
Швейцария

Миркин Борис Григорьевич

Национальный исследовательский университет
«Высшая школа экономики», Москва, Россия

Назаров Дмитрий Михайлович

Уральский государственный экономический университет,
Екатеринбург, Россия

Пальчунов Дмитрий Евгеньевич

Новосибирский государственный университет, Новосибирск,
Россия

Пардалос Панайот (Панос)

Университет Флориды, Гейнсвилл, США

Пастор Оскар

Политехнический университет Валенсии, Валенсия,
Испания

Посегга Йоахим

Университет Пассау, Пассау, Германия

Самуйлов Константин Евгеньевич

Российский университет дружбы народов, Москва, Россия

Стоянова Ольга Владимировна

Национальный исследовательский университет
«Высшая школа экономики», Санкт-Петербург, Россия

Триболе Жозе

Университет Лиссабона, Лиссабон, Португалия

Ульянов Михаил Васильевич

AVECO, Любляна, Словения

Ускенбаева Раиса Кабиевна

Казахский национальный исследовательский технический
университет им. К.И. Сатпаева, Алматы, Казахстан

Цуканова Ольга Анатольевна

Санкт-Петербургский национальный исследовательский
университет информационных технологий, механики
и оптики, Санкт-Петербург, Россия

Чхартишвили Александр Гедеванович

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

Штраус Кристина

Университет Вены, Вена, Австрия

ISSN 1998-0663 (print), ISSN 2587-8166 (online)

English version: ISSN 2587-814X (print), ISSN 2587-8158 (online)

Vol. 17 No. 3 – 2023

BUSINESS INFORMATICS

HSE Scientific Journal



Publisher:
HSE University

The journal is published quarterly

The journal is included
into the list of peer reviewed
scientific editions established
by the Supreme Certification
Commission of the Russian Federation

CONTENTS

A.S. Kaukin, P.N. Pavlov, V.S. Kosarev

Short-term forecasting of electricity prices
using generative neural networks 7

R.D. Gutgarts

Influence of algorithmization and interface
for the preparation of management decisions..... 24

L.A. Rodionova, E.D. Kopnova

Application of measures of heavy-tailedness
in problems for analysis of financial time series..... 38

G.S. Zavalin, O.V. Nedoluzhko, K.S. Solodukhin

Formation of the causal field of indicators
for an organization's intellectual capital
development: A concept and a fuzzy
economic and mathematical model..... 53

A.L. Beklaryan, L.A. Beklaryan, A.S. Akopov

Simulation model of an intelligent
transportation system for the “smart city”
with adaptive control of traffic lights
based on fuzzy clustering..... 70

A.V. Zinenko

Forecasting financial time series using
singular spectrum analysis..... 87

Editor-in-Chief
E. Zaramenskikh

Deputy Editor-in-Chief
E. Babkin

Computer making-up
O. Bogdanovich

Cover design
O. Bogdanovich
(using an image generated
by the Midjourney neural network)

Website administration
I. Khrustaleva

Address:

26-28, build. 4, Shablovka Street
Moscow 119049, Russia

Tel./fax: +7 (495) 772-9590 *28509

<http://bijournal.hse.ru>

E-mail: bijournal@hse.ru

Circulation:

English version – 100 copies,
Russian version – 100 copies,
online versions in English and Russian –
open access

Printed in HSE Printing House
44, build. 2, Izmaylovskoye Shosse,
Moscow, Russia

© HSE University

ABOUT THE JOURNAL

Business Informatics is a peer reviewed interdisciplinary academic journal published since 2007 by HSE University, Moscow, Russian Federation. The journal is administered by HSE Graduate School of Business. The journal is issued quarterly, in English and Russian.

The mission of the journal is to develop business informatics as a new field within both information technologies and management. It provides dissemination of latest technical and methodological developments, promotes new competences and provides a framework for discussion in the field of application of modern IT solutions in business, management and economics.

The journal publishes papers in the following areas: modeling of social and economic systems, digital transformation of business, innovation management, information systems and technologies in business, data analysis and business intelligence systems, mathematical methods and algorithms of business informatics, business processes modeling and analysis, decision support in management.

The journal is included into the list of peer reviewed scientific editions established by the Supreme Certification Commission of the Russian Federation.

The journal is included into Scopus, Web of Science Emerging Sources Citation Index (WoS ESCI), Russian Science Citation Index on the Web of Science platform (RSCI), EBSCO.

The journal is distributed both in printed and electronic forms.

EDITORIAL BOARD

EDITOR-IN-CHIEF

Evgeny P. Zaramenskikh

HSE University, Moscow, Russia

DEPUTY EDITOR-IN-CHIEF

Eduard A. Babkin

HSE University, Nizhny Novgorod, Russia

EDITORIAL BOARD

Sergey M. Avdoshin

HSE University, Moscow, Russia

Andranik S. Akopov

Central Economics and Mathematics Institute, Russian Academy of Sciences, Moscow, Russia

Fuad T. Aleskerov

HSE University, Moscow, Russia

Alexander P. Afanasyev

Institute for Information Transmission Problems (Kharkevich Institute), Russian Academy of Sciences, Moscow, Russia

Anton A. Afanasyev

Central Economics and Mathematics Institute, Russian Academy of Sciences, Moscow, Russia

Vladimir B. Barakhnin

Federal Research Center of Information and Computational Technologies, Novosibirsk, Russia

Alexander P. Baranov

Federal Tax Service, Moscow, Russia

Jörg Becker

University of Munster, Munster, Germany

Alexander G. Chkhartishvili

V.A. Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia

Tatiana A. Gavrilova

Saint-Petersburg University, St. Petersburg, Russia

Hervé Glotin

University of Toulon, La Garde, France

Vladimir A. Gurvich

Rutgers, The State University of New Jersey, Rutgers, USA

Laurence Jacobs

University of Zurich, Zurich, Switzerland

Iosif E. Diskin

Russian Public Opinion Research Center, Moscow, Russia

Dmitry V. Isaev

HSE University, Moscow, Russia

Alexander D. Ivannikov

Institute for Design Problems in Microelectronics, Russian Academy of Sciences, Moscow, Russia

Valery A. Kalyagin

HSE University, Nizhny Novgorod, Russia

Tatiana K. Kravchenko

HSE University, Moscow, Russia

Sergei O. Kuznetsov

HSE University, Moscow, Russia

Kwei-Jay Lin

Nagoya Institute of Technology, Nagoya, Japan

Mikhail I. Lugachev

Lomonosov Moscow State University, Moscow, Russia

Svetlana V. Maltseva

HSE University, Moscow, Russia

Peter Major

UN Commission on Science and Technology for Development, Geneva, Switzerland

Boris G. Mirkin

HSE University, Moscow, Russia

Dmitry M. Nazarov

Ural State University of Economics, Ekaterinburg, Russia

Dmitry E. Palchunov

Novosibirsk State University, Novosibirsk, Russia

Panagote (Panos) M. Pardalos

University of Florida, Gainesville, USA

Óscar Pastor

Polytechnic University of Valencia, Valencia, Spain

Joachim Posegga

University of Passau, Passau, Germany

Konstantin E. Samouylov

Peoples' Friendship University, Moscow, Russia

Kurt Sandkuhl

University of Rostock, Rostock, Germany

Olga Stoyanova

HSE University, St. Petersburg, Russia

Christine Strauss

University of Vienna, Vienna, Austria

José M. Tribolet

Universidade de Lisboa, Lisbon, Portugal

Olga A. Tsukanova

Saint-Petersburg National Research University of Information Technologies, Mechanics and Optics, St. Petersburg, Russia

Mikhail V. Ulyanov

AVECO, Ljubljana, Slovenia

Raissa K. Uskenbayeva

Kazakh National Technical University after K.I. Satpaev, Almaty, Kazakhstan

Markus Westner

Technical University for Applied Sciences (OTH Regensburg), Regensburg, Germany

DOI: 10.17323/2587-814X.2023.3.7.23

Краткосрочное прогнозирование цен на электроэнергию с использованием генеративных нейронных сетей

А.С. Каукин 

E-mail: kaukin@iep.ru

П.Н. Павлов 

E-mail: pavlov@ranepa.ru

В.С. Косарев 

E-mail: kosarev-vs@ranepa.ru

Российская академия народного хозяйства и государственной службы при Президенте Российской Федерации
Адрес: Россия, 119571, г. Москва, пр. Вернадского, д. 82.

Аннотация

В статье изучаются прогностические способности генеративно-состязательного нейросетевого подхода в отношении временных рядов на примере прогнозирования цен для узлов свободного рынка электроэнергии России на сутки вперед. В результате серии экспериментов мы приходим к выводу, что генеративно-состязательная сеть, состоящая из двух моделей (генератора и дискриминатора), позволяет достичь минимума функции ошибки с большей обобщающей способностью, чем, при прочих равных, достигается в результате оптимизации статичного аналога генеративной модели – рекуррентной нейронной сети. Собственные эмпирические результаты показывают, что при околонулевой среднеквадратической ошибке на тренировочном множестве, демонстрируемой одновременно рекуррентной и генеративной моделями, ошибка последней на тестовом множестве ниже. Состязательный подход также превзошел в точности вневыборочного прогноза альтернативные эталонные модели: сверточную нейронную сеть, адаптированную для прогнозирования временных рядов, и авторегрессионную линейную модель. Практика применения предложенного подхода показала, что генеративно-состязательная модель с заданной универсальной архитектурой и ограниченным числом объясняющих факторов при условии дообучения на данных, специфичных для целевого узла энергосистемы может использоваться для прогнозирования цен в узлах рынка на сутки вперед без существенных отклонений.

Ключевые слова: рынок электроэнергии, рынок на сутки вперед, временные ряды, генеративная нейронная сеть, рекуррентная нейронная сеть

Цитирование: Каукин А.С., Павлов П.Н., Косарев В.С. Краткосрочное прогнозирование цен на электроэнергию с использованием генеративных нейронных сетей // Бизнес-информатика. 2023. Т. 17. № 3. С. 7–23.
DOI: 10.17323/2587-814X.2023.3.7.23

Введение

На территории России действует двухуровневый рынок электроэнергии и мощности. Рынок на сутки вперед является основной площадкой, где приобретают электроэнергию и мощность промышленные производители. Рынок на сутки вперед (далее – РСВ) представляет собой систему, где осуществляется конкурентный отбор ценовых заявок поставщиков и покупателей электроэнергии за сутки до ее реальной поставки (потребителям) с определением цен и объемов поставки на каждый час суток¹. На РСВ осуществляется маргинальное ценообразование, иными словами цена определяется на основании баланса спроса и предложения и распространяется на всех участников данного рынка.

Уровень цен на электроэнергию напрямую влияет на издержки компаний. Наличие точного прогноза цен на электроэнергию необходимо для планирования операционной деятельности предприятий-представителей энергоемких отраслей, а также для построения финансовых моделей. В свою очередь нелинейная природа динамики цен на электроэнергию делает их прогнозирование сложной эмпирической задачей².

Существует множество эмпирических работ, как отечественных, так и зарубежных исследователей, где нейросетевые методы использовались для прогнозирования цен свободного рынка электроэнер-

гии и продемонстрировали свое превосходство над линейными моделями. Однако в последнее время появляются свидетельства того, что использование генеративно-состязательного подхода (generative adversarial networks, GAN)³ позволяет дополнительно повысить точность нейронной сети, построенной и оптимизированной специально для решения определенной задачи.

GAN является одновременно представителем статистических методов прогнозирования и методов имитационного моделирования⁴. Используя набор доступных предикторов, неопытный эксперт (генеративная нейронная сеть в начале обучения) делает прогнозы о цене электроэнергии в определенном узле, а другая нейронная сеть (дискриминатор) учится их отличать от реальных данных. В процессе обучения сети стимулируют друг друга ко все более точному решению поставленной перед ними индивидуальной задачи. Обучение продолжается, пока прогнозы эксперта (генератора) не достигают необходимой точности в соответствии с выбранной метрикой.

Цель представленной работы заключается в апробации методологии построения генеративно-состязательных сетей (GAN) для прогнозирования узловых цен на электроэнергию в России на рынке на сутки вперед, а также в проведении сравнительного анализа качества прогнозов на основе методологии GAN и прогнозов, полученных альтернативными (эталонными) методами.

¹ Конкурентный отбор заявок проводится коммерческим оператором (АО «АТС»). Цена РСВ определяется для каждого узла двух ценовых зон: первая ценовая зона включает территории Европейской части России и Урала (Центрального, Северо-Западного (за исключением территорий, относящихся к неценовым зонам), Южного, Северо-Кавказского, Приволжского и Уральского федеральных округов), вторая – территорию Сибири (Сибирского федерального округа).

² Характерной особенностью цен свободного рынка является значительная волатильность, сложная структура сезонности (годовая, внутринедельная, внутринедневная сезонность), а также наличие частых выбросов. Последнее может быть обусловлено, как нештатными ситуациями в энергосистеме, так и конъюнктурой отраслей-потребителей, что делает невозможным построение полной структурной модели рынка.

³ Симбиоз двух нейронных сетей: генератора и дискриминатора. Генератор предназначен непосредственно для построения прогнозов. Дискриминатор предназначен для того, чтоб отличать реальные данные от прогноза генератора и стимулировать генератор обучаться эффективнее.

⁴ Методы имитационного моделирования, как правило, используются для репрезентации структуры, связей элементов сложных экономических систем, где в явном виде необходимо учитывать феномен взаимодействия множества агентов [1–3].

Для достижения цели в представленной работе решаются следующие задачи. Во-первых, обобщаются подходы к прогнозированию цен на электроэнергию, в том числе с использованием генеративно-состязательных нейронных сетей. Во-вторых, приводится описание данных, характеризующих оптовый рынок электроэнергии (PCB) в России. В-третьих, характеризуются методологические подходы к разработке генеративно-состязательной нейросетевой модели для прогнозирования цены на электроэнергию на уровне отдельных узлов энергосистемы в России. В-четвертых, приводится обзор эмпирических результатов исследования, сравниваются результаты прогнозирования цен на электроэнергию с использованием методологии генеративно-состязательных нейронных сетей (GAN) и основных бенчмарков, включая базовую рекуррентную сеть, сверточную нейронную сеть и авторегрессионную модель (ARIMA), широко используемую в моделировании временных последовательностей в сфере экономики и финансов [4–6]. В заключении приводится краткое обобщение результатов проведенного исследования.

1. Анализ литературы

1.1. Классические подходы к прогнозированию цен на электроэнергию на PCB

Для решения задачи прогнозирования цен на электроэнергию использовались различные экономико-математические и имитационные модели. Отечественными учеными в работе [1], к примеру, использовалась имитационная модель на основе CGE (модель общего экономического равновесия) для поиска оптимальных темпов роста тарифов в электроэнергетике по регионам РФ в регулируемом сегменте рынка. Разработанная авторами система уравнений учитывала взаимодействие множества экономических агентов: потребитель, производитель, импортеры и экспортеры электроэнергии, государство.

Для прогнозирования цен свободного рынка, система одновременных уравнений использовалась в работе [7]. Авторы моделировали спотовые цены на электроэнергию в Скандинавии, используя модель на основе 29 уравнений, которая учитывала функциональные связи между климатическими факторами, развитием снежного покрова, водностью рек (гидроэлектростанции основной

источник электроэнергии в Скандинавии), параметрами энергосистемы. В работе [8] была разработана структурная модель спотовых цен на электроэнергию для Новой Англии в США, учитывающая функциональные связи между ценами на топливо, а также спросом на электроэнергию и доступностью генерирующих мощностей.

Имитационный подход является действенным инструментом моделирования экономики на основе аналитического представления взаимодействия различных агентов и позволяет учитывать физические процессы, влияющие на экономику. К примеру, в работе [2] была представлена имитационная модель для прогнозирования динамики добычи нефти по скважинам, учитывающая реализованные и плановые геолого-технические мероприятия на каждой скважине. В другом случае, имитационный подход был использован для разработки цифрового двойника завода по выпуску телевизионной техники [3]. Однако использование имитационного подхода в прогностических целях, в случае с узловыми ценами на PCB представляется затруднительным, поскольку требует наличия подробной информации о топологии энергетической сети, ее параметрах и условиях работы.

Экономико-математические подходы в части моделей временных рядов являются более распространенным для прогнозирования цен свободного рынка. Их можно обобщить в две группы: статистические и методы машинного обучения. К статистическим методам обычно относят аддитивные эконометрические модели [9]. К примеру, в работах [5, 6, 10] для прогнозирования цен свободного рынка использовались различные вариации авторегрессионных моделей (ARIMA, ARMAX, AGARCH). Однако по мере развития и популяризации методов машинного обучения (machine learning, ML) стало появляться все больше работ, где линейные эконометрические модели в сравнении с такими представителями методов машинного обучения, как модели опорных векторов, градиентный бустинг и нейронные сети [9, 11, 12], оказывались менее точными в краткосрочном прогнозировании цен.

Множество исследователей обращались к нейронным сетям для прогнозирования цен свободного рынка электроэнергии [12–14]. При этом большинство исследователей, по состоянию на 2020 г. [15], отдавало предпочтение многослойному перцептрон (multilayer perceptron, MLP⁵). Отечественные ис-

⁵ Multilayer perceptron – это искусственная нейронная сеть, которая характеризуется несколькими слоями входных узлов, соединенных в виде ориентированного графа между входным и выходным слоями.

следователи также неоднократно применяли данную архитектуру для прогнозирования цен РСВ. К примеру, Марьясин и Лукашова использовали MLP с двумя скрытыми слоями для прогноза свободных цен на электроэнергию в Ярославской области. К MLP-архитектуре обращались отечественные ученые Золотова и Дворкин [16]. В своем исследовании авторы предложили использовать перцептрон с 8 нейронами в скрытом слое для прогнозирования почасового индекса равновесных цен первой ценовой зоны.

В зарубежной литературе существует множество работ с использованием иных архитектур, хорошо зарекомендовавших себя в задачах прогнозирования временных рядов в других сферах. К примеру, комбинация сверточной и рекуррентной нейронных сетей использовалась для прогнозирования цен и спроса на электроэнергию в работах [14, 17]. Как отмечается в работе [14], подобная архитектура оправдала себя во многих областях, где для прогнозирования требовалось извлечение одновременно как временных, так и пространственных характеристик временных рядов. Авторы исследования [12] предложили использовать сверточную нейронную сеть с расширенными свертками⁶ для прогноза цен на оптовом рынке электроэнергии в канадском городе Онтарио. Однако в последнее время появляются свидетельства того, что генеративно-сопоставительный подход способен улучшить результаты сети любой архитектуры при условии использования ее в качестве генератора в GAN [18].

1.2. Генеративно-сопоставительные нейросети

1.2.1. Общая характеристика

Базовая теория генеративно-сопоставительных сетей с примерами практического использования изложена в монографии Николенко [19]. В самом простом варианте генеративно-сопоставительная модель состоит из двух искусственных нейронных сетей, которые поочередно взаимодействуют друг с другом. Одна из них, генератор — порождает объекты в пространстве данных, а вторая, дискриминатор — учится отличать порожденные генератором объекты от настоящих примеров из обучающей выборки.

Сопоставительная компонента во взаимодействии двух сетей состоит в том, что генератор должен научиться обманывать дискриминатор, а дискримина-

тор, несмотря на это, должен корректно отличать сгенерированные примеры от настоящих. Формализация подобного описания в терминах теории игр приводит к минимаксной задаче оптимизации, которую можно записать в виде уравнения 1:

$$\min_G \max_D V(D, G), \text{ где } \max V(D, G) = \\ = E_{x \sim p_{data}(x)} [\log D(x)] + E_{z \sim p_z(z)} \log (1 - D(G(z))), \quad (1)$$

где $D(x)$ — функциональная форма дискриминатора; $G(z)$ — функциональная форма генератора;

$p_z(z)$ — порождаемое генератором распределение данных;

$p_{data}(x)$ — распределение фактических данных.

На практике функциональные формы дискриминатора и генератора могут представлять собой любые архитектуры нейронных сетей. Решение минимаксной задачи обеспечивает поочередное обучение: генератора с фиксированными весами дискриминатора и наоборот — дискриминатора с фиксированными весами генератора.

Генеративно-сопоставительные сети впервые описали и применили на практике в 2014 году Гудфеллоу и Йошуа Бенжи [20]. Впоследствии их идея получила широкое практическое применение: сопоставительные сети позволили добиться значительных результатов в таких областях, как генерация изображения по текстовому описанию [21], создание лекарственных препаратов [22], генерация псевдо-реалистичных временных последовательностей с сохранением основных моментов распределения [23] и т.п.

Например, в случае генерации изображения по описанию, текст преобразуется в числовое признаковое пространство с помощью рекуррентного кодировщика, а затем эти признаки используются как условие в порождающей изображению сети GAN. В результате, к примеру, лицо человека на фотографии можно сначала отобразить в пространство признаков, а потом изменить признак возраста и сгенерировать новое изображение. Таким образом, можно искусственно «состарить» или «омолодить» человека.

Для создания нового лекарственного препарата исследователи из работы [22] использовали сопоставительный автокодировщик для порождения

⁶ Адаптация сверточной сети для прогнозирования временных рядов, которая позволяет учитывать широкий диапазон истории при прогнозировании.

молекул, которые могут быть перспективными кандидатами для создания на их основе новых лекарственных средств.

Временные ряды представляют собой уникальный объект для генеративного моделирования. В работе [23] отмечается, что модели прогнозирования временных рядов, такие как классические или нейросетевые авторегрессии, по своей сути детерминированы. Генеративные модели в свою очередь позволяют добавить в нейросетевой вывод элемент случайности.

1.2.2. Применение GAN для прогнозирования рынка электроэнергии

В зарубежной исследовательской практике существует ряд примеров использования генеративно-состязательных сетей для прогнозирования рынка электроэнергии. К примеру, в работе [24] авторы проводили апробацию модели генеративно-состязательной сети на двух наборах данных: потребление электроэнергии на уровне отдельного домохозяйства и динамика обменного курса. В результате генеративно-состязательная модель в обоих экспериментах превзошла по точности ее детерминированный эквивалент – генеративную нейронную сеть, обученную независимо.

В работе [25] исследователи предлагают модель на основе генеративно-состязательных сетей для прогнозирования поузловых цен части энергетической системы США. Нейросетевая модель использует пространственно-временные корреляции между историческими ценами в узлах и принимает в качестве входных данных исторические цены, упорядоченные в трехмерный тензор⁷. Этот тензор состоит из ряда упорядоченных по времени матриц. В свою очередь, каждая матрица – это фактически карта поузловых цен с сохранением пространственного расположения узлов. Задача генеративной модели в данном случае заключалась в генерации новой матрицы с прогнозными поузловыми ценами на электроэнергию. Базовая модель обучалась строить прогноз на час вперед.

В работе [26] генеративная состязательная сеть используется для прогноза оптовых цен на электроэнергию с интервалом в 30 минут для энергетического рынка Австралии. В отличие от предыдущих

работ, авторы строят не точечную оценку цены, а интервальную. Генеративная сеть позволила авторам получить прогнозные интервалы, покрывающие редкие и экстремальные наблюдения более точно, чем альтернативные вероятностные модели.

Следуя подходу, описанному в работе [24], для разработки собственной генеративно-состязательной модели мы используем двухшаговую процедуру. На первом шаге мы разрабатываем и оптимизируем рекуррентную нейронную сеть для решения задачи прогнозирования цены на электроэнергию в случайном узле энергосистемы. На втором шаге мы инкорпорируем полученную нейросеть в GAN в качестве генератора и проверяем устойчивость модели на некотором подмножестве узлов. Такой подход позволяет сузить пространство поиска архитектуры GAN до дискриминатора и дает возможность проверить гипотезу – возможно ли за счет включения в архитектуру GAN улучшить работу базовой нейронной сети.

2. Данные

Анализ проводится на основе почасовых отчетов администратора торговой системы (АТС) о равновесных ценах в наиболее крупных узлах энергосистемы [27]. База данных охватывает период с 13.04.2019 по 31.12.2022 и содержит информацию о 7215 узлах в 66 регионах РФ. Динамика усредненных цен по всем узлам энергосистемы РФ представлена на *рис. 1*.

Несмотря на стохастическое поведение каждого из рядов поузловых цен можно отметить тенденцию к пространственной корреляции внутри одного региона. Иллюстрацией чего служит кластеризация статистических характеристик цен: среднего и стандартного отклонения за рассматриваемый период на примере некоторых субъектов РФ (*рис. 2*).

Существует ряд особенностей, которые необходимо учесть при формировании перечня объясняющих переменных:

1. Существенная автокорреляция цен. Данный феномен подробно описан в работе Золотовой и Дврокина [16] на уровне ценовых зон. Собственный анализ показал, что в частных случаях – на уровне узлов картина в целом схожа.

2. Наличие годовой сезонности. Сезонная компонента изменяется синхронно в большинстве ре-

⁷ Под тензором в данном случае подразумевается объект машинного обучения – массив матриц или многомерный массив данных.

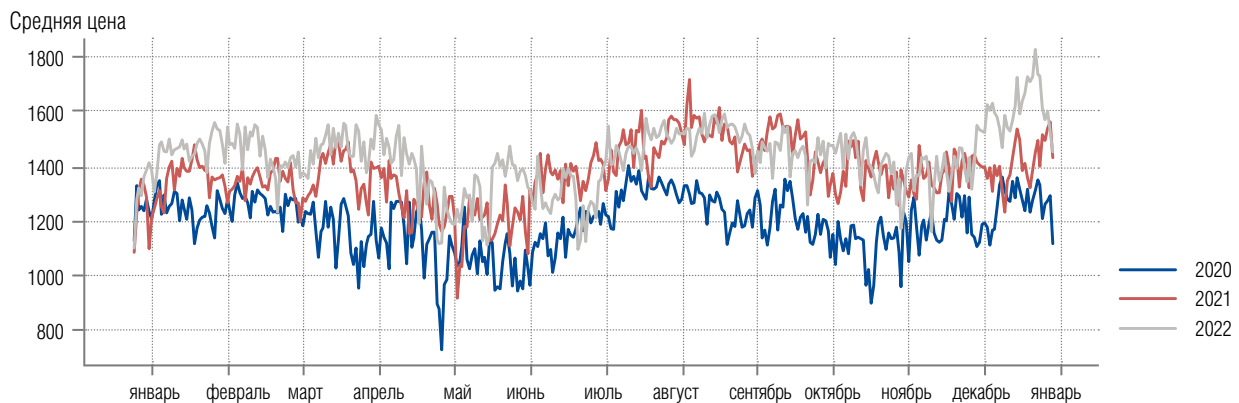


Рис. 1. Усредненные цены на РСВ по всем узлам энергетической системы РФ.

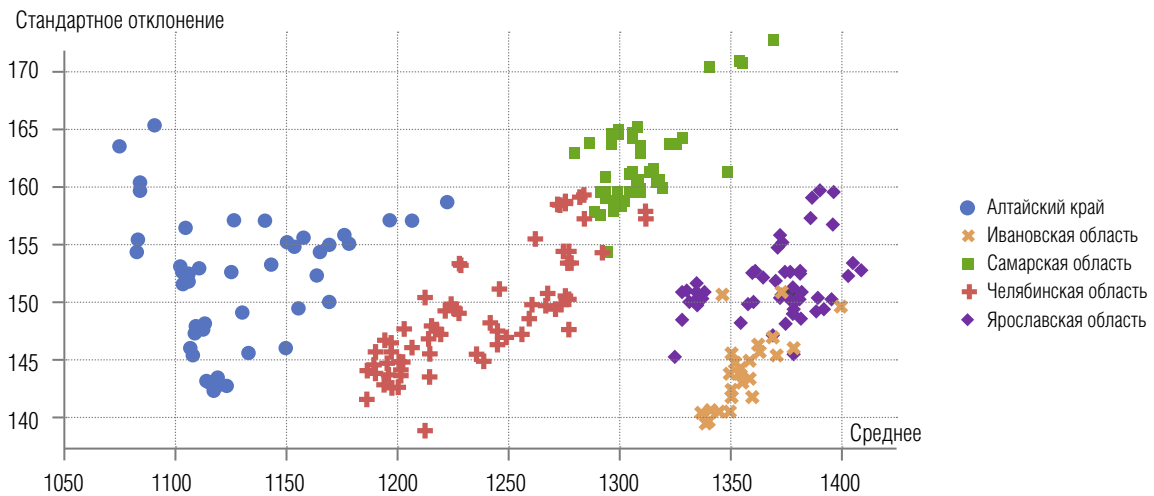


Рис. 2. Кластеризация статистических характеристик узлов внутри одного региона.

гионов России, рост цен наблюдается в третьем квартале, что связано с динамикой общероссийского производственного цикла и началом отопительного периода (более детально причины годовой сезонности проанализированы в работе Прохоровой и др. [28]). К исключениям относятся Красноярский край и Иркутская область (см. рис. 3), в которых, напротив, наблюдается снижение цен в 3 квартале, что может быть связано с доступностью в регионах относительно дешевой электроэнергии от ГЭС.

3. Поведение цен внутри недели на уровне регионов имеет некоторую относительно постоянную периодичность: в большинстве регионов в понедельник и пятницу цены находятся на максимальном уровне, в субботу — на минимальном. Исключения составляют Иркутская область и Республика Бурятия, в которых цены в субботу в среднем наи-

более высокие. Однако, в течение рабочей недели динамика может отличаться от региона к региону. В дни государственных праздников наблюдается значительное понижение цен во всех регионах, за исключением ряда регионов СКФО (рис. 4).

4. Нелинейная зависимость от температуры. В упомянутом выше исследовании Прохоровой и др. [28] отмечалась необходимость учитывать, как устойчивые изменения динамики температуры в течение года, влияющие на годовую сезонность в потреблении электричества, так и погодные факторы, которые учитывают отклонения от нормы.

3. Методология исследования

Анализ ценовой динамики в узлах энергосистемы РФ позволил выявить наличие пространственной автокорреляции. Статистические характеристики

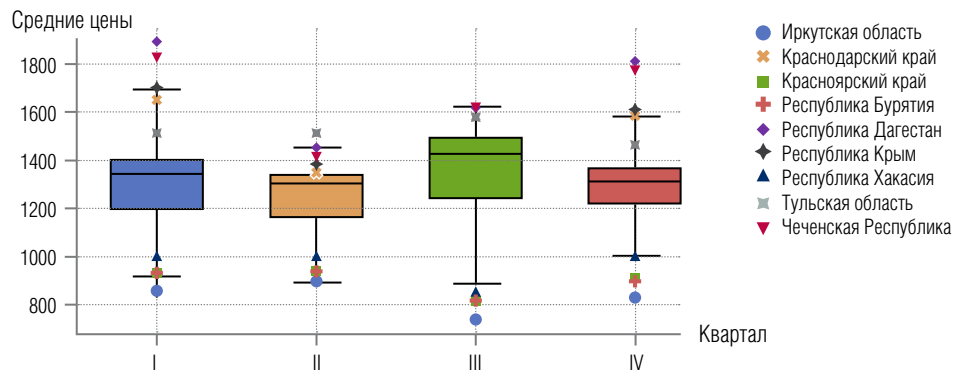


Рис. 3. Диаграмма размаха средних цен РСВ в зависимости от квартала (расчет проведен для всех регионов в выборке, выбросы показаны для отдельных регионов РФ).

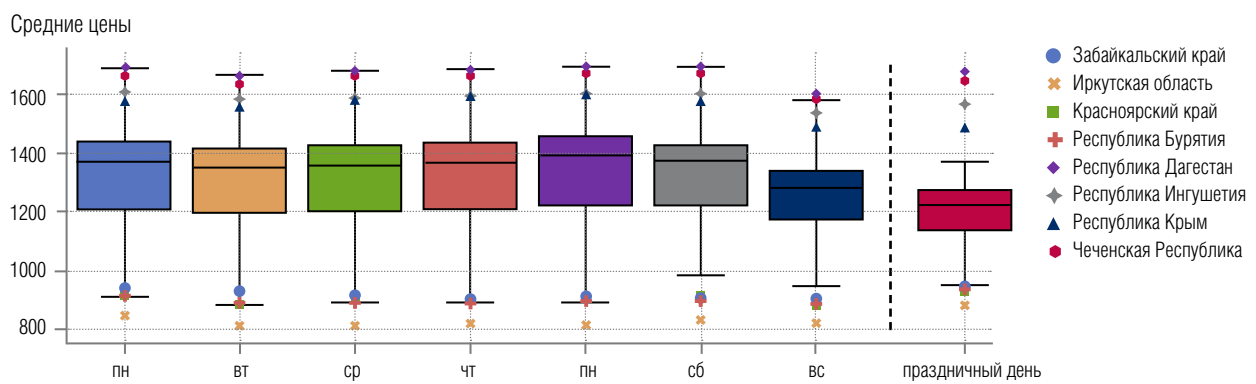


Рис. 4. Диаграмма размаха усредненных цен РСВ по дням недели и праздникам (расчет проведен для всех регионов в выборке, выбросы показаны для отдельных регионов РФ).

цен по мере удаления узлов все более отличаются, что связано с изменяющимися условиями спроса и предложения. По этой причине прогнозная модель может быть специфичной для:

а) узла и учитывать пространственный лаг (пространственную автокорреляцию),

б) региона и иметь множественный вывод – прогноз одновременно во всех узлах региональной энергосистемы, где формат входных данных предполагает учет географической связности.

Во втором случае (вариант б) для получения модели с высокой обобщающей способностью необходимо значительное количество наблюдений во времени. К сожалению, ограничения доступной ретроспективы данных на сайте администратора торговой системы (АТС) не позволяют обучить подобную модель, поэтому в своем исследовании мы остановимся на первом варианте (вариант а), предполагающем построение универсальной архитектуры нейросети, которая при условии обучения на специфичных для узла данных позволит прогнозировать

цены для любого отдельно взятого узла энергетической системы без существенной дисперсии ошибки.

С учетом приведенных выше особенностей в данных в числе объясняющих переменных будут использоваться следующие факторы:

- ◆ лаговые значения цены на электроэнергию в целевом узле;
- ◆ лаговые значения цены на электроэнергию трех ближайших узлов в рамках региональной сети;
- ◆ фиктивные переменные (дамми) на выходные и праздничные дни;
- ◆ среднесуточная температура в регионе;
- ◆ отклонение температуры от скользящего среднего.

Важно отметить, что модель для узла может включать и специфичные для местности факторы, к примеру, водность рек, что актуально для регионов Урала и Сибири [16], мощности промышленных потребителей и другие. Однако отсутствие статистики в необходимом разрезе не предоставляет этой возможности.

Как было отмечено выше, на первом шаге моделирования формируется статичная рекуррентная нейронная сеть, которая оптимизируется, а ее устойчивость проверяется на случайных узлах энергосистемы. На втором шаге моделирования формируется GAN на основе имеющейся архитектуры статичной рекуррентной нейронной сети, которая используется в качестве генератора. После оптимизации архитектуры дискриминатора и гиперпараметров GAN сравнивается точность результатов прогнозирования построенной модели с альтернативными бенчмарками.

Первоначально все модели параллельно обучаются на 5 случайных узлах в каждом регионе⁸. Таким образом, каждая модель последовательно обучается на 330 узлах⁹, а средняя ошибка и ее дисперсия, полученные на тестовых данных, используются для сравнения качества моделей. Соответственно, предобработка данных для всех узлов унифицирована: производится разделение выборки на обучающую, валидационную и тестовую в соотношении 80%, 10% и 10%; стандартизация данных и тензорное преобразование в формате скользящих окон из исходных временных рядов (размерность временного окна является гиперпараметром).

3.1. Построение статичной модели генератора (первый шаг моделирования)

Статичная модель представляет собой двухслойную рекуррентную нейронную сеть GRU¹⁰ с 55 ячейками в первом слое и 20 во втором¹¹. Первый слой на выходе сохраняет размерность данных во времени для следующего рекуррентного слоя. Второй слой в качестве выходных данных передает вектор размерности 3. Таким образом, исходную матрицу входных данных размерности (7 x 8), где 7 – это размер временного окна, 8 – число объясняющих переменных, модель отображает в пространство размерности (1 x 3). Сумма элементов данного

вектора представляет собой прогноз цены на электроэнергию на одни сутки вперед (элементы вектора суммируются с учетом весов, значения которых подбираются в ходе обучения модели). Такие гиперпараметры модели, как размер временного окна, определяющего число лагов во времени для всех переменных, и количество элементов обучающей выборки, используемых для расчета одной итерации градиентного спуска, были определены эмпирически в результате перебора, критерием выступала среднеквадратическая ошибка на тестовой выборке¹².

3.2. Построение GAN (второй шаг моделирования)

Генеративно-сопоставительная модель представляет собой последовательно соединенные генератор и дискриминатор (рис. 5). В качестве генератора используется ранее определенная модель рекуррентной нейронной сети, но в матрицу входных данных, как отдельный фактор, добавляется вектор шума, который препятствует переобучению генератора и позволяет добавить в вывод модели элемент случайности.

Архитектура дискриминатора представляет собой сверточную нейронную сеть с одним скрытым слоем, который принимает на вход размерность 7. Первые 6 элементов вектора – это фактические значения цен в ретроспективе, а последний элемент – прогноз генератора или фактическая цена, в зависимости от подающегося на вход шаблона. Дискриминатор учится классифицировать входящие данные: (X, Y) – «Истина» или 1 и $(X, G(X))$ – «Ложь» или 0. Дискриминатор обучается с использованием функции ошибки (уравнение 2), в основе которой находится кросс-энтропия (bce – binary cross entropy), уравнение 3:

$$LD = l_{bce}(D(X, Y), 1) + l_{bce}(D(X, G(X)), 0), \quad (2)$$

⁸ Учитывая, что каждый регион имеет различное количество узлов с полным покрытием данными (до 546 узлов) в целях экономии вычислительных ресурсов для тестирования модели используется ограниченное количество узлов.

⁹ 5 узлов · 66 регионов.

¹⁰ Gated Recurrent Units – вид рекуррентной нейронной сети, разработанной для моделирования временных последовательностей.

¹¹ Количество нейронов в каждом слое, вид функции активации и шаг градиентного спуска были определены в результате перебора на заданном множестве с использованием пакета для оптимизации нейронных сетей «KerasTuner».

¹² Эксперименты по определению базовой архитектуры проводились на данных случайно выбранного узла. В дальнейшем модель была дообучена для каждого узла индивидуально.

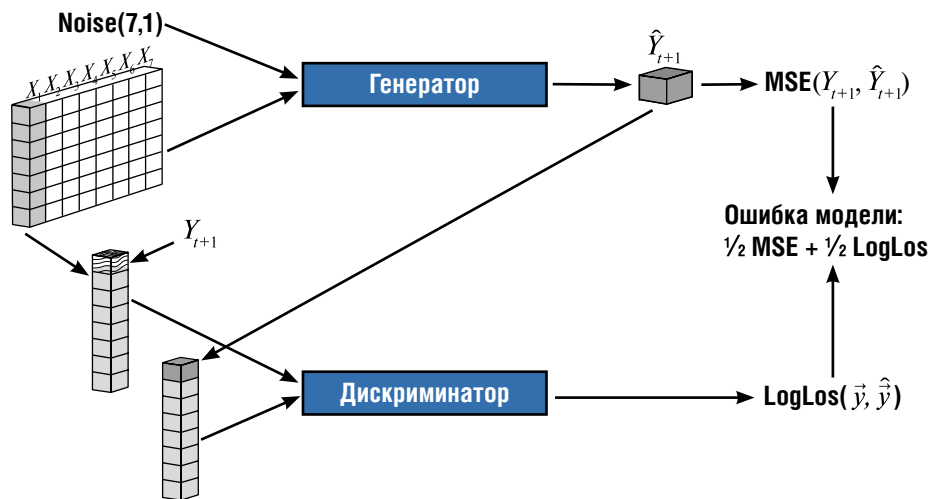


Рис. 5. Устройство генеративно-сопоставительной модели, используемой для построения прогноза поузловых цен на электроэнергию на РСВ.

$$l_{bce} = -(y \cdot \log(p) + (1 - y) \cdot \log(1 - p)). \quad (3)$$

Для обучения генератора применяется двухкомпонентная функция ошибки, состоящая из кросс-энтропии и среднеквадратической ошибки (MSE – mean squared error), уравнение 4.

$$L_G = \lambda_1 l_{bce}(D(X, G(X)), 1) + \lambda_2 \sum (Y - G(X))^2. \quad (4)$$

Метрика кросс-энтропии рассчитывается с присвоением вектору с прогнозным значением цены метки «Истина». Это необходимо для введения в «заблуждение» дискриминатора и позволяет генератору подбирать веса в процессе обучения, осложняющие задачу обучения дискриминатора на последующей итерации, что и вводит в процесс обучения элемент «сопоставительности». Важно заметить, что теоретически добавление MSE компоненты в функцию ошибки не является строго необходимым [29]. Мы добавляем MSE для более быстрой сходимости генератора и сокращения время на обучение.

Как указывалось выше, обучение дискриминатора и генератора происходит поочередно до достижения заданной точности прогноза цены, что и обеспечивает решение исходной минимаксной задачи. Укрупненное описание всего процесса обучения можно представить в виде псевдокода (таблица 1).

Более подробную информацию об алгоритме сопоставительного обучения нейронной сети можно найти в работах [30, 31].

Таким образом, дискриминатор в процессе обучения учится отличать сгенерированные образцы, все более похожие на реальные данные, от последних. Генератор, в свою очередь, на очередной итерации обучения стремится улучшить свой прогноз так, чтоб вероятностный выход модели был ближе к «Истине» при фиксированных весах дискриминатора. Такой подход позволяет изменять траекторию градиента в процессе обучения и приходиться к иному оптимуму в сравнении со статической моделью, где используется только среднеквадратическая ошибка [32].

Результаты работы модели GAN сравниваются с «эталонными» альтернативными моделями: авторегрессия ARIMA, базовая рекуррентная нейронная сеть, двухслойная сверточная нейронная сеть (convolutional neural network, CNN). Архитектура последней основана на модели, предложенной в работе [33].

4. Эмпирические результаты

Мы обучали дискриминатор в пять раз больше, чем генератор, во время каждой итерации обучения. Это позволяет дискриминатору несколько опережать генератор в точности и отличать сгенерированные данные от реальных. На рисунке 6 приведен процесс обучения генеративно-сопоставительной модели для некоторого случайного узла.

На левой оси приведена общая ошибка модели, которая представляет собой взвешенную сумму

Таблица 1.

**Укрупненный алгоритм обучения
генеративно-состязательной нейронной сети**

Задание условий: задается шаг градиентного спуска ρ_D, ρ_G ; задаются параметры весов функций ошибок для генеративной модели λ_1, λ_2 ; инициализируются веса случайным образом в моделях дискриминатора и генератора.

While (выполняется пока алгоритм не сойдется до заданной точности):

Обучение дискриминатора (D):

Получение M наборов данных из тренировочной выборки X^{train} :

$$(X, Y) = (X^1, Y^1), \dots, (X^m, Y^m) \subset X^{train}$$

Выполнение шага стохастического градиентного спуска и обновление весов D при фиксированных весах G :

$$W_D = W_d - \rho_D \sum_{i=1}^M \frac{\partial L_D(X^{(i)}, Y^{(i)})}{\partial W_D}$$

Обучение генератора (G):

Получение M новых наборов данных из X^{train} :

$$(X, Y) = (X^1, Y^1), \dots, (X^m, Y^m) \subset X^{train}$$

Выполнение шага стохастического градиентного спуска и обновление весов G при фиксированных весах D :

$$W_G = W_G - \rho_G \sum_{i=1}^M \frac{\partial L_G(X^{(i)}, Y^{(i)})}{\partial W_G}$$

End While

ошибки дискриминатора и генератора (mse+logloss). На правой оси приведено значение ошибки дискриминатора на каждой итерации обучения (logloss).

Для того, чтоб проверить устойчивость предложенной модели был проведен следующий эксперимент: в каждом регионе случайным образом отбиралось по 5 узлов энергосистемы. Единственный критерий отбора – это отсутствие пропусков в данных на анализируемом временном горизонте. Далее для всех узлов в соответствии с унифицированной процедурой формировался набор объясняющих переменных, проводилась предобработка данных, а затем сохраненная ранее генеративная модель дообучалась на данных каждого из 5 узлов. По итогам эксперимента рассчитывалась ошибка прогноза цены на электроэнергию на тестовой выборке (параллельно обучение и расчет ошибки проводился для статичной модели генератора и других бенчмарков).

Результаты эксперимента показали, что генеративно-состязательная модель демонстрирует наименьшую среднюю ошибку на тестовой выборке и минимальный разброс значений ошибки на уровне отдельных узлов энергосистемы. В таблице 2 приведены усредненные ошибки на тестовой выборке рассматриваемых моделей и их стандартные отклонения.

В работе [25] авторы связывают превосходство GAN с возможностью обеспечивать требуемый градиент для оптимизации генератора во время обучения: градиент, направляемый дискриминатором, позволяет достичь более широкого¹³ минимума, чем, при прочих равных, достигается в результате оптимизации статичного аналога генеративной модели. Собственные эмпирические результаты также свидетельствуют в пользу данного тезиса, поскольку при околонулевой

¹³ Дмитрий Ветров [36] выдвинул гипотезу о наличии в ландшафте функции потерь нейронной сети минимумов различной кривизны: узких и широких. В узких минимумах, при нулевой ошибке на тренировочной выборке, ошибка на тестовой заметно выше; в широких – ошибки мало отличимы.

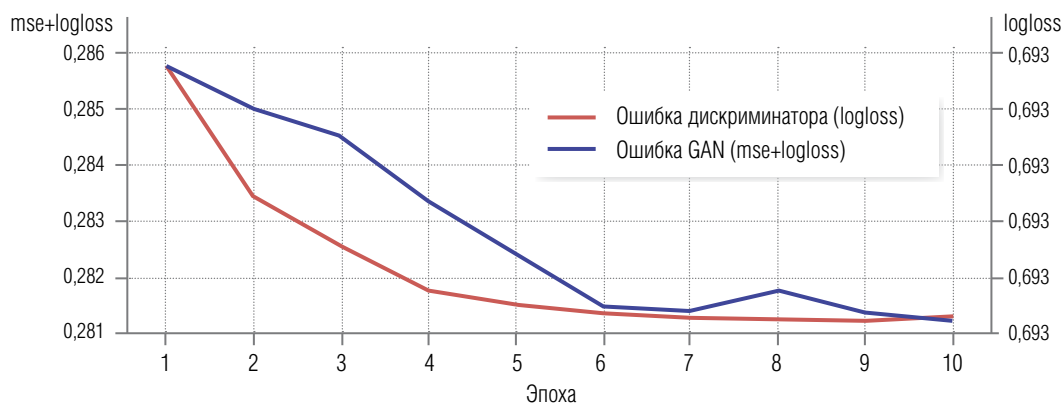


Рис. 6. Значение ошибок генеративно-сопоставительной модели в процессе обучения на тренировочной выборке.

Таблица 2.

Сравнительная таблица усредненных ошибок и их разброса при применении модели к пяти случайным узлам в анализируемых регионах

Показатель		Ошибка на тестовой выборке			
		GAN	GRU	CNN	ARIMA
MAE	mean	0,0497	0,062	0,0681	0,0724
	std	0,0002	0,0009	0,0063	0,0092
MAPE	mean	0,092	0,0973	0,1055	0,1114
	std	0,0005	0,0018	0,0113	0,0242
MSE	mean	0,0046	0,0059	0,0073	0,0082
	std	0	0,0002	0,0013	0,0034

ошибке на тренировочном множестве, достигаемой одновременно статичной моделью генератора и GAN, ошибка последней на тестовой множестве несколько ниже.

Усредненные ошибки модели GAN и их стандартные отклонения визуализированы¹⁴ на рис. 7: градации цвета на условной карте регионов соответствует значению средней ошибки, посчитанной для пяти случайно выбранных узлов в соответствующем регионе РФ (см. левую часть рисунка 7) и дисперсии средних ошибок (см. правую часть рисунка 7).

Наилучшие результаты модель демонстрирует в большей части регионов ЦФО, СЗФО и ПФО. В свою очередь, наибольшую среднюю ошибку модель демонстрирует в таких регионах, как Республика Татарстан, Башкирия, ряде регионов Сибири и Кавказа. Нестабильность модели, проявляющаяся

в высокой дисперсии средней ошибки, также проявляется преимущественно в перечисленных регионах. Неудовлетворительная работа модели в данных регионах, как представляется, обусловлена недостаточностью набора объясняющих факторов, который во всех случаях являлся стандартным. Возможно, для приграничных регионов некоторую роль в ценообразовании на рынке на сутки вперед играют энергетически мосты с другими странами, связанными с отечественной энергетической системой; для центральных регионов Сибири важным фактором является водность рек, поскольку значительная доля генерации приходится на ГЭС.

Также важно отметить тот факт, что на тестовую часть выборки (с 19.08.2022 по 31.12.2022) приходится период с явным структурным сдвигом — резким ростом цены на электроэнергию в конце

¹⁴ Для визуализации использовалось приложение QGIS — Географическая информационная система (ГИС) с открыты исходным кодом.

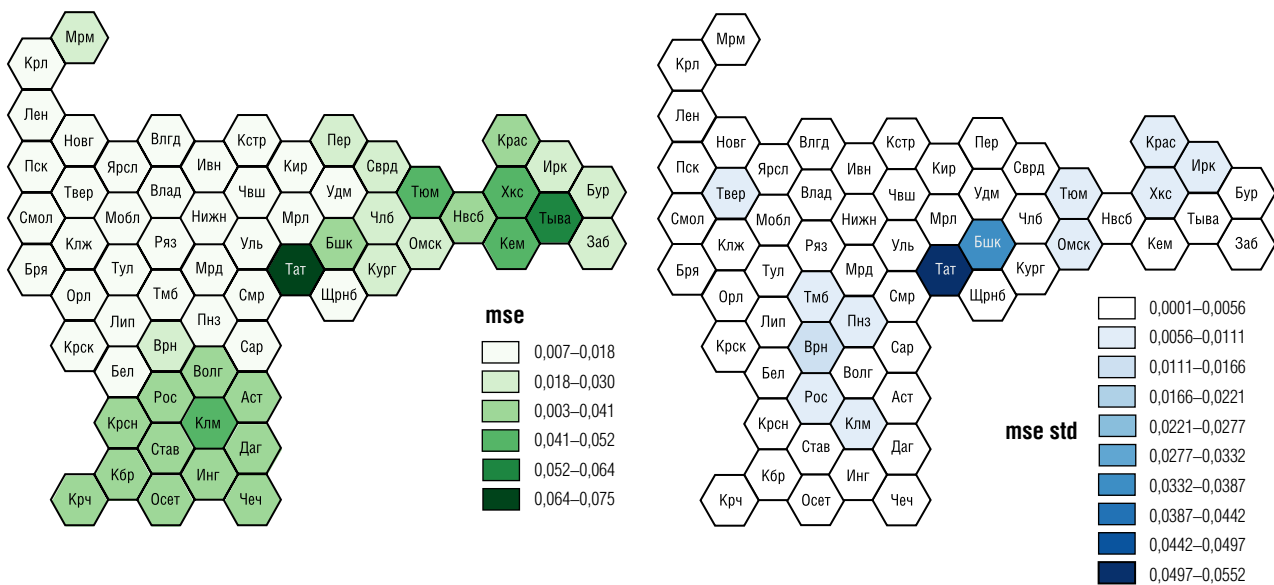


Рис. 7. Средняя ошибка генеративно-состязательной модели на тестовой выборке (левый график) и ее дисперсия (правый график) для пяти случайных узлов в регионе.

2022 в таких регионах, как Иркутская, Тюменская и Томская область, Красноярский край, Ставропольский край, Краснодарский край¹⁵. Резкий рост тренда цены в данном случае был обусловлен конъюнктурными факторами, влияние которых не могло присутствовать в обучающей выборке.

В завершение эмпирической части исследования проверялась гипотеза о групповом равенстве дисперсий Левена для полученных среднеквадратических ошибок на случайно выбранных узлах энергосистемы внутри отдельного взятого региона, где нулевая гипотеза состоит в том, что «все подвыборки имеют равные дисперсии». При исключении выбросов (12 регионов с наибольшим стандартным отклонением ошибки), для оставшихся 54 регионов нулевая гипотеза теста не отвергается¹⁶. Таким образом, для большинства регионов предложенная архитектура генеративной модели является универсальной и способна обеспечить относительно низкую ошибку на тестовой выборке для различных узлов энергосистемы страны.

Заключение

Результаты проведенного исследования демонстрируют, что генеративно-состязательная

модель с заданной универсальной архитектурой (не изменяющейся при переходе от региона к региону РФ) и ограниченным числом объясняющих факторов при условии дообучения на данных, специфичных для целевого узла может использоваться для прогнозирования цен в узлах рынка на сутки вперед без существенных отклонений в точности для 54 из 66 рассматриваемых регионов России. В состав нейросетевой модели целесообразно включать следующий набор переменных: исторические значения цен в целевом и географически близких к нему узлах энергосистемы (поузловые цены на оптовом рынке электроэнергии коррелированы как пространственно, так и во времени), температура окружающей среды и сезонные факторы.

Предложенная генеративно-состязательная модель позволила на 22% уменьшить среднеквадратическую ошибку на тестовой выборке статичной модели генератора, имеющего в своей основе рекуррентную нейронную сеть, а также превзошла по качеству альтернативные (эталонные) модели-бенчмарки: сверточную нейронную сеть и авторегрессионную линейную модель (ARIMA). ■

¹⁵ Например, в Иркутской области, среднесуточные значения тренда цены в декабре 2022 г. были в среднем на 68% выше, чем годом ранее.

¹⁶ Значение статистики теста составляет 1,29.

Литература

1. Акопов А.С., Бекларян Г.Л. Анализ эффективности регулирующей политики государства с помощью региональной модели CGE поведения естественных монополий (на примере электроэнергетики) // *Экономическая наука современной России*. 2005. № 4 (31). С. 130–139.
2. Акопов А.С., Бекларян А.Л., Хачатрян Н.К., Фомин А.В. Система прогнозирования динамики добычи нефти с использованием имитационного моделирования // *Информационные технологии*. 2017. Т. 23. № 6. С. 431–436.
3. Макаров В.Л., Бахтизин А.Р., Бекларян Г.Л., Акопов А.С. Цифровой завод: методы дискретно-событийного моделирования и оптимизации производственных характеристик // *Бизнес-информатика*. 2021. Т. 15. № 2. С. 7–20. <https://doi.org/10.17323/2587-814X.2021.2.7.20>
4. Хачатрян Н.К., Кравченко Т.К., Акопов А.С., Уварова О.М. Прогнозирование основных показателей фондового рынка России авторегрессионными моделями с распределенными лагами // *Аудит и финансовый анализ*. 2016. № 3. С. 128–133.
5. Gonzalez J.P., San Roque A.M., Perez E.A. Forecasting functional time series with a new Hilbertian ARMAX model: Application to electricity price forecasting // *IEEE Transactions on Power Systems*. 2017. Vol. 33. No. 1. P. 545–556. <https://doi.org/10.1109/TPWRS.2017>
6. Zhao Z., Wang C., Nokleby M., Miller C.J. Improving short-term electricity price forecasting using day-ahead LMP with ARIMA models // *2017 IEEE Power & Energy Society General Meeting*. 2017. P. 1–5. <https://doi.org/10.1109/PESGM.2017.8274124>
7. Vehviläinen I., Pyykkönen T. Stochastic factor model for electricity spot price – the case of the Nordic market // *Energy Economics*. 2005. No. 2. P. 351–367.
8. Howison S., Coulon M. Stochastic behaviour of the electricity bid stack: from fundamental drivers to power prices // *The Journal of Energy Markets*. 2009. P. 29–69.
9. Weron R. Electricity price forecasting: A review of the state-of-the-art with a look into the future // *International Journal of Forecasting*. 2014. Vol. 30. No. 4. P. 1030–1081. <https://doi.org/10.1016/j.ijforecast.2014.08.008>
10. Garcia R.C., Contreras J., Akkeren M., Garcia J.B. A GARCH forecasting model to predict day-ahead electricity prices // *IEEE Transactions on Power Systems*. 2005. Vol. 20. No. 2. P. 867–874. <https://doi.org/10.1109/TPWRS.2005.846044>
11. Lynch C., O’Leary C., Sandareshan P.G., Akin Y. Experimental analysis of GBM to expand the time horizon of Irish electricity price forecasts // *Energies*. 2021. Vol. 14. No. 22. 7587. <https://doi.org/10.3390/en14227587>
12. Deng Z., Liu C., Zhu Z. Inter-hours rolling scheduling of behind-the-meter storage operating systems using electricity price forecasting based on deep convolutional neural network // *International Journal of Electrical Power & Energy Systems*. 2021. Vol. 125. 106499. <https://doi.org/10.1016/j.ijepes.2020.106499>
13. Zhang J., Tan Z., Yang S. Day-ahead electricity price forecasting by a new hybrid method // *Computers and Industrial Engineering*. 2012. Vol. 63. P. 695–701. <https://doi.org/10.1016/j.cie.2012.03.016>
14. Kim T.Y., Cho S.B. Predicting residential energy consumption using CNN-LSTM neural networks // *Energy*. 2019. Vol. 182. P. 72–81. <https://doi.org/10.1016/j.energy.2019.05.230>
15. Марьясин О.Ю., Лукашов А.И. Прогнозирование свободных цен на электроэнергию с использованием нейронных сетей // *Сборник научных трудов XXII международной научно-технической конференции «Нейроинформатика–2020»*. 2020. С. 294–303.
16. Золотова И.Ю., Дворкин В.В. Краткосрочное прогнозирование цен на российском оптовом рынке электроэнергии на основе нейронных сетей // *Проблемы прогнозирования*. 2017. № 6. С. 47–57.
17. Guo X., Zhao Q., Zheng D., Ning Y., Gao, Y. A short-term load forecasting model of multi-scale CNN-LSTM hybrid neural network considering the real-time electricity price // *Energy Reports*. 2020. Vol. 6. P. 1046–1053. <https://doi.org/10.1016/j.egy.2020.11.078>
18. Wu W., Huang F., Kao Y., Chen Z., Wu Q. Prediction method of multiple related time series based on generative adversarial networks // *Information*. 2021. Vol. 12. No. 2. 55. <https://doi.org/10.3390/info12020055>
19. Николенко С., Кадурин А., Архангельская Е. Глубокое обучение. Санкт-Петербург: Питер, 2019.
20. Goodfellow I., Pouget-Abadie I., Mirza M., Xu B., Warde-Farley D., Ozair S., Courville A., Bengio Y. Generative adversarial networks // *arXiv:1406.2661*. 2014. <https://doi.org/10.48550/arXiv.1406.2661>
21. Zhang H., Zhang H., Yu Koh J., Baldridge J., Lee H., Yang Y. Cross-modal contrastive learning for text-to-image generation // *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2021. P. 833–842.
22. Kadurin A., Nikolenko S., Khrabrov K., Aliper A., Zhavoronkov A. DruGAN: An advanced generative adversarial autoencoder model for de novo generation of new molecules with desired molecular properties in silico // *Molecular Pharmaceutics*. 2017. Vol. 14. No. 9. P. 3098–3104. <https://doi.org/10.1021/acs.molpharmaceut.7b00346>
23. Yoon J., Jarrett D., van der Schaar M. Time-series generative adversarial networks // *Neural Information Processing Systems*. 2019.

24. Koochali A., Dengel A., Ahmed S. If you like it, GAN it – Probabilistic multivariate times series forecast with GAN // Engineering Proceedings. 2021. Vol. 5. No. 1. 40. <https://doi.org/10.3390/engproc2021005040>
25. Zhang Z., Wu M. Real-time locational marginal price forecasting using generative adversarial network // 2020 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids. 2020. P. 1–6. <https://doi.org/10.1109/SmartGridComm47815.2020.9302938>
26. Kabir H.M., Khosravi A., Nahavardini S., Kavousi-Fard A. Partial adversarial training for neural network-based uncertainty quantification // IEEE Transactions on Emerging Topics in Computational Intelligence. 2019. Vol. 5. No. 4. P. 595–606. <https://doi.org/10.1109/TETCI.2019.2936546>
27. Отчёт о равновесных ценах в наиболее крупных узлах расчётной модели. Администратор Торговой Системы. [Электронный ресурс] https://www.atsenergo.ru/nreport?rname=big_nodes_prices_pub®ion=eur (дата обращения 02.03.2023).
28. Порохова Н.В., Рудаков Е.Н., Саакян Ю.З. Альтернативные методы мониторинга динамики промышленного производства // Проблемы прогнозирования. 2012. № 4. С. 36–53.
29. Zhang K., Zhong G., Dong J., Wang S., Wang Y. Stock market prediction based on generative adversarial network // Procedia Computer Science. 2019. Vol. 147. P. 400–406. <https://doi.org/10.1016/j.procs.2019.01.256>
30. Mathieu M., Couprie C., LeCun Y. Deep multi-scale video prediction beyond mean square error // arXiv:1511.05440. 2015. <https://doi.org/10.48550/arXiv.1511.05440>
31. Radford A., Metz L., Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks // arXiv:1511.06434. 2015. <https://doi.org/10.48550/arXiv.1511.06434>
32. Staffini A. Stock price forecasting by a deep convolutional generative adversarial network // Frontiers in Artificial Intelligence. 2022. Vol. 5. 837596. <https://doi.org/10.3389/frai.2022.837596>
33. Liu S., Ji H., Wang M.C. Nonpooling convolutional neural network forecasting for seasonal time series with trends. IEEE Transactions on Neural Networks and Learning Systems. 2020. Vol. 31. No. 8. P. 2879–2888. <https://doi.org/10.1109/tnnls.2019.2934110>
34. Izmailov P., Maddox J., Kirichenko P., Garipov T., Vetrov D., Gordon A. Subspace inference for Bayesian deep learning. Proceedings of the 35th Uncertainty in Artificial Intelligence Conference. Proceedings of Machine Learning Research. 2020. Vol. 115. P. 1169–1179.

Об авторах

Каукин Андрей Сергеевич

кандидат экономических наук;

заведующий лабораторией отраслевых рынков и инфраструктуры, Институт экономической политики имени Е.Т. Гайдара, Россия, 125993, Москва, Газетный пер., д. 3-5, стр. 1;

заведующий лабораторией системного анализа отраслевых рынков, Институт прикладных экономических исследований Российской академии народного хозяйства и государственной службы при Президенте Российской Федерации, Россия, 119571, г. Москва, пр. Вернадского, д. 82;

E-mail: kaukin@ier.ru

ORCID: 0000-0003-2892-5278

Павлов Павел Николаевич

кандидат экономических наук;

старший научный сотрудник, лаборатория системного анализа отраслевых рынков, Институт прикладных экономических исследований Российской академии народного хозяйства и государственной службы при Президенте Российской Федерации, Россия, 119571, г. Москва, пр. Вернадского, д. 82;

E-mail: pavlov@ranepa.ru

ORCID: 0000-0001-6200-6481

Косарев Владимир Сергеевич

научный сотрудник, лаборатория системного анализа отраслевых рынков, Институт прикладных экономических исследований Российской академии народного хозяйства и государственной службы при Президенте Российской Федерации, Россия, 119571, г. Москва, пр. Вернадского, д. 82;

E-mail: kosarev-vs@ranepa.ru

ORCID: 0000-0001-5435-9076

Short-term forecasting of electricity prices using generative neural networks

Andrej S. Kaukin

E-mail: kaukin@iep.ru

Pavel N. Pavlov

E-mail: pavlov@ranepa.ru

Vladimir S. Kosarev

E-mail: kosarev-vs@ranepa.ru

Russian Presidential Academy of National Economy and Public Administration

Address: 82, Vernadskogo Prospect, Moscow 119571, Russia

Abstract

This article studies the predictive abilities of the generative-adversarial neural network approach in relation to time series using the example of price forecasting for the nodes of the Russian free electricity market for the day ahead. As a result of a series of experiments, we came to the conclusion that a generative adversarial network, consisting of two models (generator and discriminator), allows one to achieve a minimum of the error function with a greater generalizing ability than, all other things being equal, is achieved as a result of optimizing the static analogue of the generative model – recurrent neural network. Our own empirical results show that with a near-zero mean square error on the training set, which is demonstrated simultaneously by the recurrent and generative models, the error of the latter on the test set is lower. The adversarial approach also outperformed alternative reference models in out-of-sample forecasting accuracy: a convolutional neural network adapted for time series forecasting and an autoregressive linear model. Application of the proposed approach has shown that a generative-adversarial model with a given universal architecture and a limited number of explanatory factors, subject to additional training on data specific to the target node of the power system, can be used to predict prices in market nodes for the day ahead without significant deviations.

Keywords: electricity market, day-ahead market, time series, generative neural network, recurrent neural network

Citation: Kaukin A.S., Pavlov P.N., Kosarev V.S. (2023) Short-term forecasting of electricity prices using generative neural networks. *Business Informatics*, vol. 17, no. 3, pp. 7–23. DOI: 10.17323/2587-814X.2023.3.7.23

References

1. Akopov A.S., Beklarjan G.L. (2005) Analysis of the effectiveness of the regulatory policy of the state using the regional CGE model of the behavior of natural monopolies (on the example of the electric power industry). *Jekonomicheskaja nauka sovremennoj Rossii*, no. 4(31), pp. 130–139 (in Russian).
2. Akopov A.S., Beklarjan A.L., Hachatrjan N.K., Fomin A.V. (2017) Oil production dynamics forecasting system using simulation modeling. *Informacionnye tehnologii*, vol. 23, no. 6, pp. 431–436 (in Russian).

3. Makarov V.L., Bahtizin A.R., Beklarjan G.L., Akopov A.S. (2021) Digital plant: methods of discrete-event modeling and optimization of production characteristics. *Business Informatics*, vol. 15, no. 2, pp. 7–20. <https://doi.org/10.17323/2587-814X.2021.2.7.20>
4. Hachatryan N.K., Kravchenko T.K., Akopov A.S., Uvarova O.M. (2016) Forecasting of the main indicators of the Russian stock market by autoregressive models with distributed lags. *Audit and financial analysis*, vol. 3, pp. 128–133 (in Russian).
5. Gonzalez J.P., San Roque A.M., Perez E.A. (2017) Forecasting functional time series with a new Hilbertian ARMAX model: Application to electricity price forecasting. *IEEE Transactions on Power Systems*, vol. 33, no. 1, pp. 545–556. <https://doi.org/10.1109/TPWRS.2017>
6. Zhao Z., Wang C., Nokleby M., Miller C.J. (2017) Improving short-term electricity price forecasting using day-ahead LMP with ARIMA models. *2017 IEEE Power & Energy Society General Meeting*, pp. 1–5. <https://doi.org/10.1109/PESGM.2017.8274124>
7. Vehviläinen I., Pyykkönen T. (2005) Stochastic factor model for electricity spot price – the case of the Nordic market. *Energy Economics*, no. 2, pp. 351–367.
8. Howison S., Coulon M. (2009) Stochastic behaviour of the electricity bid stack: from fundamental drivers to power prices. *The Journal of Energy Markets*, pp. 29–69.
9. Weron R. (2014) Electricity price forecasting: A review of the state-of-the-art with a look into the future. *International Journal of Forecasting*, vol. 30, no. 4, pp. 1030–1081. <https://doi.org/10.1016/j.ijforecast.2014.08.008>
10. Garcia R.C., Contreras J., Akkeren M., Garcia J.B. (2005) A GARCH forecasting model to predict day-ahead electricity prices. *IEEE Transactions on Power Systems*, vol. 20, no. 2, pp. 867–874. <https://doi.org/10.1109/TPWRS.2005.846044>
11. Lynch C., O’Leary C., Sandareshan P.G., Akin Y. (2021) Experimental analysis of GBM to expand the time horizon of Irish electricity price forecasts. *Energies*, vol. 14, no. 22, 7587. <https://doi.org/10.3390/en14227587>
12. Deng Z., Liu C., Zhu Z. (2021) Inter-hours rolling scheduling of behind-the-meter storage operating systems using electricity price forecasting based on deep convolutional neural network. *International Journal of Electrical Power & Energy Systems*, vol. 125, 106499. <https://doi.org/10.1016/j.ijepes.2020.106499>
13. Zhang J., Tan Z., Yang S. (2012) Day-ahead electricity price forecasting by a new hybrid method. *Computers and Industrial Engineering*, vol. 63, pp. 695–701. <https://doi.org/10.1016/j.cie.2012.03.016>
14. Kim T.Y., Cho S.B. (2019) Predicting residential energy consumption using CNN-LSTM neural networks. *Energy*, vol. 182, pp. 72–81. <https://doi.org/10.1016/j.energy.2019.05.230>
15. Mar’yasyn O.Yu., Lukashov A.I. (2020) Prediction of free electricity prices using neural networks. *Neiroinformatika-2020. Sbornik nauchnykh trudov. XXII mezhdunarodnaya nauchno-tehnicheskaya konferentsiya*, pp. 294–303 (in Russian).
16. Zolotova I.Yu., Dvorkin V.V. (2017) Short-term forecasting of prices in the Russian wholesale electricity market based on neural networks. *Problemy prognozirovaniya*, no. 6, pp. 47–57 (in Russian).
17. Guo X., Zhao Q., Zheng D., Ning Y., Gao, Y. (2020) A short-term load forecasting model of multi-scale CNN-LSTM hybrid neural network considering the real-time electricity price. *Energy Reports*, vol. 6, pp. 1046–1053. <https://doi.org/10.1016/j.egy.2020.11.078>
18. Wu W., Huang F., Kao Y., Chen Z., Wu Q. (2021) Prediction method of multiple related time series based on generative adversarial networks. *Information*, vol. 12, no. 2, 55. <https://doi.org/10.3390/info12020055>
19. Nikolenko S., Kadurin A., Arkhangel’skaya E. (2019) *Glubokoe obuchenie*. Sankt-Peterburg, Piter (in Russian).
20. Goodfellow I., Pouget-Abadie I., Mirza M., Xu B., Warde-Farley D., Ozair S., Courville A., Bengio Y. (2014) Generative adversarial networks. *arXiv:1406.2661*. <https://doi.org/10.48550/arXiv.1406.2661>
21. Zhang H., Zhang H., Yu Koh J., Baldrige J., Lee H., Yang Y. (2021) Cross-modal contrastive learning for text-to-image generation. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 833–842.
22. Kadurin A., Nikolenko S., Khrabrov K., Aliper A., Zhavoronkov A. (2017) DruGAN: An advanced generative adversarial autoencoder model for de novo generation of new molecules with desired molecular properties in silico. *Molecular Pharmaceutics*, vol. 14, no. 9, pp. 3098–3104. <https://doi.org/10.1021/acs.molpharmaceut.7b00346>
23. Yoon J., Jarrett D., van der Schaar M. (2019) Time-series generative adversarial networks. *Neural Information Processing Systems*.
24. Koochali A., Dengel A., Ahmed S. (2021) If you like it, GAN it – Probabilistic multivariate times series forecast with GAN. *Engineering Proceedings*, vol. 5, no. 1, 40. <https://doi.org/10.3390/engproc2021005040>
25. Zhang Z., Wu M. (2020) Real-time locational marginal price forecasting using generative adversarial network. *2020 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids*, pp. 1–6. <https://doi.org/10.1109/SmartGridComm47815.2020.9302938>
26. Kabir H.M., Khosravi A., Nahavardini S., Kavousi-Fard A. (2019) Partial adversarial training for neural network-based uncertainty quantification. *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 5, no. 4, pp. 595–606. <https://doi.org/10.1109/TETCI.2019.2936546>
27. *Otchet o ravnovesnykh tsenakh v naibolee krupnykh uzlakh raschetnoi modeli*. Administrator Torgovoi Sistemy. Available at: https://www.atsenergo.ru/nreport?rname=big_nodes_prices_pub®ion=eur (accessed 02.03.2023) (in Russian).

28. Porokhova N.V., Rudakov E.N., Saakyan Yu.Z. (2009) Alternative methods for monitoring the dynamics of industrial production. *Problemy Prognozirovaniya*, no. 4, pp. 36–53 (in Russian).
29. Zhang K., Zhong G., Dong J., Wang S., Wang Y. (2019) Stock market prediction based on generative adversarial network. *Procedia Computer Science*, vol. 147, pp. 400–406. <https://doi.org/10.1016/j.procs.2019.01.256>
30. Mathieu M., Couprie C., LeCun Y. (2015) Deep multi-scale video prediction beyond mean square error. *arXiv:1511.05440*. <https://doi.org/10.48550/arXiv.1511.05440>
31. Radford A., Metz L., Chintala S. (2015) Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv:1511.06434*. <https://doi.org/10.48550/arXiv.1511.06434>
32. Staffini A. (2022) Stock price forecasting by a deep convolutional generative adversarial network. *Frontiers in Artificial Intelligence*, vol. 5, 837596. <https://doi.org/10.3389/frai.2022.837596>
33. Liu S., Ji H., Wang M.C. (2020) Nonpooling convolutional neural network forecasting for seasonal time series with trends. *IEEE Transactions on Neural Networks and Learning Systems*, vol. 31, no. 8, pp. 2879–2888. <https://doi.org/10.1109/tnnls.2019.2934110>
34. Izmailov P., Maddox J., Kirichenko P., Garipov T., Vetrov D., Gordon A. (2020) Subspace inference for Bayesian deep learning. *Proceedings of the 35th Uncertainty in Artificial Intelligence Conference. Proceedings of Machine Learning Research*, vol. 115, pp. 1169–1179.

About the authors

Andrej S. Kaukin

Cand. Sci. (Econ.);

Head of the Laboratory of Industry Markets and Infrastructure, Gaidar Institute for Economic Policy, 3–5, Gazetnyy Lane, Moscow 125993, Russia;

Head of the laboratory of System Analysis of Industrial Markets, Institute of Applied Economic Research, Russian Presidential Academy of National Economy and Public Administration, 82, Vernadskogo Prospect, Moscow 119571, Russia;

E-mail: kaukin@iep.ru

ORCID: 0000-0003-2892-5278

Pavel N. Pavlov

Cand. Sci. (Econ.);

Senior researcher, Laboratory of System Analysis of Industrial Markets, Institute of Applied Economic Research, Russian Presidential Academy of National Economy and Public Administration, 82, Vernadskogo Prospect, Moscow 119571, Russia;

E-mail: pavlov@ranepa.ru

ORCID: 0000-0001-6200-6481

Vladimir S. Kosarev

Research Fellow, Laboratory of System Analysis of Industrial Markets, Institute of Applied Economic Research, Russian Presidential Academy of National Economy and Public Administration, 82, Vernadskogo Prospect, Moscow 119571, Russia;

E-mail: kosarev-vs@ranepa.ru

ORCID: 0000-0001-5435-9076

DOI: 10.17323/2587-814X.2023.3.24.37

Влияние алгоритмизации и интерфейса на подготовку управленческих решений*

Р.Д. Гутгарц 

E-mail: gutgarc@gmail.com

Иркутский национальный исследовательский технический университет

Адрес: Россия, 664074, г. Иркутск, ул. Лермонтова, д. 83

Аннотация

В современных условиях принятие управленческих решений осуществляется при использовании автоматизированных систем под общим названием «системы поддержки принятия решений» (СППР). При их создании важно учитывать два ключевых момента. Первый – это алгоритмическая составляющая, отражающая логику работы системы в целом и отдельных ее частей. Второй – это интерфейс приложения, посредством которого пользователь с ним взаимодействует. Интерфейс представляет собой графическую интерпретацию алгоритмов, которые реализованы в рамках системы. Поэтому очень важно спроектировать и создать такую взаимосвязь между алгоритмом и интерфейсом, чтобы пользователю было максимально комфортно использовать СППР для решения своих текущих задач (ввод информации, ее обработку, представление и анализ для принятия решений). Таким образом, между интерфейсом и алгоритмом имеется прямо пропорциональная зависимость. И, несмотря на тот факт, что по данным аспектам имеется множество исследований, как теоретической, так и практической направленности, еще остаются вопросы, на которые следует обращать внимание в прикладном плане. Целью данного исследования является формулирование практических рекомендаций для предотвращения ввода некорректной информации в базу данных СППР и представления результатов в виде, удобном для ее анализа. Основные задачи работы – показать на примерах, какие ошибки могут способствовать вводу в базу данных недостоверной информации, а также каким наилучшим образом представить информацию на экране монитора в соответствии с психофизиологическими характеристиками человека, чтобы сократить время на ее анализ и принятие решений. Проанализированы типовые ошибки, возникающие в процессе ввода информации. Приведены рекомендации по визуализации аналитической информации на экране монитора.

* Статья опубликована при поддержке Программы НИУ ВШЭ «Университетское партнерство»

Ключевые слова: системы поддержки принятия решений, особенности алгоритмизации, взаимосвязь интерфейса и алгоритмов, представление информации, обработка ошибок при вводе информации

Цитирование: Гутгарц Р.Д. Влияние алгоритмизации и интерфейса на подготовку управленческих решений // Бизнес-информатика. 2023. Т. 17. № 3. С. 24–37. DOI: 10.17323/2587-814X.2023.3.24.37

Введение

Ключевые показатели в принятии управленческих решений заключаются в том, чтобы они были обоснованными и своевременными. Этого возможно добиться только при использовании соответствующей информации, удовлетворяющей всем своим классическим требованиям: достоверности, непротиворечивости, полноты, актуальности, ценности, понятности, доступности и пр.

Поэтому, с одной стороны, очень важно алгоритмически отслеживать и устранять все потенциально возможные ошибки на этапе начального ввода информации в базу данных, так как последующая обработка неверной информации неизбежно будет порождать и неверные результаты. Их анализ, в свою очередь, приведет к неправильным выводам в принятии решений.

С другой стороны, не менее важным является аспект представления информации для анализа и принятия решений. А это зависит от особенностей интерфейса, посредством которого осуществляется визуальная интерпретация результатов обработки информации, предназначенной для принятия решений.

Алгоритмизация и интерфейс – очень многогранные области. В статье внимание уделено только отдельным моментам, связанным с некоторыми прикладными вопросами, знание которых позволит лицам, принимающим решения (ЛПР), учитывать их при формулировании задач по подготовке аналитической информации, а также сокращать в дальнейшем время для ее анализа.

Изложение материалов по тематике принятия управленческих решений можно рассматривать в двух взаимосвязанных аспектах. Первый аспект связан с вопросами организационно-методического характера (в частности, используемый математический аппарат, особенности применения системного анализа, организационное взаимодействие между заинтересованными уровнями управления). Второй аспект – это специализированное программное обеспечение, т.е. системы поддержки принятия решений (СППР), ориентированное на обработку

информации для анализа и принятия решений. В настоящее время они основываются на приложениях в форме экспертных систем и искусственном интеллекте [1]. Кроме того, остаются востребованными и такие канонические формы аналитического контента, как таблицы (сформированные по определенным правилам) и графики. Они могут отражать результаты обработки информации в рамках экспертных систем или систем искусственного интеллекта, систем бизнес-аналитики или любых других расчетов, выполненных в условиях каких-либо специализированных систем.

1. О роли информации в процессе принятия решений

Кравченко Т.К. и Исаев Д.В. приводят 13 этапов в процессе принятия решений [2, с. 21–26]. Первый этап заключается в осознании ЛПР необходимости принятия решения (неформализуемая часть). Второй и третий этапы («Получение информации» и «Анализ информации» соответственно), с которых, собственно, и начинается сам процесс, связаны с информационным аспектом. Последующие этапы определяют классические действия при проведении системного анализа, позволяющие прийти к выбору варианта решения.

Говоря об использовании компьютера при решении прикладных задач в [3, с. 17] отмечается, что среди соответствующего программного обеспечения (ПО) для предметной области «Экономическая деятельность» представлены бухгалтерские программы, программы расчета смет, оценки эффективности инвестиционных проектов, оценки недвижимости и др. Кроме того, достаточно эффективно может быть использован табличный процессор MS Excel.

В [4] указывается, что термин «Управленческое решение» (УР) «... употребляется в двух основных значениях – как процесс и как явление. УР как процесс – это выполнение следующих основных процедур: информационной подготовки, разработки вариантов, согласование вариантов, выбор одного варианта, утверждения, реализации, кон-

троля выполнения УР и информирование инициатора решения. УР как явление — это набор мероприятий, направленных на разрешение рассматриваемой экономической проблемы в форме постановления, приказа или распоряжения, данной в устном или письменном виде» [4, с. 22].

В работе [5] имеется специальный раздел, посвященный информационной поддержке принятия решений, где информация для обозначенных целей представлена как ключевой фактор в контексте своих канонических особенностей, включая современные ее интерпретации в формате Интернета.

Табекин А.В. отмечает, что одним из факторов, определяющим качество управленческих решений, является «... объем и ценность располагаемой информации; для успешного принятия управленческого решения главным является не столько объем информации, сколько ее ценность (релевантность) и своевременность в сочетании с уровнем профессионализма, опыта, интуиции кадров, принимающих и реализующих управленческое решение...» [6, с. 36].

Крушанов А.А. обращает внимание, что «... ранней кибернетике были присущи две разноплановые спорные задачи...», первая из которых заключалась в том, что были выделены обобщенные процессы управления и феномен информации, как новые предметы научного познания. Это обозначено как управленческо-информационный аспект кибернетики. Автор понимает под управлением «... целенаправленное информационное воздействие, осуществляемое по схеме обратной связи» [7].

Из приведенных определений можно заключить, что информация является «главным действующим лицом» при решении задач, связанных с анализом и принятием управленческих решений.

Собственно процесс принятия решений базируется на следующих ключевых моментах:

1. Своевременно предоставленная информация, представленная в корректном и удобном для восприятия виде.
2. Математический инструментарий, позволяющий обрабатывать исходную информацию в целях ее интерпретации для принятия решений.
3. Программное обеспечение для автоматизированной обработки исходной информации согласно применяемым алгоритмам.
4. Не формализуемые аспекты, которые зависят, например, от интуиции ЛПП, от их компетенций по трактовке специальной информации в зависимости от конечных или промежуточных целей анализа.

2. Влияние ошибок при вводе информации в базу данных на ее последующую обработку в целях принятия решений

На примере *таблицы 1* прокомментируем некоторые виды алгоритмических ошибок, которые могут возникнуть в процессе начального формирования базы данных (БД) СППР и в дальнейшем провоцировать появление некорректных результатов при обработке информации.

Таблица 1.

Виды и примеры ошибок (по материалам [8, с. 38–39])

Вид ошибки	Пример ошибки	Комментарий
1. Ошибки в записи данных	Записано 5 вместо 4. Вместо кода 36, обозначающего предприятие «А» введен код 35.	Это моторные ошибки. Выявить их очень сложно. Можно только случайно или в контексте последующих решений, в которых неправильно введенная информация отразится на результатах и привлечет внимание аналитика. Вторая ошибка возможна даже в том случае, когда код (или название предприятия) будет выбираться из выпадающего списка.
2. Неустановленные данные	Получен платеж от покупателя под № 153, но в списке счетов к оплате покупатель с таким номером не обнаружен.	Появление ошибки такого рода может быть обусловлено двумя причинами. Первая: поскольку платеж сформирован «на стороне» и проверить правильность его заполнения нет возможности, то информацию из него внести в базу данных не получится. Если документ критически важен и игнорировать его содержимое нельзя, то выход только один: через коммуникационные каналы связаться с источником информации и выяснить вопрос «вручную». Вторая: некорректно сформирован список счетов к оплате. Решения: 1) перестроить список; 2) проверить алгоритм формирования списка; 3) проверить базу данных (возможно ошибка в номере счета была моторной).

Вид ошибки	Пример ошибки	Комментарий
3. Искажение регламентированного лимита.	По определенным видам поставок для покупателей определены ограничения на количество соответствующего товара в пределах одной покупки (например, 100 изделий). Но если менеджер выписал одному покупателю два наряда (например, по 75 изделий), общий объем которых превышает установленную норму, то этот покупатель сможет получить количество товара, превышающее установленное лимитированное значение.	Некорректный алгоритм. При попытке выписать второй наряд на количество изделий, превышающих 100, система должна либо предложить клерку выписать наряд на 25 изделий, либо заблокировать его действия при попытке выписать второй наряд на 75 изделий и проинформировать об этом пользователя.
4. Пропущенные записи	Запись в базе данных по каким-то причинам была удалена.	Такая ошибка может быть причиной появления ошибки № 2. Действия, связанные с удалением определенной информации, должны алгоритмически жестко регламентироваться. Например, применить вопрос «Вы действительно хотите удалить?». Или удаленная информация какое-то время может находиться в «корзине» с возможностью ее восстановления при необходимости.
5. Ошибки при формировании отчетов	ЛПР при анализе отчета предполагает, что в него включена информация по проданной и оплаченной продукции. Однако фактически в документе помимо оплаченных товаров присутствуют товары, проданные в кредит.	Аналогичные ошибки могут встречаться достаточно часто вследствие некорректного описания задачи для программирования. Для формирования отчетных (аналитических, статистических) документов целесообразно предлагать всегда их графическую структуру с указанием правильных названий граф, а также при возможности необходимо указать реквизиты БД, откуда будет выбираться информация.
6. Ошибки при вводе данных	Повышение заработной платы для работника с табельным номером 174 по некоторым причинам записано в БД для работника с табельным номером 175.	С одной стороны, такая ошибка может быть моторной. С другой стороны – преднамеренной. В любом случае, если работник знает о повышении зарплаты, то при ее первом после повышения получении ошибка будет выявлена. Если же работник не уведомлен о повышении зарплаты, то ситуация относится к морально-этической сфере.
7. Повторение ошибки	В БД введена неверная цена на изделие, поэтому изделия этого вида выставляются к оплате по этой неверной цене.	Ошибка могла быть моторной или преднамеренной. Поскольку источником для ввода подобного рода информации обычно является бумажный документ, то его отсканированную копию всегда можно на некоторое время сохранять в специальных папках или файлах. Периодически можно осуществлять «ручные» проверки. Зная об этом, сотрудник вряд ли будет рисковать.
8. Неверное разграничение по периодам	Информация о заказе была внесена в БД в последний день месяца. Но фактическая отгрузка товара по данному заказу осуществилась только через несколько дней уже в следующем месяце. Тем не менее, информация о продаже отражена в БД за предыдущий месяц.	Причина ошибки – некорректный алгоритм. Это – типичный пример игнорирования обработки нестандартных ситуаций. Решение должно опираться на соблюдение условий, указанных потенциальным пользователем.
9. Фальсификация данных	Сотрудник организации, имеющий доступ к финансовой информации других работников и обладающий правами на изменение информации, удалил записи о неоплаченных счетах своего коллеги, что обеспечило последнему возможность не платить за приобретенные товары и (или) услуги.	Причина ошибки – некорректное разграничение прав доступа для сотрудника при работе с информацией. Удаление информации – всегда критически важный аспект. Поэтому все действия, связанные с удалением, подлежат жесткой регламентации и контролю. Это может быть сделано алгоритмически через процедуру протоколирования (журналирования) работ.
10. Некорректный учет	Деньги, которые были переданы представителю покупателя для одних целей, были потрачены на совершенно другие цели (например, вместо оплаты инструментов для работы был оплачен деловой ужин в ресторане).	Ошибки такого рода могут быть выявлены либо случайно, либо по результатам анализа отчетных документов, либо в процессе аудита. Поскольку источником для ввода подобного рода информация, вероятно, является бумажный документ, то всегда можно определить логическую зависимость между его реквизитами. И при заполнении соответствующей формы на экране алгоритмически такую связь выявлять. Если для ввода информации из документа используется процедура распознавания, то также может быть применен специальный алгоритм, определяющий корреляцию между логически связанными реквизитами.

Приведенные в *таблице 1* примеры ошибок относятся к предметной области «Бухгалтерия». Однако нетрудно спроецировать их и на любые другие предметные области.

3. Алгоритмические особенности интерфейса

Более 20 лет назад Якоб Нильсен предложил 10 основных принципов, которые необходимо учитывать при проектировании взаимодействия пользователя с системой [9]. Данные принципы универсальны по своей сути. Их можно считать классическими, поэтому они не утратили своего значения и в настоящее время. Игнорирование этих принципов может провоцировать различного рода проблемы при использовании самых разных типов автоматизированных информационных систем, включая СППР.

Кратко прокомментируем те принципы, которые должны найти свое отражение в алгоритмическом контексте, и которые потенциальный пользователь может учесть при формулировании функциональных требований к СППР.

Принцип 1: видимость состояния системы. Длительность решения разных по сути задач может быть различной. Есть задачи, решение которых осуществляется настолько быстро, что при существующей скорости обработки информации на современных компьютерах результат выдается практически мгновенно после нажатия клавиши «Enter». Однако имеются и другие задачи (например, составление производственных планов, расчет каких-либо показателей для большого количества работников, обработка значительных объемов статистической информации для составления прогнозных оценок и др.), решения которых могут быть достаточно продолжительными. В этом случае для пользователя, ответственного за выполнение задачи, полезно представить визуальную информацию о ее хронологии. Данная информация может быть как дискретной, так и непрерывной. Можно использовать горизонтальный индикатор, круговые указатели (часы или секторальное заполнение) или какие-либо другие.

Принцип 2: соответствие между системой и реальным миром. Нужно всегда помнить о том, что СППР предназначена для использования в определенной предметной области, в которой обязательно присутствует специфическая терминология. Поэтому вся информация, представленная пользователю на экране монитора, должна полностью соответствовать тем смысловым значениям (понятия, определе-

ния, обозначения, названия объектов, показателей, коэффициентов, процессов, ситуаций, явлений, зависимостей и т.п.), которые приняты именно в этой области. В интерфейсе не должно быть терминологии, характерной для ИТ-сферы, за исключением процедур и действий, ставших стандартами «де-факто» (копировать, печатать, вырезать, сохранить и т.п.) при условии, что их терминологическая замена нецелесообразна. Во всех других случаях всегда необходимо найти аналогию между терминологией ИТ-сферы и предметной областью. Например, в интерфейсе не должно присутствовать слово «База данных», поскольку в реальности ей может соответствовать «План счетов», «Личные дела сотрудников», «Производственное оборудование», «График платежей» и т.п.

Принцип 3: пользовательский контроль и свобода. Пользователь не должен бояться совершить ошибку. Система всегда должна его «подстраховывать». Это может быть:

- ♦ предупредительное сообщение (например, «Вы действительно хотите удалить запись о сотруднике с табельным № ...?», «После изменения показателя «Ω» восстановить его начальное значение будет невозможно», «Проверьте синхронизацию времени с сетью Интернет», «Вы не указали период расчета», «Точность расчетов не должна превышать 2 знака в дробной части»);
- ♦ контекстная подсказка при вводе информации (например, «Значение реквизита не должно превышать 10 символов», «Ввод только цифровых данных»);
- ♦ возможность выбора значения из выпадающего списка при заполнении поля (применение необходимых справочников, словарей, классификаторов и других упорядоченных совокупностей определенных объектов), существенно уменьшающая вероятность ввода неверного значения (в СППР это может быть использовано для составления запроса на получение необходимой информации);
- ♦ возможность возврата на предыдущие этапы обработки.

Принцип 4: согласованность и стандарты. В рамках одной системы все используемые названия объектов (процессов, явлений, ситуаций и т.п.) должны быть обязательно унифицированы. Для этих целей создаются, в частности, классификаторы, справочники, словари, которые принято относить к нормативно-справочной информации. Упорядо-

ченные совокупности соответствующих значений позволяют экономить время при вводе информации, минимизировать ошибки, оптимизировать объемы памяти компьютера при хранении введенных данных, унифицировать требуемую информацию во всех документах, отражающим любого рода аналитику (рис. 1).

Технологически использование значений из классификаторов или их аналогов осуществляется через выпадающие списки.

Принцип 5: предотвращение ошибок. Грамотно спроектированный интерфейс должен либо вообще оградить пользователя от возможности допустить ошибки, либо минимизировать их появление. При вводе информации это может быть сделано, например, за счет контекстных подсказок, выбора значений из предложенных списков (перечней), использования масок или шаблонов. Для данных типа «год» (при вводе значения вручную) целесообразно задавать дополнительный диапазон изменения, так как неправильно введенное значение может оказаться «за пределами» допустимого исторического периода.

Для указания дат и (или) временных периодов, как правило, применяются современные инструментальные средства, которые позволяют практически полностью исключить ошибку при выборе дат. Дата в формате «ДД.ММ.ГГГГ» может выбираться либо из встроенного календаря, либо «собираться» из частей (день, месяц, год) — из последовательностей логически обоснованных предлагаемых чисел. Пользователю остается только указать нужные значения. Ошибка при этом может быть случайной по невнимательности.

Достаточно просто выявлять ошибку в показателе логического типа (т. е. он может принимать одно из двух значений — «0» или «1»).

В любом случае «ручной» ввод, особенно для текстовой и (или) символьной информации должен быть сведен к минимуму.

Принцип 6: распознавание, а не вспоминание. Пользователь должен чувствовать себя комфортно и безопасно при работе с интерфейсом. Поэтому ему нужно создать условия, которые не потребуют от него запоминания точной последовательности каких-либо своих действий при выполнении определенных задач. Он должен быть всегда уверен в том, что система либо подскажет ему при совершении ошибки простой логический выход, либо предоставит ссылку на соответствующий фрагмент инструкции по работе с приложением.

Принцип 7: гибкость и эффективность использования. В любой СППР всегда могут присутствовать параметры, которые пользователь может настроить под свои индивидуальные предпочтения, если он использует соответствующую информацию для решения только своих текущих задач. Например, в системе 1С:Предприятие предусмотрено 76 (!) объектов интерфейса, для которых пользователь самостоятельно может изменить их цветовое оформление согласно своим требованиям [10]. Если же информация (в частности, сформированный отчет) передается во внешнюю часть системы, в которой осуществляются расчеты, или в систему, находящуюся за периметром соответствующей организации, то она должна соответствовать принятым корпоративным стандартам.

Оставшиеся три принципа («Эстетичный и минималистический дизайн», «Помощь пользователям в распознавании, диагностировании и исправлении ошибок» и «Помощь и документация») не имеют непосредственного отношения к алгоритмизации и поэтому не влияют на формулирование функциональных требований пользователя. Данные принципы должны быть реализованы профессиональными разработчиками приложений «по определению».

<p>Возможные варианты представления названия структурного подразделения при «ручном» вводе</p> <ol style="list-style-type: none">1. отдел кадров;2. отд. Кадров;3. ОК;4. Отд. кадр.;5. Отдел кадров;6. Отдел Кадров и т.п.	<p>Единственный вариант представления названия структурного подразделения в системе с использованием справочника структурных подразделений</p> <p>Отдел кадров</p>
--	---

Рис. 1. Иллюстрация унификации для названия управленческого подразделения на предприятии (в организации).

4. Особенности использования текстовой и символьной информации в интерфейсе

Информация, предназначенная для анализа, должна удовлетворять определенным требованиям. Это объясняется тем, что работа в виртуальной среде на должна создавать дискомфорта, обусловленного психофизиологическими характеристиками человека.

Интерфейс, являясь виртуальной средой, с точки зрения пользователя не должен противоречить его действиям в реальном мире. Это связано с общими принципами мыслительной деятельности человека (в широком смысле), а также с особенностями восприятия информации через органы чувств. Именно такого рода аспекты отражены Сьюзан Уэйншенк в [11].

В качестве примера рассмотрим влияние различных шрифтов и маркеров, которые могут использоваться для представления аналитической информации.

Существуют очень большое разнообразие шрифтов. На ресурсе [12] их представлено больше 80 типов.

Приведем несколько рекомендаций по использованию некоторых шрифтов [13]:

1. Шрифты с засечками (самый популярный – Times New Roman) обладают универсальным назначением и могут применяться для самого разного текста. Но особенно важно, что они хорошо представляют текст даже при использовании маленького размера кегля и поэтому целесообразны для отображения большого объема текста.
2. Шрифты без засечек или рубленые (примеры: Arial, Tahoma, Verdana) и моноширинные (пример: Courier New) рекомендуется использовать для заголовков или выделенных фрагментов текста, но не для текста в целом.
3. Шрифты, стилизованные под рукописные, затрудняют чтение больших текстов. Поэтому рекомендуются в качестве заголовков или выделения фрагментов текста.
4. Декоративные шрифты (например, готические или старославянские) следует применять с особой осторожностью, поскольку они не соответствуют современности и поэтому могут восприниматься с трудом.

5. Контурный шрифт, шрифт с тенью, а также подчеркнутый текст не рекомендуется применять для больших фрагментов текста, так как это затрудняет чтение (теряется четкость текста, что усложняет его восприятие и понимание).

6. Существует такое понятие, как «выворотка», когда, например, белые буквы располагаются на черном фоне. Использование такого приема для длинного текста, представленного мелким шрифтом также не очень комфортно и будет затруднять чтение. Однако данный прием может использоваться для заголовков.

В Интернете есть специальные бесплатные ресурсы по преобразованию текста в любой шрифт, например, [14, 15].

Предоставление публичной информации, особенно в рамках функционального ПО, должно соответствовать общим правилам ее восприятия. В качестве примера приведем текст, расположенный по ссылке [16] (рис. 2). На рис. 3 тот же текст преобразован к виду, удобному для восприятия и понимания (орфография источника сохранена). Это было сделано за счет структурирования информации, т.е. разбиения текста на фрагменты и выделения однотипных смысловых значений в форме списков. Разница в представлении одной и той же информации очевидна.

В качестве маркеров для списков рекомендуется применять символы, закрашенные внутри (например: ■ ● ◆ ◇ -), т.е. контрастные по отношению к тексту, поскольку их начертание принципиально отличается от букв, а суть маркера состоит именно в том, чтобы выделить элементы списка. В противовес этому не рекомендуется использовать для этих же целей символы, «пустые» внутри, так как по своему начертанию они напоминают буквы (например: ❖ □ □ ☒ ◇ ○), хотя таковыми не являются. Однако четкость списков в этом случае будет «размытой».

Немаловажное значение имеет использование иконок и картинок. Можно либо рисовать свои оригинальные, либо выбирать только те, которые представлены на специальных бесплатных ресурсах. Нельзя копировать готовые рисунки, картинки, иконки, логотипы или их фрагменты, если они авторские. В противном случае это будет противоречить Закону РФ «О товарных знаках, знаках обслуживания и наименованиях мест происхождения товаров».

**Описание программы:
«Orakl-Кадры» («ОРАКуЛ-Кадры»)**

«Система кадрового документооборота и управления персоналом» – мощнейшая кадровая программа под ОС Microsoft Windows 95/98/NT4/2000/ME/XP, охватывающая все аспекты ведения кадрового учета и документооборота. Программа создана при помощи ведущих менеджеров по HR г.С-Петербурга, пользуется популярностью более 3,5 лет. Программа возьмёт на себя все то, что Вам раньше приходилось делать вручную, перебирая «массы бумаги». «Orakl-К» выгодно отличается простота и удобство управления, она универсальна и легка в работе любому пользователю компьютера, как опытному, так и новичку. В «стандартную версию» входят: пароль на входе, 4 базы данных (штатных, кадрового резерва, уволенных, архива) с учетными карточками сотрудников (более 130 тем&окон!), ведение всего учета (табеля, отпуска, командировки, больничные, ссуды, мат.помощь, компенсации и т.п.), огромная библиотека кадровых документов-шаблонов, автоматическое создание приказов, настраиваемое штатное расписание (любой сложности!), Ваши личные заметки, весь воинский учет, стажи, блок анализа, более 33 тысяч (!!!) выборок-отчетов (т.е. нахождение любой информации).

Рис. 2. Пример «нечитабельного» текста (фрагмент) [16].

Описание программы: «Orakl-Кадры» («ОРАКуЛ-Кадры»)

«Система кадрового документооборота и управления персоналом» — мощнейшая кадровая программа под ОС Microsoft Windows 95/98/NT4/2000/ME/XP, охватывающая все аспекты ведения кадрового учета и документооборота.

Программа создана при помощи ведущих менеджеров по HR г.С-Петербурга, пользуется популярностью более 3,5 лет. Программа возьмёт на себя все то, что Вам раньше приходилось делать вручную, перебирая «массы бумаги». «Orakl-К» выгодно отличается простота и удобство управления, она универсальна и легка в работе любому пользователю компьютера, как опытному, так и новичку.

В «стандартную версию» входят:

- пароль на входе, 4 базы данных (штатных, кадрового резерва, уволенных, архива) с учетными карточками сотрудников (более 130 тем & окон!);
- ведение всего учета (табеля, отпуска, командировки, больничные, ссуды, мат. помощь, компенсации и т.п.);
- огромная библиотека кадровых документов-шаблонов;
- автоматическое создание приказов;
- настраиваемое штатное расписание (любой сложности!);
- Ваши личные заметки;
- весь воинский учет;
- стажи;
- блок анализа;
- более 33 тысяч (!!!) выборок-отчетов (т.е. нахождение любой информации);
- блоки тестирования и аттестации и многое-многое другое.

Возможна как стандартная поставка, так и комплектация под Ваш конкретный заказ. Программа учитывает требования Постановления Правительства РФ и соответствующего Постановления Госкомстата. Для нестандартных решений предлагаются дополнительные 35 стандартных модулей расширения или функции под заказ.

Рис. 3. Пример преобразованного текста (фрагмент).

5. Особенности представления информации для анализа и принятия решений

Поскольку качество принятия решений имеет прямую зависимость от используемой при этом информации, то очень важна ее визуализация в определенной форме. С одной стороны, она должна содержать действительно только те показатели (зависимости, результаты расчетов, статистику и т.п.), которые необходимы для анализа и принятия управленческих решений в конкретных ситуациях. С другой стороны, представление информации должно быть настолько наглядным, чтобы затраты времени на ее рассмотрение были минимальными независимо от того, как быстро решение необходимо принять.

Сценарий диалога в рамках СППР может включать в себя указание некоторых оперативных условий для выбора информации и ее обработки. К ним можно отнести, например:

- ◆ период времени (с ... по ...);
- ◆ структурное подразделение (производственное или управленческое любого уровня);
- ◆ группу товаров или услуг;
- ◆ рынок сбыта (район, город, регион, страну);
- ◆ ценовые или стоимостные характеристики;
- ◆ способ доставки товаров;
- ◆ упорядочение по возрастанию или убыванию показателей;
- ◆ подведение промежуточных и (или) окончательных итогов;
- ◆ выделение результатов, выходящих за пределы установленных ограничений (в большую или меньшую сторону) или имеющих отклонения, превышающие регламентированные границы.
- ◆ и др.

Сценарий диалога может также предусматривать необходимость и (или) вариативность последующего решения в зависимости от промежуточных результатов, полученных либо на определенном этапе решения, либо до регламентированного момента времени.

Как известно, любая результирующая информация для анализа может быть получена двумя путями. Первый: структуры документов известны и поэтому алгоритм их формирования заранее запрограммирован. В этом случае ЛПР достаточно выбрать нужное название документа (в широком

смысле, так как помимо традиционной таблицы документ может быть представлен, например, в виде графика, «плоского» текста или анимированной «картинки») из предлагаемого списка. Второй: структура формы документа задается самим пользователем непосредственно перед формированием нужного документа. Для этого пользователь должен иметь навыки составления запроса на выбор информации (т.е. владеть специальным языком запросов, предусмотренным в конкретной СППР), что позволяет без специального программирования оперативно генерировать любые формы документов по желанию самого пользователя.

В табличных документах, предназначенных для анализа, должен отсутствовать «информационный шум», который не несет никакой семантической нагрузки, но на подсознательном уровне отвлекает внимание пользователя и замедляет его работу. Пример показан на *рис. 4*. Вариант избавления от такого рода «информационного шума», затрудняющего восприятие информации, достаточно прост. Для приведенной таблицы нужно в названии соответствующей графы через «запятую» добавить то значение (в примере это – единица измерения), которое является повторяющимся (*рис. 5*).

Компьютерные технологии позволяют особым образом выделять отдельные фрагменты аналитических документов, способствующие привлечению внимания аналитиков. Для этих целей можно использовать, например, изменение шрифта (по стилю, цвету, размеру), «закрашивание» или обрамления цветом определенных клеток (строк, столбцов) таблицы, а также другие возможности. Такой «дизайн» целесообразно применять для того, чтобы показать значения, существенно отличающиеся от прочих, изменение каких-либо показателей на критически допустимую величину, приближение значения показателя к регламентированному пределу и др.

В отдельных случаях может использоваться цветное кодирование. Это будет полезно, например, в том случае, когда в одном аналитическом документе одновременно нужно отразить различные статусы объектов и (или) их характеристик. Таким образом, в рамках одного документа можно отметить, в частности, товары: ждущие отгрузки, отгруженные, ожидающие оплаты, оплаченные.

При проектировании таблиц необходимо учитывать одну особенность. Если получается, что в одной из граф во всех строках присутствует одинаковое значение (текстовое, цифровое, логическое, «дата»

Year Result: 2022 Smartphone shipments						
Rank	OEM	Y2022		Y2021		YoY
		Shipment	M/S	Shipment	M/S	
1	Samsung	259	21%	272	20%	-4.8%
2	Apple	231	19%	235	18%	-1.3%
3	Xiaomi	152	13%	190	14%	-20.0%
4	Oppo Group	107	9%	144	11%	-25.6%
5	vivo	98	8%	134	10%	-27.1%
6	Transsion	68	6%	75	6%	-8.7%

Рис. 4. Пример аналитического документа (фрагмент), содержащего «информационный шум» [17].

Year Result: 2022 Smartphone shipments						
Rank	OEM	Y2022		Y2021		YoY, %
		Shipment	M/S, %	Shipment	M/S, %	
1	Samsung	259	21	272	20	-4.8
2	Apple	231	19	235	18	-1.3
3	Xiaomi	152	13	190	14	-20.0
4	Oppo Group	107	9	144	11	-25.6
5	vivo	98	8	134	10	-27.1
6	Transsion	68	6	75	6	-8.7

Рис. 5. Пример аналитического документа (фрагмент) без «информационного шума».

или какое-либо другое), то это противоречит одному из свойств информации – информативности. В подобных случаях целесообразно название такой графы отразить в заголовке таблицы и данную графу из структуры таблицы исключить. Таблица при этом станет более компактной и удобной для анализа.

Известно, что в отдельных случаях для аналитической информации более предпочтительной и наглядной будет графическая форма. Она может существовать в качестве единственного варианта представления соответствующей информации или дополнительного к классической таблице. Хорошим образцом аналитики и статистики являются статические ежегодники, издаваемые Высшей школой экономики. На рис. 6 и 7 показаны такие примеры.

Заключение

Тематика, отражающая теоретические и практические аспекты в области СППР, представляет интерес для разных групп специалистов. Основная часть публикаций посвящена общим вопросам канониче-

ского характера в данной предметной области, например [20], использованию конкретных методов и моделей, описанию применения и сопутствующих особенностей СППР в самых различных функциональных сферах (медицине [21–23], специализиро-

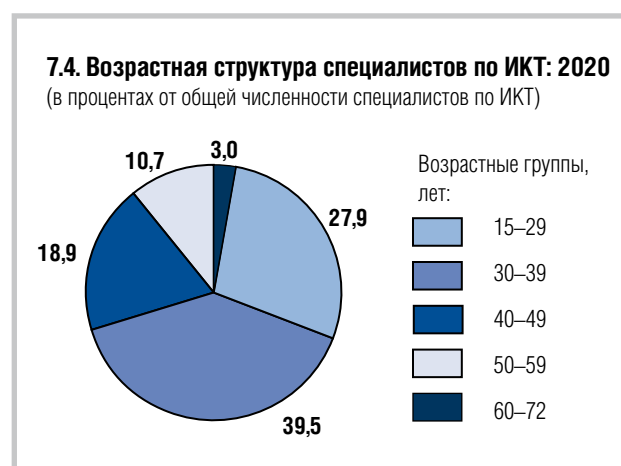


Рис. 6. Пример 1 графического представления аналитической информации [18, с. 80].



Рис. 7. Пример 2 графического представления аналитической информации [18, с. 93].

ванных университетах [24], оборонной технике [25], углеобогатительных предприятиях [26], лесном хозяйстве [27–28], транспорте [29], финансовых вопросах [30–31] и др.), а также описанию используемых СППР в виде конкретного ПО.

Тем не менее, недостаточно внимания уделяется прикладным вопросам, которые необходимо учитывать как при формулировании функциональных требований и оценке СППР со стороны потенциальных пользователей, так и при проектировании и разработке подобного рода систем со

стороны разработчиков. В статье рассматриваются отдельные вопросы именно в этих аспектах.

Алгоритмизация в сфере поддержки принятия управленческих решений и сопутствующий интерфейс представляют собой динамичные и развивающиеся области, которые включают множество нюансов, часто остающиеся вне поля зрения пользователей и разработчиков. Именно это может создавать проблемы при применении различных информационных систем, включая СППР. Поэтому обсуждаемые в статье вопросы, касающиеся алгоритмизации и интерфейса, можно рассматривать с различной степенью глубины и дифференциации, а также критического анализа присутствующих на рынке СППР, что может стать предметом дальнейших исследований.

Материалы статьи могут быть полезны специалистам, которые выступают в роли ЛПР. Они могут использовать их при формулировании функциональных требований к разработке СППР по заказу, а также при оценке интерфейса уже существующих систем, приобретаемых предприятием или организацией для решения своих задач по поддержке принятия управленческих решений.

Сведения, изложенные в статье, могут быть также применены в прикладном аспекте студентами, магистрантами и аспирантами, которые занимаются исследованиями в сфере принятия управленческих решений и (или) создания СППР.

Отдельные моменты, изложенные в статье, могут представлять интерес для разработчиков СППР. ■

Литература

1. Демидовский А.В., Бабкин Э.А. Интегрированные нейросимволические системы поддержки принятия решений: проблемы и перспективы // Бизнес-информатика. 2021. Т. 15. № 3. С. 7–23. <https://doi.org/10.17323/2587-814X.2021.3.7.23>
2. Кравченко Т.К., Исаев Д.В. Системы поддержки принятия решений: учебник и практикум для вузов. М.: Юрайт, 2022.
3. Набатова Д.С. Математические и инструментальные методы поддержки принятия решений: учебник и практикум для вузов. М.: Юрайт, 2023.
4. Теория принятия решений. В 2 томах. Т. 1: учебник и практикум для вузов / В.Г. Халин [и др.]; ответственный редактор В.Г. Халин. М.: Юрайт, 2023.
5. Теория принятия решений. В 2 томах. Т. 2: учебник и практикум для вузов / В.Г. Халин [и др.]; ответственный редактор В.Г. Халин. М.: Юрайт, 2023.
6. Табекин А.В. Методы принятия управленческих решений: учебник для вузов. М.: Юрайт, 2023.
7. Крушанов А.А. Понятие «управление» в кибернетическом контексте // Vox. Философский журнал. 2017. № 23. <https://doi.org/10.24411/2077-6608-2017-00020>
8. Перри У. ЭВМ и организация бухгалтерского учета. М.: Финансы и статистика, 1986.
9. 10 основных принципов Якоба Нильсена для проектирования взаимодействия. [Электронный ресурс] <https://www.nngroup.com/articles/usability-heuristics-complex-applications/> (дата обращения 24.07.2023).

10. ДемOVERсия системы 1С: Предприятие. [Электронный ресурс] <https://v8.1c.ru/podderzhka-i-obuchenie/demonstratsionnye-konfiguratsii/> (дата обращения 24.07.2023).
11. Уэйншенк С. 100 главных принципов дизайна: как удержать внимание. СПб: Питер, 2015.
12. Подбор шрифтов онлайн. [Электронный ресурс] <https://ffont.ru/fonts> (дата обращения 24.07.2023).
13. Массель Г.Г. Психологические аспекты разработки компьютерных систем: науч. тр. Институт систем энергетики им. Л.А. Мелентьева СО РАН. Иркутск: ИСЭМ СО РАН, 2007.
14. Красивые шрифты онлайн. [Электронный ресурс] <https://beautifulfonts.net/#page1> (дата обращения 24.07.2023)
15. Красивые шрифты онлайн. [Электронный ресурс] <https://prettyfont.net/> (дата обращения 24.07.2023).
16. Электронный портал Kazus.ru. [Электронный ресурс] http://kazus.ru/programs/viewdownloadetails/kz_0/lid_4350.html (дата обращения 24.07.2023).
17. Смартфоны (мировой рынок). TAdviser. [Электронный ресурс] <https://www.tadviser.ru/a/47257> (дата обращения 24.07.2023).
18. Цифровая экономика: 2022: краткий статистический сборник / Г.И. Абдрахманова, С.А. Васильковский, К.О. Вишнеvский и др.; Нац. исслед. ун-т «Высшая школа экономики». М.: НИУ ВШЭ, 2022.
19. Заславская В.Л. Системы поддержки принятия решений и их роль в информационно-управляющих системах // Экономика и управление: проблемы, решения. 2022. Т. 2. № 12. С. 144–153. <https://doi.org/10.36871/ek.up.prg2022.12.02.017>
20. Ahmed A., Spiess B., Kortsmit J., Van den Ham R., Erdoes G., Klein A. A narrative review of clinical decision support systems for perioperative bleeding management in cardiac surgery // Journal of Cardiothoracic and Vascular Anesthesia. 2023. Vol. 37. No. 9. P. 1804–1812. <https://doi.org/10.1053/j.jvca.2023.05.008>
21. Toffaha K.M., Simsekler M.C.E., Omar M.A. Leveraging artificial intelligence and decision support systems in hospital-acquired pressure injuries prediction: A comprehensive review // Artificial Intelligence in Medicine. 2023. Vol. 141. 102560. <https://doi.org/10.1016/j.artmed.2023.102560>
22. Малых В.Л. Системы поддержки принятия решений в медицине // Программные системы: теория и приложения. 2019. Т. 10. № 2(41). С. 155–184. <https://doi.org/10.25209/2079-3316-2019-10-2-155-184>
23. Вечерская С.Е. Выбор критериев системы поддержки принятия решений для художественного вуза // Информатика и образование. 2021. № 3. С. 56–62. <https://doi.org/10.32517/0234-0453-2021-36-3-56-62>
24. Оркин В.В., Нестеренко О.Е., Платонов С.А. Модель системы ситуационного управления в автоматизированной системе поддержки принятия решений // Вопросы оборонной техники. Серия 16: Технические средства противодействия терроризму. 2021. № 1–2(151–152). С. 40–45.
25. Лепило Н.Н., Катан К.С. Повышение эффективности деятельности углеобогатительного предприятия на основе СППР // Экономический вестник Донбасского государственного технического университета. 2019. № 2. С. 52–59.
26. Punit B., Reddy A.S., Dave T.N. A spreadsheet-based decision support system for selection of optimal soil liquefaction mitigation technique // Decision Analytics Journal. 2023. Vol. 6. 100154. <https://doi.org/10.1016/j.dajour.2022.100154>
27. Aldea J., Bianchi S., Nilsson U., Hynynen J., Lee D., Holmström E., Huuskonen S. Evaluation of growth models for mixed forests used in Swedish and Finnish decision support systems // Forest Ecology and Management. 2023. Vol. 529. 120721. <https://doi.org/10.1016/j.foreco.2022.120721>
28. Kaltsidis A., Ketikidis K., Basbas S., Aifadopoulou G., Grau J.M.S. A decision support system for taxi drivers // Transportation Research Procedia. 2023. Vol. 69. P. 123–130.
29. Siconolfi S.F. An analytics-based decision support system for evaluating the fiscal health of academic programs // Decision Analytics Journal. 2022. Vol. 4. 100091. <https://doi.org/10.1016/j.dajour.2022.100091>
30. Borrero-Domínguez C., Escobar-Rodríguez T. Decision support systems in crowdfunding: A fuzzy cognitive maps (FCM) approach // Decision Support Systems. 2023. 114000. <https://doi.org/10.1016/j.dss.2023.114000>

Об авторе

Гутгарц Римма Давыдовна

доктор экономических наук;

профессор, институт информационных технологий и анализа данных, Иркутский национальный исследовательский технический университет, Россия, 664074, г. Иркутск, ул. Лермонтова, д. 83;

E-mail: gutgarc@gmail.com

ORCID: 0000-0001-9881-1976

Influence of algorithmization and interface for the preparation of management decisions

Rimma D. Gutgarts

E-mail: gutgarc@gmail.com

Irkutsk National Research Technical University

Address: 83, Lermontov St., Irkutsk 664074, Russia

Abstract

In modern conditions, managerial decision-making is carried out using automated systems under the general name “Decision Support Systems” (DSS). When creating them, it is important to consider two key points. The first is the algorithmic component, which reflects the logic of the system as a whole and its individual parts. The second is the application interface through which the user interacts with it. The interface is a graphical interpretation of the algorithms that are implemented within the system. Therefore, it is very important to design and create such a relationship between the algorithm and the interface so that the user is as comfortable as possible using the DSS to solve current tasks (information input, its processing, presentation and analysis for decision making). Thus, there is a directly proportional relationship between the interface and the algorithm. Moreover, despite the fact that there are many studies on these aspects, both theoretical and practical, there are still questions to which one should pay attention to in terms of application. The purpose of this study is to formulate practical recommendations to prevent the entry of incorrect information into the DSS database and to present the results in a form convenient for its analysis. The main tasks of the work are to show by means of examples which errors can contribute to the entry of unreliable information into the database, as well as how best to present information on the monitor screen in accordance with the psychophysiological characteristics of a person in order to reduce the time for its analysis and decision-making.

Keywords: decision support systems, features of algorithmization, interrelation of the interface and algorithms, presentation of information, error handling when entering information

Citation: Gutgarts R.D. (2023) Influence of algorithmization and interface for the preparation of management decisions. *Business Informatics*, vol. 17, no. 3, pp. 24–37. DOI: 10.17323/2587-814X.2023.3.24.37

References

1. Demidovskij A.V., Babkin E.A. (2021) Integrated neurosymbolic decision support systems: problems and opportunities. *Business Informatics*, vol. 15, no 3, pp. 7–23. <https://doi.org/10.17323/2587-814X.2021.3.7.2>
2. Kravchenko T.K., Isaev D.V. (2022) *Decision support systems: textbook and workshop for universities*. Moscow: Urait (in Russian).
3. Nabatova D.S. (2023) *Mathematical and instrumental methods of decision support: textbook and workshop for universities*. Moscow: Urait (in Russian).
4. Khalin V.G. et al. (2023) *Theory of decision making. In two volumes. Vol. 1: textbook and workshop for universities*. Moscow: Urait (in Russian).
5. Khalin V.G. et al. (2023) *Theory of decision making. In two volumes. Vol. 2: textbook and workshop for universities*. Moscow: Urait (in Russian).
6. Tabekin A.V. (2023) *Methods of making managerial decisions: a textbook for universities*. Moscow: Urait (in Russian).
7. Krushanov A.A. (2017) The concept of «management» in the cybernetic context. *Vox. Filosofskiy zhurnal* (electronic journal), vol. 23 (in Russian). <https://doi.org/10.24411/2077-6608-2017-00020>

8. Perry W. (1986) *Computer and the organization of accounting*. Moscow: Finance and statistics (in Russian).
9. Nielsen J. (1994) *10 Usability Heuristics for User Interface Design*. Nielsen Norman Group. Available at: <https://www.nngroup.com/articles/ten-usability-heuristics/> (accessed 24 July 2023).
10. *IC:Enterprise. Demo applications*. Available at: <https://v8.1c.ru/podderzhka-i-obuchenie/demonstratsionnye-konfiguratsii/> (accessed 24 July 2023).
11. Weinshank S. (2015) *100 main principles of design: how to keep attention*. St. Petersburg: Piter (in Russian).
12. *Free fonts*. Available at: <https://ffont.ru/en/fonts> (accessed 24 July 2023).
13. Massel G.G. (2007) *Psychological aspects of the development of computer systems*. Melentiev Energy Systems Institute Siberian Branch of the Russian Academy of Sciences. Irkutsk: ESI SB RAS (in Russian).
14. *Beautiful fonts online*. Available at: <https://beautifulfonts.net/en/> (accessed 24 July 2023).
15. *Beautiful fonts online*. Available at: <https://prettyfont.net/> (accessed 24 July 2023).
16. *Kazus.ru*. Available at: http://kazus.ru/programs/viewdownloaddetails/kz_0/lid_4350.html (accessed 07/24/2023).
17. *Smartphones (Global Market)*. TAdviser. Available at: <https://tadviser.com/a/e.php?id=47257> (accessed 24 July 2023).
18. Gokhberg L., Kuzminov Y.I., et al. (eds.) (2022) *Digital economy, 2022: Pocket data book*. Moscow: HSE.
19. Zaslavskaya V.L. (2022) Decision support systems and their role in information management systems. *Economics and management: problems, solutions*, vol. 2, no. 12, pp. 144–153. <https://doi.org/10.36871/ek.up.pr2022.12.02.017>
20. Ahmed A., Spiess B., Kortsmitt J., van den Ham R., Erdoes G., Klein A. (2023) A narrative review of clinical decision support systems for perioperative bleeding management in cardiac surgery. *Journal of Cardiothoracic and Vascular Anesthesia*, vol. 37, no. 9, pp. 1804–1812. <https://doi.org/10.1053/j.jvca.2023.05.008>
21. Toffaha K.M., Simsekler M.C.E., Omar M.A. (2023) Leveraging artificial intelligence and decision support systems in hospital-acquired pressure injuries prediction: A comprehensive review. *Artificial Intelligence in Medicine*, vol. 141, 102560. <https://doi.org/10.1016/j.artmed.2023.102560>
22. Malykh V.L. (2019) Decision support systems in medicine. *Program systems: Theory and applications*, vol. 10, no. 2(41), pp. 155–184. <https://doi.org/10.25209/2079-3316-2019-10-2-155-184>
23. Vecherskaya S.E. (2021) Selection of criteria for a decision support system for an art university. *Informatics and Education*, no. 3, pp. 56–62. <https://doi.org/10.32517/0234-0453-2021-36-3-56-62>
24. Orkin V.V., Nesterenko O.E., Platonov S.A. (2021) A model of a situational control system in an automated decision support system. *Questions of defense technology. Series 16: Technical means of countering terrorism*, no. 1–2(151–152), pp. 40–45.
25. Lepilo N.N., Katan K.S. (2019) Improving the efficiency of a coal preparation enterprise based on the DSS. *Economic Bulletin of the Donbass State Technical University*, no. 2, pp. 52–59.
26. Punit B., Reddy A.S., Dave T.N. (2023) A spreadsheet-based decision support system for selection of optimal soil liquefaction mitigation technique. *Decision Analytics Journal*, vol. 6, 100154. <https://doi.org/10.1016/j.dajour.2022.100154>
27. Aldea J., Bianchi S., Nilsson U., Hynynen J., Lee D., Holmström E., Huuskonen S. (2023) Evaluation of growth models for mixed forests used in Swedish and Finnish decision support systems. *Forest Ecology and Management*, vol. 529, 120721. <https://doi.org/10.1016/j.foreco.2022.120721>
28. Kaltsidis A., Ketikidis K., Basbas S., Aifadopoulou G., Grau J.M.S. (2023) A decision support system for taxi drivers. *Transportation Research Procedia*, vol. 69, pp. 123–130.
29. Siconolfi S.F. (2022) An analytics-based decision support system for evaluating the fiscal health of academic programs. *Decision Analytics Journal*, vol. 4, 100091. <https://doi.org/10.1016/j.dajour.2022.100091>
30. Borrero-Domínguez C., Escobar-Rodríguez T. (2023) Decision support systems in crowdfunding: A fuzzy cognitive maps (FCM) approach. *Decision Support Systems*, 114000. <https://doi.org/10.1016/j.dss.2023.114000>

About the author

Rimma D. Gutgarts

Dr. Sci. (Econ.);

Professor, Institute of Informatics Technologies and Data Analysis, Irkutsk National Research Technical University, 83, Lermontov St., Irkutsk 664074, Russia;

E-mail: gutgarc@gmail.com

ORCID: 0000-0001-9881-1976

DOI: 10.17323/2587-814X.2023.3.38.52

Применение мер «тяжелохвостости» в задачах анализа финансовых временных рядов

Л.А. Родионова 

E-mail: lrodionova@hse.ru

Е.Д. Копнова 

E-mail: ekopnova@hse.ru

Национальный исследовательский университет «Высшая школа экономики»

Адрес: Россия, 109028, г. Москва, Покровский бульвар, д. 11

Аннотация

Важной особенностью при работе с финансовыми данными является тот факт, что остатки GARCH-моделей часто имеют более «толстые хвосты», чем хвосты нормального распределения из-за большого числа «выбросов» в данных, что требует более детального изучения. В статье анализировались и сравнивались куртозис и мера «тяжелохвостости», основанная на квантилях, применительно к задаче выбора спецификации GARCH(1,1)-модели. Были рассмотрены данные индексов Мосбиржи за период 01.04.2019 по 22.02.2022 гг., значения куртозиса которых варьировались от 3 до 52. Как показали эмпирические данные, куртозис очень чувствителен к «выбросам» в данных, что затрудняло предположения относительно вида распределения остатков модели. Рассматриваемый в работе подход на основе меры «тяжелохвостости» позволил обосновать выбор числа степеней свободы t -распределения остатков модели для объяснения «толстых хвостов» в финансовых данных. Было получено, что часто встречающимися являются GARCH(1,1)-модели с $t(5)$ -распределением в остатках.

Ключевые слова: GARCH-модели, куртозис, «тяжелохвостость», t -распределение остатков, степени свободы, толстые хвосты

Цитирование: Родионова Л.А., Копнова Е.Д. Применение мер «тяжелохвостости» в задачах анализа финансовых временных рядов // Бизнес-информатика. 2023. Т. 17. № 3. С. 38–52. DOI: 10.17323/2587-814X.2023.3.38.52

Введение

Известен факт, что многие финансовые временные ряды имеют определенные закономерности: доходности активов слабо стационарны, наблюдается кластеризация волатильности, нормальность распределения отвергается в пользу распределения с толстыми хвостами и др. [1]. Для описания и прогнозирования процессов, обладающих такими свойствами, широко используется класс моделей с условной гетероскедастичностью (ARCH, GARCH-модели), предложенных Энглом [2] и Боллерслевом [3], и их модификации. Важной особенностью при работе с финансовыми данными, которую мы бы хотели подробно рассмотреть в настоящей статье, является тот факт, что остатки ARCH/GARCH моделей имеют более «толстые хвосты», чем хвосты нормального распределения из-за большого числа «выбросов» в данных, и этот факт требует более детального изучения. Для учета «толстых хвостов» в эконометрической практике было предложено несколько альтернативных распределений: t -распределение Стьюдента [3, 4], обобщенное распределение ошибок (GED) [5, 6], скошенное t -распределение Стьюдента [7] и др. Отметим, что возможность выбора t -распределения Стьюдента и GED-распределения при оценивании GARCH-модели реализована в эконометрических пакетах (например, Stata16), что представляет практический интерес в обосновании выбора соответствующего распределения при моделировании и прогнозировании. Предложенные распределения различаются по свойствам, следовательно, предполагаемые распределения не будут одинаково хорошо характеризовать «толщину» хвоста распределения. Возникает *проблема исследования*, как выбрать тип распределения, наилучшим образом характеризующего «тяжелохвостость» распределения. Правильная спецификация GARCH-модели с учетом «тяжелых хвостов» позволяет получить более точные прогнозы доходностей, а инвесторам в дальнейшем — максимальную прибыль, что обуславливает *актуальность* исследования.

Основная *цель* статьи состоит в анализе поведения меры «тяжелохвостости», основанной на квантилях, применительно к выбору степеней свободы t -распределения Стьюдента в остатках GARCH-модели. Отметим, что меры «тяжелохвостости» широко обсуждаются в зарубежной литературе и являются альтернативным подходом выбора числа степеней свободы t -распределения. Проверим,

насколько эти меры применимы в эконометрической практике при анализе финансовых данных и сравним с классическим подходом подбора степеней свободы t -распределения на основе сравнения оценок метода максимального правдоподобия.

1. Измерение «тяжелохвостости» распределения

В данном разделе разберем, какие подходы существуют для измерения «тяжести» хвоста распределения. Под «тяжелохвостостью» (heavy-tailedness) распределения для случайной величины (с.в.) X обычно понимают

$$P(|X| > x) \sim \frac{C}{x^\xi}, \quad (1)$$

где константы $C, \xi > 0$, а $f(x) \sim g(x)$, что означает:

$$\lim_{x \rightarrow \infty} \frac{f(x)}{g(x)} = 1.$$

Параметр ξ принято называть «*хвостовым индексом*» (tail index) распределения X , который характеризует скорость затухания хвостов степенного распределения (1) и вероятность наблюдения экстремальных значений с.в. Чем больше вероятностная масса в хвостах, тем меньше параметры индекса хвоста, и наоборот [8].

Заметим, что в литературе распределения с «тяжелыми хвостами» разделяют на три подкласса: распределения с «толстым хвостом», распределения с «длинным хвостом» и субэкспоненциальные распределения [9, 10]. Распределение с «толстым хвостом» (fat tail) демонстрирует большую асимметрию или эксцесс по сравнению с нормальным распределением, а в статьях по анализу финансов часто термины «толстый хвост» и «тяжелый хвост» используют как синонимы. В своей работе будем придерживаться термина «тяжелохвостость», а «толстые хвосты» распределения будем рассматривать как частный случай «тяжелых хвостов».

На практике возникает вопрос, как измерить «тяжелохвостость» вообще, и как оценить степень «тяжести» хвоста распределения. Существуют параметрические и непараметрические подходы к оцениванию хвостового индекса [11]. В своей статье проанализируем «тяжелохвостость» в контексте моделирования временных рядов на основе GARCH-моделей. Одной из мер, используемых для выявления «выбросов» во временных рядах является куртозис. В 1905 году Пирсон ввел понятие куртозиса через моменты 4-го порядка:

$$K = \frac{\mu_4}{\sigma^4}, \quad (2)$$

где μ_4 – центральный момент 4-го порядка;
 σ^4 – квадрат дисперсии;

и классифицировал все распределения как платокуртические, мезокуртические, или лептокуртические в зависимости от того, насколько «плосковершинным» является распределение по отношению к нормальному [12].

Для нормального распределения $K = 3$, в связи с чем часто используют модифицированный показатель эксцесса (excess kurtosis):

$$E_k = \frac{\mu_4}{\sigma^4} - 3.$$

Далее в статье будет использоваться куртозис K в виде (2). «Толстые хвосты» (как частный случай «тяжелых хвостов») характеризуются избыточным куртозисом $K > 3$, а распределение называют лептокуртическим (leptokurtic) [13].

В данной работе будем сравнивать и исследовать остатки GARCH(1,1)-моделей¹, как часто используемых в эконометрической практике [14]. Напомним определение. Процесс ε_t следует обобщенной модели авторегрессионной условной гетероскедастичности или GARCH(1,1)-модели, если $\varepsilon_t = \sigma_t z_t$, $t = 1, 2, \dots$, где $z_t \sim N(0,1)$ – независимые нормально распределенные случайные величины, а условная дисперсия процесса имеет вид:

$$\sigma_t^2 = \omega + \gamma \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2. \quad (3)$$

Отметим, что на практике модель (3) лишь отчасти объясняет «толстые хвосты» и необходимо уточнение спецификации распределения остатков. В качестве альтернативы нормальному распределению часто используют t -распределение Стьюдента [3, 4]. Стандартизированное t -распределение Стьюдента с нулевым средним и единичной дисперсией имеет плотность:

$$f(z_t, v) = \frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{1}{2}v\right)\sqrt{\pi(v-1)}} \left[1 + \frac{z_t^2}{v-2}\right]^{-\frac{v+1}{2}},$$

где $\Gamma(\cdot)$ – гамма-функция Эйлера;
 $v > 2$ – число степеней свободы.

Куртозис распределения z_t равен

$$K = 3 \frac{v-2}{v-4} \text{ при } v > 4.$$

Куртозис ошибки ε_t равен

$$K_u = 3 \frac{v-2}{v-4} \frac{E(\sigma_t^4)}{[E(\sigma_t^2)]^2}. \quad (4)$$

Множитель, содержащий v в выражении (4), позволяет учитывать избыточный куртозис [1]. Из формулы (4) видно, что куртозис t -распределения зависит от степени свободы и степени волатильности процесса. Таким образом, варьируя степени свободы t -распределения, можно получать разные значения куртозиса и разную степень «тяжелохвостости». Однако, куртозис, основанный на моментах 4-го порядка, очень чувствителен к выбросам, и, следовательно, при наличии выбросов может приводить к ложным выводам о характере распределения остатков.

Вопрос, измеряет ли куртозис «тяжелохвостость», и как на основе куртозиса определить, у какого распределения «тяжелее хвост», является дискуссионным. Если, например, на куртозис смотреть как на среднее выбросов, то большой куртозис свидетельствует о «больших тяжелых хвостах» [15]. Некоторые авторы описывают куртозис как меру и «островершинности», и «толщины хвоста» [16]. В целом существует три подхода к сравнению «тяжелохвостости» распределений: обычный куртозис K , меры «тяжести» хвостов и отношение межквантильных интервалов.

В статье рассмотрим один из подходов – меру «тяжелохвостости», основанную на квантилях, в виде:

$$K_\alpha(y_t) = \frac{Q_{1-\alpha}(y_t) - Q_\alpha(y_t)}{Q_{1-\tau}(y_t) - Q_\tau(y_t)}, \quad (5)$$

где $Q_\theta(y_t)$ – θ -квантиль y_t , $0 < \alpha < \tau < 0,5$ [16].

Следуя статье [17], выберем $\tau = 0,25$ и $\alpha = 0,01$, $\alpha = 0,05$. Считается, что квантильная мера (5) свободна от предположений о распределениях и от значений куртозиса, поэтому устойчива к неверным классификациям распределений и может быть

¹ Отметим, что авторы статьи допускают, что «тяжелохвостость» может быть результатом ошибок выбора спецификации модели, но в данной работе рассматривается только варьирование числа степеней свободы t -распределения для того, чтобы подобрать оптимальную модель с точки зрения минимизации информационных критериев. Подбор других спецификаций и модификаций GARCH-моделей остается за рамками статьи.

использована для сравнения распределений. Рассчитаем и изучим поведение куртозиса (2) и меры «тяжелохвостости» (5) для нормального распределения и t -распределения Стьюдента, как часто используемого распределения в эконометрической практике для учета «толстых хвостов», и сравним с распределениями эмпирических данных доходностей, что позволит сделать предположение о числе степеней свободы t -распределения для эмпирических данных при спецификации GARCH(1,1)-моделей.

2. Анализ тяжелых хвостов распределения в отечественных исследованиях

В данном разделе проанализируем, как учитывают тяжелые хвосты распределений на практике при прогнозировании финансовых рядов доходностей в исследованиях последних лет. Как было отмечено выше, для учета «толстых хвостов» в эконометрической практике часто используют t -распределение Стьюдента и его разновидности. Шведов [18] отмечал важность использования t -распределения для оценок ММП (метод максимального правдоподобия) в случае наличия в данных резко выделяющихся наблюдений. Автор сравнивал EM-алгоритм (англ. expectation-maximization (EM) algorithm) и МНК (метод наименьших квадратов) для оценок линейной регрессионной модели на сгенерированных данных при различных распределениях ошибок. Балаев [19] рассматривал и сравнивал двумерное t -распределение с вектором и скаляром степеней свободы, обобщенное распределение ошибки и распределение Грамма–Шарлье по данным дневных цен закрытия фондовых индексов различных стран: S&P 500, FTSE 100, CAC 40, DAX, Hang Seng, Nikkei за период 26.11.1990–18.11.2012. Автор отмечал, что распределения всех рассматриваемых доходностей имеют «тяжелые хвосты»: коэффициент эксцесса варьировался от 5,21 до 9,45. Было получено, что модель на основе t -распределения с вектором степеней свободы более предпочтительна. Работа Фантаццини [20] имеет обзорный характер и посвящена моделированию многомерных распределений на основе копул-функций. В работе введены понятия «верхней» и «нижней» «хвостовой зависимости» для случайных величин с некоторой вероятностью возникновения выбросов и рассмотрены копула-функции Стьюдента. Балаш [21] в своей работе анализировал

влияние внешних источников информации (новости и объемы торгов) на волатильность ценных бумаг с помощью моделей GARCH(1,1). На основе анализа коэффициентов асимметрии и куртозиса, было отмечено, что гипотеза о нормальности отвергается для большинства рассматриваемых ценных бумаг (доходности акций 19 компаний из списка FTSE100 за период 01.07.2005–01.07.2008). Эксцесс варьировался от 0 до 5,8. В ряде работ для моделирования доходностей авторы использовали класс специальных моделей Go-GARCH, GJR-GARCH, позволяющих оценивать параметр числа степеней свободы t -распределения наряду с другими параметрами модели [22, 23]. В работе Лакшиной [22] моделировались доходности с дальнейшим расчетом динамического коэффициента хеджирования для восьми акций российских компаний, торгующихся на РТС за период 01.01.2007–01.10.2014, эксцесс варьировался от 18 до 42. На основе модели GO-GARCH было рассчитано, что остатки распределены по Стьюденту с 2 степенями свободы.

Заметим, что авторы рассматриваемых работ в качестве «индикатора» «тяжелых хвостов» использовали коэффициент эксцесса, а вопрос о выборе числа степеней свободы t -распределения обычно оставался за рамками таких исследований. Также известно, что включение числа степеней свободы t -распределения в состав аргументов функции правдоподобия в ММП не всегда является корректным [24]. Таким образом, необходимы некоторые альтернативные характеристики «тяжести» хвостов, легко реализуемые на практике. В своей работе мы решили сделать попытку восполнить этот пробел.

3. Анализ меры «тяжелохвостости» и куртозиса для распределения Стьюдента на сгенерированных данных

Рассмотрим поведение меры «тяжелохвостости» в виде (5) и куртозиса (2) для $t(\nu)$ -распределения Стьюдента для разных степеней свободы ν . Методом Монте-Карло сгенерируем 5000 повторений по $N = 200, 750$ и 1000 наблюдений, рассчитаем и сравним куртозис K и меры «тяжелохвостости» $K_{0,1}$, $K_{0,5}$ для квантилей $\alpha = 0,01$, $\alpha = 0,05$ соответственно: интервал варьирования от минимума до максимума $(\min - \max)^2$ и среднее (mean). Результаты моделирования приведены в *таблице 1*. Случайные величины

² Отметим, что в работе также рассматривался 95% доверительный интервал мер, но интерпретация практически не изменилась.

генерировались в пакете Stata16. Генерирование псевдослучайных чисел реализовано на основе алгоритма, предложенного в работе [25]. В силу свойств t -распределения рассматривались степени свободы от 3 до 10, заметим, что теоретический куртозис существует для $\nu > 4$.

Значения мер из *таблицы 1* далее будут использованы для сравнения с мерами «тяжелохвостости» анализируемых доходностей индексов Мосбиржи для дальнейшей спецификации GARCH(1,1)-модели в выборе предположения о распределении остатков.

Таблица 1.

**Меры «тяжелохвостости» (5) и куртозис (2)
для $t(\nu)$ -распределения Стьюдента и нормального распределения**

	$K_{01min-max}$	K_{01mean}	$K_{05min-max}$	K_{05mean}	$K_{min-max}$	K_{mean}
$N = 1000$						
$t(3)$	4,8–7,7	5,96	2,71–3,62	3,08	5,7–464,07	29,883
$t(4)$	3,9–7,08	5,085	2,51–3,37	2,875	4,09–304,55	12,653
$t(5)$	3,76–5,61	4,645	2,41–3,17	2,772	3,88–115,6	7,550
$t(6)$	3,71–5,35	4,388	2,35–3,13	2,708	3,5–127,91	6,166
$t(7)$	3,57–5,02	4,224	2,29–3,11	2,662	3,35–26,54	4,679
$t(8)$	3,45–4,97	4,103	2,26–3,13	2,633	3,30–16,28	4,486
$t(9)$	3,37–4,76	4,016	2,28–3,13	2,609	3,04–13,85	4,186
$t(10)$	3,38–4,68	3,947	2,29–3,11	2,589	3,01–13,83	4,013
$N(0,1)$	3,01–4,07	3,467	2,18–2,74	2,442	2,57–3,84	3,004
$N = 750$						
$t(3)$	4,37–8,94	5,989	2,57–3,75	3,084	4,28–572,55	27,596
$t(4)$	3,81–7,52	5,080	2,39–3,49	2,881	3,34–614,12	12,881
$t(5)$	3,45–6,49	4,649	2,31–3,37	2,771	3,31–541,35	7,391
$t(6)$	3,39–5,82	4,397	2,27–3,25	2,707	3,26–147,19	5,672
$t(7)$	3,34–5,31	4,216	2,27–3,21	2,663	3,06–48,49	4,909
$t(8)$	3,21–5,42	4,105	2,23–3,25	2,631	3,02–55,67	4,411
$t(9)$	3,21–5,21	4,013	2,15–3,07	2,606	2,89–23,87	4,152
$t(10)$	3,12–5,11	3,951	2,18–3,20	2,591	2,89–24,09	3,975
$N(0,1)$	2,82–4,37	3,465	2,13–2,84	2,442	2,49–4,36	2,995
$N = 200$						
$t(3)$	3,87–12,45	6,259	2,25–4,31	3,108	3,28–137,78	13,783
$t(4)$	3,42–9,54	5,237	2,21–3,74	2,897	3,12–141,47	8,775
$t(5)$	3,17–7,75	4,756	2,14–3,73	2,790	2,78–62,72	6,123
$t(6)$	2,81–6,46	4,459	1,97–3,64	2,714	2,71–65,19	5,200
$t(7)$	2,97–7,72	4,307	1,99–3,75	2,670	2,52–21,59	4,648
$t(8)$	2,96–6,31	4,155	2,01–3,40	2,633	2,61–26,43	4,266
$t(9)$	2,96–5,56	4,058	2,05–3,62	2,611	2,51–26,48	4,086
$t(10)$	2,73–6,24	3,987	2,00–3,51	2,588	2,31–30,24	3,943
$N(0,1)$	2,72–4,92	3,504	2,01–3,23	2,446	2,32–4,28	3,007

Поведение мер «тяжелохвостости» – основные выводы:

- ♦ Меры «тяжелохвостости» $K_{0,1}$, $K_{0,5}$ (5), основанные на оценках квантилей, более устойчивы к выбросам по сравнению с куртозисом, рассчитываемым по оценкам моментов распределения. Меры лежат в фиксированных диапазонах в зависимости от степени свободы, в то время как значения куртозиса сильно варьируются при всех рассмотренных степенях свободы ν от 3 до 10. По значениям куртозиса довольно затруднительно делать предположение относительно степеней свободы t -распределения.
- ♦ Более информативны меры «тяжелохвостости» $K_{0,1}$ для квантиля 0,01, так как меры $K_{0,5}$ имеют много пересекающихся интервалов. В дальнейшем для обоснования выбора степеней свободы t -распределения ν будет использоваться мера $K_{0,1}$, а мера $K_{0,5}$ будет использоваться для контроля области значений меры по эмпирическим данным.
- ♦ Полученные интервалы варьирования мер «тяжелохвостости» для теоретических распределений в дальнейшей работе будут использованы для сравнения с мерами «тяжелохвостости» по эмпирическим данным и обоснования предположения относительно степеней свободы t -распределения остатков GARCH(1,1)-модели.
- ♦ Можно предположить, что подход сравнения мер «тяжелохвостости» хорошо работает на больших выборках ($N = 700$ и больше). Отметим, что в статье [17], рассчитывались также меры «тяжелохвостости» на основе дробно-квантильных интервалов (interfractile range) ($\alpha = 0,125$), что

требует выборку как минимум 1000 наблюдений. Это можно отнести к недостаткам данного подхода, так как экономические данные не всегда выражены длинными временными рядами.

4. Меры «тяжелохвостости» и значения куртозисов для индексов Мосбиржи

В работе были рассмотрены доходности (логарифмические разности) индексов Мосбиржи: основных и отраслевых за период с 01.04.2019 по 23.02.2022 гг. (732 торговых дня) [26]. Изучались цены открытия и закрытия (цена первой и последней транзакции в торговый день). На *рис. 1* представлен график доходностей индекса нефти и газа (цена закрытия). Ряд доходностей имеет неизменной нулевое среднее значение, наблюдается кластеризация волатильности, а период $t < 200$ (01.04.2019–20.02.2020) отмечен как период низкой волатильности. Динамика доходностей других индексов ведет себя похожим образом. Периоды высокой волатильности, как правило, характеризуются аномально высокими значениями (по модулю) доходностей, что ведет к высоким значениям куртозиса и появлению «толстых хвостов» распределения. Отметим, что для анализа были взяты индексы, имеющие различные значения куртозисов.

Введем обозначения в *таблицы 2 и 3*:

- ♦ период 1: 01.04.2019–23.02.2022 гг. ($N = 732$), период 2: 01.04.2019–20.02.2020 гг. ($N = 200$);
- ♦ K_1 , K_2 – куртозисы (2) доходностей индексов Мосбиржи за период 1 и 2 соответственно;

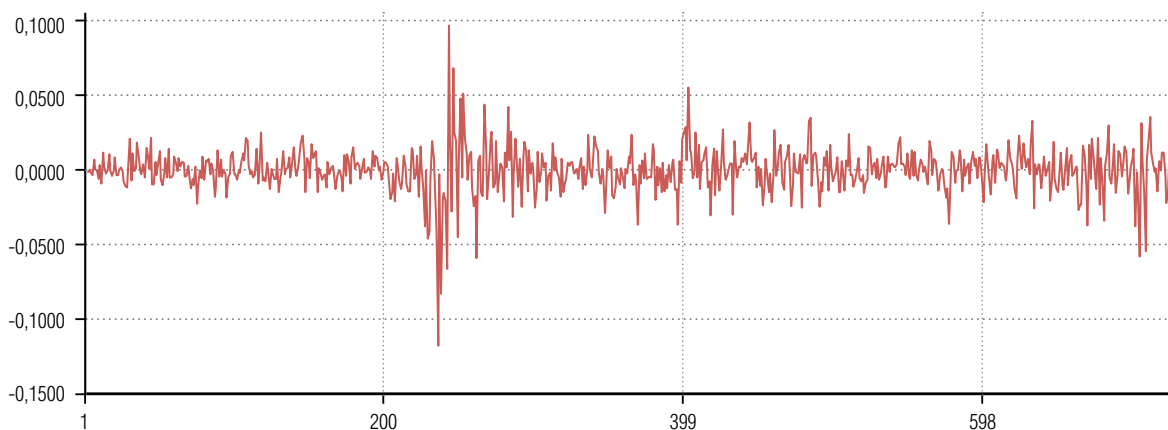


Рис. 1. Логарифмическая разность индекса нефти и газа (цены закрытия) (ось y) за период 04.2019–23.02.2022 (ось x).

- ♦ K_{101} , K_{105} и K_{201} , K_{205} – меры «тяжелохвостости» (5) для периодов 1 и 2 для квантилей 0,01 и 0,05 соответственно.

Таблицы 2 и 3 содержат значения куртозиса K_1 и K_2 доходностей индексов за два периода: период 1 (01.04.2019–23.02.2022) – весь период наблюдений с ярко выраженной кластеризацией волатильности, период 2 (01.04.2019–20.02.2020) – период низкой волатильности. Отметим, что для периода 1 характерно существенное различие в величине куртозиса между ценой открытия и закрытия почти в два раза, а куртозис варьируется от 6,77 до 51,87. В дальнейшем для моделирования использовалась

и цена открытия (в табл. 2, 3 переменные с индексом 2), и цена закрытия (в табл. 2, 3 переменные с индексом 1) как разные переменные. Очевидно, что величина куртозиса показателей для периода 1 свидетельствует о том, что нельзя использовать предположение о нормальности остатков в GARCH(1,1). В период 2 расхождения в величине куртозиса для цены открытия и закрытия незначительные, и для некоторых индексов совпадают с куртозисом нормального распределения ($K = 3$) (табл. 2, 3). Таким образом, дальнейший анализ и сравнение оценок GARCH(1,1)-моделей в работе проводился для периода 1.

Таблица 2.

Значения куртозисов (2) и мер «тяжелохвостости» (5) доходностей индексов (логарифмических разностей) для основных индексов

Индексы Мосбиржи		Период 1				Период 2		
		$K1$	$K1_{01}$	$K1_{05}$	Предполагаемое $t(\nu)$ для остатков	$K2$	$K2_{01}$	$K2_{05}$
blue1	Индекс МосБиржи голубых фишек	15,194	6,009	2,946	$t(3)-t(5)$	3,597	3,991	2,287
blue2		33,434	6,225	2,966	$t(3)-t(5)$	3,579	4,541	2,509
imoe1	Индекс МосБиржи	15,596	6,347	3,019	$t(3)-t(5)$	3,406	3,792	2,347
imoe2		32,844	6,763	2,993	$t(3)-t(4)$	3,392	4,041	2,319
rts1	Индекс РТС	14,432	6,716	3,179	$t(3)-t(4)$	4,982	4,547	2,411
rts2		29,391	6,508	3,087	$t(3)-t(5)$	4,177	5,333	2,789

Замечание: $var1$ – цена закрытия, $var2$ – цена открытия. Период 1: 01.04.2019–23.02.2022 ($N = 732$), период 2: 01.04.2019–20.02.2020 ($N = 200$).

Отметим, что меры «тяжелохвостости» $K2_{01}$ (5) большинства анализируемых индексов (85%), рассчитанные для периода 2 попадают в интервалы варьирования мер для нормального распределения ($N = 200$): 2,72–4,92 (табл. 1). В таблице 2 приведены меры «тяжелохвостости» $K2_{01}$ (5) некоторых переменных, рассчитанные для периода 2. Таким образом, можно предположить, что наилучшей моделью для периода с низкой волатильностью будут GARCH(1,1)-модели в предположении нормальности остатков. В дальнейшем анализе рассматривался только период 1.

Для периода 1 на основании рассчитанных мер были сделаны предположения относительно степеней свободы $t(\nu)$ -распределения остатков для дальнейшей спецификации GARCH(1,1)-моделей

путем сравнения рассчитанных мер «тяжелохвостости» со значением мер теоретических распределений (табл. 1). Например, индекс blue1 имеет меру «тяжелохвостости» $K1_{01} = 6,009$ (табл. 2), эта мера соответствует интервалу варьирования $K01_{\min-\max}$ ($N = 750$) для $t(3)$: 4,37–8,94; $t(4)$: 3,81–7,52; $t(5)$: 3,45–6,49 (табл. 1), следовательно, предполагаемыми распределениями остатков при оценивании GARCH(1,1)-моделей будут $t(3)-t(5)$. Рассмотрим, насколько хорошо рассмотренная мера «тяжелохвостости» позволяет верно специфицировать GARCH(1,1). Из таблиц 2, 3 видно, что меры «тяжелохвостости» не дают однозначного ответа, но позволяют сузить количество моделей, которые необходимо в дальнейшем оценивать.

Таблица 3.

**Значения куртозисов (2) и мер «тяжелохвостости» (5)
доходностей индексов (логарифмических разностей)
для отраслевых индексов**

Индексы Мосбиржи		Период 1				Период 2
		$K1$	$K1_{01}$	$K1_{05}$	Предполагаемое $t(\nu)$ для остатков	$K2$
gaz1	Нефти и газа	14,669	5,602	3,007	$t(3)-t(6)$	3,233
gaz2		27,428	6,221	3,388	$t(3)-t(5)$	3,529
chem1	Химии и нефтехимии, в рублях	6,772	5,785	3,528	$t(3)-t(6)$	5,527
chem2		10,558	6,034	3,270	$t(3)-t(5)$	7,789
electro1	Электроэнергетики, в рублях	21,194	6,604	3,107	$t(3)-t(4)$	3,519
electro2		36,633	6,775	2,937	$t(3)-t(4)$	4,044
telecom1	Телекоммуникаций, в рублях	13,667	6,373	3,109	$t(3)-t(5)$	7,295
telecom2		51,869	7,142	3,461	$t(3)-t(4)$	5,977
metal1	Металлов и добычи, в рублях	12,399	5,621	2,842	$t(3)-t(6)$	3,295
metal2		25,064	5,975	3,049	$t(3)-t(5)$	3,472
finan1	Финансов, в рублях	13,063	5,973	3,309	$t(3)-t(5)$	4,097
finan2		37,475	6,557	3,327	$t(3)-t(5)$	4,421
potreb1	Потребительского сектора, в рублях	12,399	5,621	2,842	$t(3)-t(6)$	3,295
potreb2		25,064	5,975	3,049	$t(3)-t(5)$	3,472
trans1	Транспорта, в рублях	16,832	7,081	3,457	$t(3)-t(4)$	5,569
trans2		18,288	7,712	3,368	$t(3)$	7,355

Замечание: $var1$ – цена закрытия, $var2$ – цена открытия. Период 1: 01.04.2019–23.02.2022 ($N = 732$), период 2: 01.04.2019–20.02.2020 ($N = 200$).

5. Результаты

В данном разделе сравним спецификации GARCH(1,1) для индексов Мосбиржи в предположении разных видов распределения в остатках: нормального и t -распределения со степенями свободы от 3 до 9. В качестве примера приведены результаты оценивания и сравнения моделей для показателя gaz1 в таблице 3. Для сравнения моделей использовался критерий отношения правдоподобия [27, с. 171] и сравнение информационных критериев Акаике и Шварца. Результаты дали один и тот же результат, поэтому далее в тексте приведены только значение максимума функции правдоподобия. В таблице 4 приведены оценки параметров γ и β GARCH(1,1)-модели в виде (3), значения LLF – значение максимума функции правдоподобия для

текущей модели. Отметим, что в Stata рассчитывается полная функция логарифма правдоподобия с включением слагаемых без оптимизационных параметров. Вид функции логарифма правдоподобия в предположении нормального и t -распределения приведен в [28]. Все коэффициенты модели статистически значимы на 1% уровне значимости.

Результаты показывают (табл. 4), что для различных спецификаций коэффициент $\beta \approx 0,84-0,89$, что свидетельствует об инерционности волатильности во времени, $(\gamma + \beta)$ превосходит 0,9, что указывает на наличие ярко выраженного GARCH-эффекта. Коэффициенты γ и β для различных спецификаций ведут себя достаточно устойчиво. Для случая $t(3)$ $(\gamma + \beta) > 1$, что нарушает условие положительности условной дисперсии модели. Также эта модель неадекватна с точки зрения моделирования «тя-

Таблица 4.
**Оценки параметров GARCH(1,1)-модели
для переменной gaz1
в предположении различных
степеней свободы t -распределения остатков**

Распределение остатков	γ	β	$\gamma + \beta$	LLF
N	0,155	0,836	0,991	2130,876
$t(3)$	0,136	0,892	1,028	2149,819
$t(4)$	0,110	0,884	0,993	2153,393
$t(5)$	0,104	0,879	0,982	2154,087
$t(6)$	0,102	0,875	0,978	2153,702
$t(7)$	0,103	0,871	0,974	2151,963
$t(8)$	0,103	0,871	0,974	2151,963
$t(9)$	0,104	0,869	0,973	2150,997

желых хвостов», так как куртозис t -распределения определен и больше 3 для степеней свободы $\nu > 4$. С точки зрения минимальных значений информационных критериев наилучшей моделью является модель с $t(5)$ -распределением в остатках, для которой $AIC = -4300,17$ и $BIC = -4281,79$, для данной модели также наблюдается максимум $LLF = 2154,09$. Наиболее предпочтительной моделью для прогнозирования волатильности будет GARCH(1,1) с предположением $t(5)$ -распределения в остатках. Отметим, что получившаяся модель соответствует предположениям относительно модели по мерам «тяжелохвостости»: $t(3)$ – $t(6)$ (табл. 3).

Отметим, что показатель gaz1 имеет куртозис $K1 = 14,67$, что свидетельствует о выбросах в данных и не позволяет использовать предположение о нормальности остатков в GARCH(1,1)-модели. GARCH(1,1) с нормальностью в остатках имеет самые большие значения AIC , BIC и минимальные LLF . Меры «тяжелохвостости» $K1_{01} = 5,60$ и $K1_{05} = 3,01$ попадают в интервалы варьирования для мер «тяжелохвостости» распределений $t(3)$ – $t(6)$ (табл. 1), что в данном случае совпало с результатами оценивания GARCH(1,1)-модели методом перебора.

Аналогичным образом оценивались GARCH(1,1)-модели для всех других показателей. В таблице 5

приведены наилучшие модели с максимумом LLF . Сравнение различных спецификаций GARCH(1,1)-моделей для каждого показателя приведены в приложении.

Как показал анализ оценок параметров GARCH(1,1)-модели для индексов Мосбиржи, наиболее распространенными и наилучшими моделями являются модели с предположением $t(5)$ -распределения в остатках. Такие спецификации модели составили 60%, при этом куртозис логарифмических доходностей показателей варьировался от 6 до 51 (табл. 2, 3). В Stata16 есть возможность оценивать GARCH(1,1)-модели в предположении $t(\nu)$ -распределения в остатках при $\nu \rightarrow \infty$ (например, можно выбрать $\nu = 1000$), но, как показывает проведенный анализ эмпирических данных, степени свободы ν t -распределений для рассмотренных показателей варьируются от 4 до 7 (табл. 5), а по мерам «тяжелохвостости» при $\nu = 10$ t -распределение приближается к нормальному. Предположения относительно степеней свободы на основе меры «тяжелохвостости» совпали с результатами эмпирического анализа методом перебора моделей в 68% случаев: исключением стали несколько индексов, например, trans2. Модели с предположением $t(3)$ -распределения в остатках, для которого не существует теоретического куртозиса, не оказались в числе наилучших. Также можно заметить, что нет ни одной модели с предположением нормальности остатков для периода 1, такие модели имели, как правило, наихудшие характеристики LLF (приложение).

Отметим, что в работе не рассматривались альтернативные подходы к оцениванию «тяжелохвостости» распределения, например, другие типы мер, что также представляет научный интерес для дальнейшего исследования. Еще одним возможным распределением при спецификации GARCH-моделей является GED-распределение, которое в данном исследовании не рассматривалось. Часто распределение финансовых показателей имеют асимметрию, что также необходимо учитывать при выборе спецификации GARCH-моделей, но нами рассмотрено не было. Как уже было отмечено выше, в работах также используют меры «тяжелохвостости» на основе дробно-квантильных интервалов (interfractile range), что требует выборку объема $N \geq 1000$ [17]. Эти меры также могут стать предметом дальнейшего исследования.

Таблица 5.

**Оценки параметров GARCH(1,1)-модели для индексов Мосбиржи
в предположении различных степеней свободы
 t -распределения остатков**

Индексы	Предполагаемое распределение остатков по мере «тяжелохвостости» (5)	Распределение остатков наилучшей модели по LLF	γ	β	$\gamma + \beta$	LLF
Основные индексы акций						
blue1	$t(3)-t(5)$	$t(6)$	0,115	0,869	0,984	2233,683
blue2	$t(3)-t(5)$	$t(5)$	0,123	0,861	0,984	2219,847
imoex1	$t(3)-t(5)$	$t(6)$	0,116	0,871	0,987	2295,943
imoex2	$t(3)-t(4)$	$t(7)$	0,112	0,863	0,975	2294,071
rts1	$t(3)-t(4)$	$t(5)$	0,097	0,901	0,998	2052,176
rts2	$t(3)-t(5)$	$t(5)$	0,121	0,871	0,992	2046,424
Отраслевые индексы						
gaz1	$t(3)-t(6)$	$t(5)$	0,104	0,879	0,982	2154,087
gaz2	$t(3)-t(5)$	$t(5)$	0,145	0,84	0,985	2109,548
chem1	$t(3)-t(6)$	$t(5)$	0,05	0,955	1,005	2345,464
chem2	$t(3)-t(5)$	$t(5)$	0,149	0,829	0,978	2313,152
electro1	$t(3)-t(4)$	$t(5)$	0,136	0,842	0,978	2350,902
electro2	$t(3)-t(4)$	$t(5)$	0,137	0,843	0,98	2350,902
telecom1	$t(3)-t(5)$	$t(4)$	0,119	0,864	0,983	2437,535
telecom2	$t(3)-t(4)$	$t(4)$	0,152	0,857	1,009	2418,431
metal1	$t(3)-t(6)$	$t(5)$	0,096	0,879	0,975	2275,479
metal2	$t(3)-t(5)$	$t(5)$	0,142	0,806	0,948	2262,375
finan1	$t(3)-t(5)$	$t(5)$	0,102	0,898	1,000	2167,337
finan2	$t(3)-t(5)$	$t(4)$	0,124	0,883	1,007	2129,232
potreb1	$t(3)-t(6)$	$t(4)$	0,104	0,881	0,985	2274,596
potreb2	$t(3)-t(5)$	$t(5)$	0,142	0,806	0,948	2262,375
trans1	$t(3)-t(4)$	$t(4)$	0,212	0,732	0,944	2251,718
trans2	$t(3)$	$t(6)$	0,301	0,668	0,969	2217,500

Заключение

В статье были рассмотрены обобщенные авторегрессионные модели условной гетероскедастичности для 22 индексов Мосбиржи (основных и отраслевых), имеющих различные значения куртозиса от 3 до 52, с целью изучения «тяжелохвостости» распределений и влияния куртозиса на выбор предположения о виде распределения в остатках модели для объяснения «толстых хвостов». Как по-

казал анализ, куртозис лишь отчасти является «индикатором» «толстых хвостов»: на его основании сложно делать предположение о виде распределения остатков, так как он чувствителен к выбросам. Так, например, куртозисы для chem1 и blue2 составили 6,77 и 33,43, но для этих показателей наилучшей моделью оказалась одна и та же спецификация модели – GARCH(1,1) с $t(5)$ -распределением в остатках. В работе было показано, что рассмотрен-

ные меры «тяжелохвостости» достаточно устойчивы к выбросам, позволяют частично обосновать выбор степени свободы для t -распределения при оценивании GARCH(1,1)-моделей. Следует заметить, что использование подхода сравнения моделей на основе оценок метода максимального правдоподобия дает похожие результаты в 68% случаев (табл. 5) и возможно в эконометрической практике для анализа финансовых временных рядов на выборках объема $N < 1000$ именно классический подход предпочтителен. Однако проведенный анализ мер «тяжелохвостости» имеет важное

практическое значение для моделирования временных рядов с «тяжелыми хвостами» и обоснования выбора степеней свободы t -распределения, так как куртозис не является хорошим количественным измерителем «тяжести» хвостов распределения. На взгляд авторов, меры «тяжелохвостости» и их свойства могут быть полезны широкому кругу исследователей, работающих с финансовыми временными рядами, для получения более точных прогнозов доходностей, а данная статья является небольшим вкладом в дальнейшее развитие инструментария анализа временных рядов. ■

Приложение.

Сравнение GARCH(1,1)-моделей для индексов Мосбиржи по LLF

Распределение остатков	Индексы										
	blue1	blue2	imoex1	imoex2	rts1	rts2	gaz2	chem1	chem2	electro1	electro2
N	2217,00	2189,29	2279,85	2277,39	2021,94	2017,52	2067,12	2321,80	2285,55	2321,49	2266,05
$t(3)$	2225,95	2213,30	2288,15	2284,26	2048,65	2041,62	2104,99	2342,08	2309,40	2347,75	2322,18
$t(4)$	2231,36	2218,40	2293,59	2290,61	2051,96	2045,63	2108,85	2345,10	2312,76	2350,74	2325,95
$t(5)$	2233,23	2219,85	2295,48	2293,09	2052,18	2046,42	2109,55	2345,46	2313,15	2350,90	2326,57
$t(6)$	2233,68	2219,83	2295,94	2293,95	2051,30	2046,02	2109,00	2344,89	2312,49	2350,07	2325,93
$t(7)$	2233,51	2219,18	2295,78	2294,07	2050,04	2045,15	2107,95	2343,97	2311,44	2348,89	2324,75
$t(8)$	2233,06	2218,27	2295,33	2293,83	2048,67	2044,10	2106,71	2342,94	2310,28	2347,63	2323,37
$t(9)$	2232,49	2217,26	2294,77	2293,41	2047,32	2043,02	2105,41	2341,91	2309,12	2346,38	2321,91

Распределение остатков	Индексы									
	telecom1	telecom2	metal1	metal2	finan1	finan2	potreb1	potreb2	trans1	trans2
N	2383,40	2355,57	2244,45	2216,95	2142,97	2077,17	2244,45	2259,31	2198,72	2192,53
$t(3)$	2435,90	2416,07	2270,33	2259,31	2166,45	2126,57	2270,33	2259,31	2250,95	2210,66
$t(4)$	2437,54	2418,43	2274,60	2262,29	2166,45	2129,23	2274,60	2262,29	2251,72	2215,79
$t(5)$	2436,19	2417,67	2275,48	2262,38	2167,34	2129,05	2275,48	2262,38	2249,93	2217,34
$t(6)$	2433,95	2415,87	2275,07	2261,43	2167,05	2127,84	2275,07	2261,43	2247,41	2217,50
$t(7)$	2431,52	2413,74	2274,14	2260,10	2166,30	2126,24	2274,14	2260,10	2244,78	2217,06
$t(8)$	2429,13	2411,56	2273,01	2258,66	2165,37	2124,53	2273,01	2258,66	2242,26	2216,37
$t(9)$	2426,88	2409,43	2271,83	2257,22	2164,41	2122,84	2271,83	2257,22	2239,91	2215,57

Замечание. Для gaz1 сравнение моделей приведено в табл. 4.

Литература

1. Росси Э. Одномерные GARCH-модели: обзор // Квантиль. 2010. № 8. С. 1–67.
2. Engle R. Autoregressive conditional heteroscedasticity with estimates of the variance of the United Kingdom inflation // *Econometrica*. 1982. Vol. 50. P. 987–1007. <https://doi.org/10.2307/1912773>
3. Bollerslev T. Generalized autoregressive conditional heteroskedasticity // *Journal of Econometrics*. 1986. Vol. 31. P. 307–327. [https://doi.org/10.1016/0304-4076\(86\)90063-1](https://doi.org/10.1016/0304-4076(86)90063-1)
4. Bollerslev T., Baillie R. The message in daily exchange rates: A conditional-variance tale // *Journal of Business & Economic Statistics*. 1989. Vol. 7(3). P. 297–305. <https://doi.org/10.2307/1391527>
5. Hsieh D.A. Modeling heteroscedasticity in daily foreign-exchange rates // *Journal of Business & Economic Statistics*. 1989. Vol. 7(3). P. 307–317. <https://doi.org/10.2307/1391528>
6. Nelson D. Conditional heteroskedasticity in asset returns: A new approach // *Econometrica*. 1991. Vol. 59(2). P. 347–370. <https://doi.org/10.2307/2938260>
7. Hansen B. Autoregressive conditional density estimation // *International Economic Review*. 1994. Vol. 35(3). P. 705–730. <https://doi.org/10.2307/2527081>
8. Ibragimov M., Ibragimov R., Walden J. Heavy-tailed distributions and robustness in economics and finance. Springer International Publishing, 2015.
9. Asmussen S.R. Steady-state properties of GI/G/1 // *Applied Probability and Queues. Stochastic Modelling and Applied Probability*. 2003. Vol. 51. P. 266–301. https://doi.org/10.1007/0-387-21525-5_10
10. Foss S., Korshunov D., Zachary S. An introduction to heavy-tailed and subexponential distributions. Springer Science & Business Media, 2013.
11. Novak S.Y. Extreme value methods with applications to finance. London: CRC, 2011.
12. Fiori A.M., Zenga M. Karl Pearson and the origin of kurtosis // *International Statistical Review*. 2009. Vol. 77(1). P. 40–50.
13. Цыплаков А. Мини-словарь англоязычных эконометрических терминов // Квантиль. 2008. № 5. С. 45–48.
14. Engle R. GARCH 101: The use of ARCH/GARCH models in applied econometrics // *Journal of economic perspectives*. 2001. Vol. 15(4). P. 157–168. <https://doi.org/10.1257/jep.15.4.157>
15. Hogg R.V. Adaptive robust procedures: A partial review and some suggestions for future applications and theory // *Journal of the American Statistical Association*. 1974. Vol. 69. P. 909–927.
16. Ruppert D. What is kurtosis?: An influence function approach // *The American Statistician*. 1987. Vol. 41(1). P. 1–5. <https://doi.org/10.2307/2684309>
17. Xiaochun L. On tail fatness of macroeconomic dynamics // *Journal of Macroeconomics*. 2019. Vol. 62(C). <https://doi.org/10.1016/j.jmacro.2019.103154>
18. Шведов А.С. Робастная регрессия с применением t-распределения и EM-алгоритма // *Экономический журнал Высшей школы экономики*. 2011. Т. 15. № 1. С. 68–87.
19. Балаев А.И. Анализ многомерных временных рядов финансовых доходностей: сравнение различных подходов к моделированию тяжелых хвостов // *Экономический журнал Высшей школы экономики*. 2013. Т. 17. № 2. С. 239–263.
20. Фантащини Д. Моделирование многомерных распределений с использованием копула-функций. II. // *Прикладная эконометрика*. 2011. №23(3). С. 98–132.
21. Балаш В.А., Сидоров С.П. Использование данных новостной аналитики в GARCH моделях // *Прикладная эконометрика*. 2013. № 1. С. 82–96.
22. Лакшина В.В. Динамическое хеджирование с учетом степени неприятия риска // *Экономический журнал Высшей школы экономики*. 2016. Т. 20. № 1. С. 156–174.
23. Асатуров К.Г., Теплова Т.В. Построение коэффициентов хеджирования для высоколиквидных акций российского рынка на основе моделей класса GARCH // *Экономика и математические методы*. 2014. Т. 50. № 1. С. 37–54.
24. Lucas A. Robustness of the student-t based m-estimator // *Communications in Statistics – Theory and Methods*. 1997. Vol. 26(5). P. 1165–1182. <https://doi.org/10.1080/03610929708831974>
25. Matsumoto M., Nishimura T. Mersenne twister: A 623-dimensionally equidistributed uniform pseudorandom number generator // *ACM Transactions on Modeling and Computer Simulation*. 1998. Vol. 8. P. 3–30.
26. Индексы Мосбиржи. [Электронный ресурс] <https://www.moex.com/ru/indices> (дата обращения 10.12.22).
27. Айвазян С.А. Методы эконометрики. М.: Магистр, 2010.
28. Документация Stata. [Электронный ресурс] <https://www.stata.com/manuals/tsarch.pdf> (дата обращения 10.12.22).

Об авторах

Родионова Лилия Анатольевна

кандидат экономических наук;

доцент, департамент статистики и анализа данных, факультет экономических наук, Национальный исследовательский университет «Высшая школа экономики», Россия, 109028, г. Москва, Покровский бульвар, д. 11;

E-mail: lrodionova@hse.ru

ORCID: 0000-0002-0310-6359

Копнова Елена Дмитриевна

кандидат технических наук;

доцент, департамент статистики и анализа данных, факультет экономических наук, Национальный исследовательский университет «Высшая школа экономики», Россия, 109028, г. Москва, Покровский бульвар, д. 11;

E-mail: ekopnova@hse.ru

ORCID: 0000-0002-8429-141X

Application of measures of heavy-tailedness in problems for analysis of financial time series

Lilia A. Rodionova

E-mail: lrodionova@hse.ru

Elena D. Kopnova

E-mail: ekopnova@hse.ru

HSE University

Address: 11, Pokrovsky Bulvar, Moscow 109028, Russia

Abstract

An important feature when working with financial data is the fact that the residuals of GARCH-models often have fatter tails than the tails of a normal distribution due to the large number of “outliers” in the data. This requires more detailed study. Kurtosis and quantile-based measure of heavy-tailedness were analyzed and compared in the article in relation to the problem of choosing the GARCH(1,1)-model specification. The data of indices of the Moscow Exchange were considered for the period from April 01, 2019 to February 22, 2022. Kurtosis values ranged from 3 to 52. Empirical data showed that kurtosis was very sensitive to “outliers” in the data, which made it difficult to make assumptions about the distribution of model residuals. The approach considered in this paper based on the heavy-tailedness measure made

it possible to justify the choice of degrees of freedom of the t-distribution for the model residuals to explain the fat tails in financial data. It was found that GARCH(1,1)-models with t(5)-distribution in the residuals are common.

Keywords: GARCH, kurtosis, quantile-based measure of heavy-tailedness, t-distribution of residuals, degrees of freedom, fat tails

Citation: Rodionova L.A., Kopnova E.D. (2023) Application of measures of heavy-tailedness in problems for analysis of financial time series. *Business Informatics*, vol. 17, no. 3, pp. 38–52. DOI: 10.17323/2587-814X.2023.3.38.52

References

1. Rossi E. (2010) One-dimensional GARCH models: A review. *Quantile*, vol. 8, pp. 1–67 (in Russian).
2. Engle R. (1982) Autoregressive conditional Heteroscedasticity with estimates of the variance of the United Kingdom inflation. *Econometrica*, vol. 50, pp. 987–1007. <https://doi.org/10.2307/1912773>
3. Bollerslev T. (1986) Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, vol. 31, pp. 307–327. [https://doi.org/10.1016/0304-4076\(86\)90063-1](https://doi.org/10.1016/0304-4076(86)90063-1)
4. Bollerslev T., Baillie R. (1989) The message in daily exchange rates: A conditional-variance tale. *Journal of Business & Economic Statistics*, vol. 7(3), pp. 297–305. <https://doi.org/10.2307/1391527>
5. Hsieh D.A. (1989) Modeling heteroscedasticity in daily foreign-exchange rates. *Journal of Business & Economic Statistics*, vol. 7(3), pp. 307–317. <https://doi.org/10.2307/1391528>
6. Nelson D. (1991) Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, vol. 59, no. 2, pp. 347–370. <https://doi.org/10.2307/2938260>
7. Hansen B. (1994) Autoregressive conditional density estimation. *International Economic Review*, vol. 35, no. 3, pp. 705–730. <https://doi.org/10.2307/2527081>
8. Ibragimov M., Ibragimov R., Walden J. (2015) *Heavy-tailed distributions and robustness in economics and finance*. Springer International Publishing.
9. Asmussen S.R. (2003) Steady-state properties of GI/G/1. *Applied Probability and Queues. Stochastic Modelling and Applied Probability*, vol. 51, pp. 266–301. https://doi.org/10.1007/0-387-21525-5_10
10. Foss S., Korshunov D., Zachary S. (2013) *An introduction to heavy-tailed and subexponential distributions*. Springer Science & Business Media.
11. Novak S.Y. (2011) *Extreme value methods with applications to finance*. London: CRC.
12. Fiori A.M., Zenga M. (2009) Karl Pearson and the origin of kurtosis. *International Statistical Review*, vol. 77(1), pp. 40–50.
13. Tsyplakov A. (2008) Mini-dictionary of English econometric terms. *Quantile*, vol. 5, pp. 45–48 (in Russian).
14. Engle R. (2001) GARCH 101: The use of ARCH/GARCH models in applied econometrics. *Journal of economic perspectives*, vol. 15, no 4, pp. 157–168. <https://doi.org/10.1257/jep.15.4.157>
15. Hogg R.V. (1974) Adaptive robust procedures: A partial review and some suggestions for future applications and theory. *Journal of the American Statistical Association*, vol. 69, pp. 909–927.
16. Ruppert D. (1987) What is kurtosis?: An influence function approach. *The American Statistician*, vol. 41(1), pp. 1–5. <https://doi.org/10.2307/2684309>
17. Xiaochun L. (2019) On tail fatness of macroeconomic dynamics. *Journal of Macroeconomics (electronic journal)*, vol. 62(C). <https://doi.org/10.1016/j.jmacro.2019.103154>
18. Shvedov A.S. (2011) Robust regression using t-distribution and EM-algorithm. *The HSE Economic Journal*, vol. 1, pp. 68–87 (in Russian).
19. Balaev A.I. (2013) Multivariate time series analysis of financial returns: Comparison of different heavy tail modeling approaches. *The HSE Economic Journal*, vol. 2, pp. 239–263 (in Russian).
20. Fantazzini D. (2011) Modeling multivariate distributions using copula functions. II. *Applied Econometrics*, vol. 3, pp. 98–132 (in Russian).

21. Balash V.A., Sidorov S.P. (2013) Using news analytics data in GARCH models. *Applied Econometrics*, vol. 1, pp. 82–96 (in Russian).
22. Lakshina V.V. (2016) Dynamic hedging with risk aversion. *The HSE Economic Journal*, vol. 1, pp. 156–174 (in Russian).
23. Asaturov K.G., Teplova T.V. (2014) Construction of hedging ratios for highly liquid shares of the Russian market based on GARCH class models. *Economics and mathematical methods*, vol. 1, pp. 37–54 (in Russian).
24. Lucas A. (1997) Robustness of the Student-t based M-estimator. *Communications in Statistics – Theory and Methods*, vol. 26, no. 5, pp. 1165–1182. <https://doi.org/10.1080/03610929708831974>
25. Matsumoto M., Nishimura T. (1998) Mersenne Twister: A 623-dimensionally equidistributed uniform pseudorandom number generator // *ACM Transactions on Modeling and Computer Simulation*, vol. 8, pp. 3–30.
26. *Moscow Exchange Indices*. Available at: <https://www.moex.com/ru/indices> (accessed 10 December 2022).
27. Ayvazyan S.A. (2010) *Methods of econometrics*. Moscow: Magistr (in Russian).
28. *State documentation*. Available at: <https://www.stata.com/manuals/tsarch.pdf> (accessed 10 December 2022).

About the authors

Lilia A. Rodionova

Cand. Sci. (Econ.);

Associate Professor, Department of Statistics and Data Analysis, HSE University, 11, Pokrovsky Bulvar, Moscow 109028, Russia;

E-mail: lrodionova@hse.ru

ORCID: 0000-0002-0310-6359

Elena D. Kopnova

Cand. Sci. (Tech.);

Associate Professor, Department of Statistics and Data Analysis, HSE University, 11, Pokrovsky Bulvar, Moscow 109028, Russia;

E-mail: ekopnova@hse.ru

ORCID: 0000-0002-8429-141X

DOI: 10.17323/2587-814X.2023.3.53.69

Формирование каузального поля показателей развития интеллектуального капитала организации: концепция и нечеткая экономико-математическая модель

Г.С. Завалин 

E-mail: georgiy.zavalin08@vvsu.ru

О.В. Недолужко 

E-mail: olga.nedoluzhko25@gmail.com

К.С. Солодухин 

E-mail: k.solodukhin@mail.ru

Владивостокский государственный университет

Адрес: Россия, 690014, г. Владивосток, ул. Гоголя, д. 41

Аннотация

Развитие теории интеллектуального капитала за счет введения в нее концепции имплицитности предполагает рассмотрение интеллектуального капитала как имплицитного фактора, процесс формирования которого в значительной степени определяется влиянием специфических скрытых факторов, чье воздействие выражено неявно и трудно формализуется. В настоящее время процесс отбора эксплицитных и имплицитных факторов, воздействующих на интеллектуальный капитал, в отечественных и зарубежных исследованиях не формализован, что обуславливает актуальность настоящей работы. Цель данного исследования состояла в разработке схемы отбора эксплицитных и имплицитных факторов развития интеллектуального капитала организации в увязке с ее стратегией на основе модифицированной сбалансированной системы показателей с учетом распределения показателей по типам когнитивной активности. Операционализация данной схемы осуществлена путем разработки нечеткой экономико-математической модели, пригодной к практическому применению. Основной особенностью модели является возможность нечеткого задания «границ отсечения» эксплицитных и имплицитных факторов. Представлены результаты апробации модели на примере крупного регионального университета. Приведены наборы эксплицитных и имплицитных факторов интеллектуального капитала университета для различных «границ отсечения» с использованием различных методов дефазификации.

Ключевые слова: когнитивная активность, нечеткая модель, имплицитный фактор, эксплицитный фактор, интеллектуальный капитал, стратегический менеджмент

Цитирование: Завалин Г.С., Недолужко О.В., Солодухин К.С. Формирование каузального поля показателей развития интеллектуального капитала организации: концепция и нечеткая экономико-математическая модель // Бизнес-информатика. 2023. Т. 17. № 3. С. 53–69. DOI: 10.17323/2587-814X.2023.3.53.69

Введение

В условиях формирования экономики знаний основными источниками конкурентных преимуществ организации являются невоспроизводимые факторы производства, в том числе, интеллектуальный капитал (далее – ИК) организации. ИК является инструментальным стержнем экономики знаний. Характер развития ИК в значительной степени определяется влиянием специфических скрытых факторов, чье воздействие на процесс развития выражено неявно и трудно формализуется. Данное обстоятельство обуславливает необходимость выявления таких факторов [1–3].

Развитие теории интеллектуального капитала за счет введения в нее концепции имплицитности позволило обеспечить объяснение и интерпретацию бизнес-процессов экономических систем на принципиально новом уровне обобщения. В рамках теории интеллектуальный капитал сам является имплицитным фактором, процесс формирования которого в значительной степени определяется воздействием ряда имплицитных факторов, влияющих на него [4]. Концепция имплицитности изначально была сформулирована в когнитивной психологии [5], получила свое развитие в лингвистике [6, 7], а на современном этапе нашла практическое применение в экономических науках [8, 9].

Под имплицитными факторами в рамках концепции понимаются неявные факторы, оказывающие существенное влияние на бизнес-процессы хозяйствующего субъекта, в основе которых лежит скрытая информация [10]. В условиях экономики знаний, когда наиболее значительным становится влияние информации как наиболее важного ресурса, влияние имплицитных факторов в системе управления деятельностью организации усиливается [11, 12]. Соответственно, представляется целесообразным выделить, кроме явно воздействующих на развитие ИК факторов (эксплицитных), факторы скрытого, опосредованного влияния (имплицитные). В совокупности, выделенные группы

ключевых показателей развития ИК, эксплицитных, или очевидных факторов, оказывающих непосредственное воздействие на развитие ИК, а также имплицитных факторов составляют каузальное поле показателей развития ИК организации [13].

Развитие ИК как источника конкурентных преимуществ организации осуществляется в рамках стратегического управления организацией. Важнейшим инструментом структуризации и операционализации стратегии является сбалансированная система показателей (ССП – Balanced ScoreCard, BSC). Данный системный метод управления, предложенный Р. Капланом и Д. Нортеном, позволяет перевести сформулированные стратегические цели и задачи организации, учитывающие все аспекты ее дальнейшего развития, в конкретные действия [14–16]. За свою тридцатилетнюю историю концепция СПП претерпела значительную эволюцию, не только не растеряв, но упрочив свою популярность. Ее используют практически все известные консалтинговые компании, а все крупные разработчики корпоративных информационных систем предлагают инструментальную поддержку СПП [17–19].

Возможность использования СПП в отношении оценки ИК организаций обусловлена, во-первых, акцентом на нематериальные показатели, во-вторых, прослеживающейся взаимосвязью между традиционно выделяемыми основными структурными компонентами ИК (человеческий капитал, организационный капитал, отношенческий капитал) и перспективами СПП [15, 20–23].

Не смотря на признанные достоинства, СПП не свободен от недостатков, подвергаясь критике на протяжении своей эволюции [13, 24–27]. Выделим два из них, критичные с точки зрения целей настоящей работы.

Во-первых, в традиционной модели СПП не учитываются опосредованные влияния имплицитных факторов на ключевые показатели деятельности организации. Причинно-следственные связи в стратегических картах отражают факторы прямого влияния (явные, очевидные зависимости).

Во-вторых, в классической версии ССП изначально заложено неравноправие стейкхолдеров организации с точки зрения учета их интересов. Между тем, важнейшим структурным компонентом ИК является отношенческий капитал, определяемый характером взаимосвязей организации с субъектами внешней среды [28–32].

Решение первой проблемы предлагает Д.М. Назаров в работе [13], разработав модель рефлексивного отбора имплицитных показателей управленческой деятельности организации и применив ее к разработке модифицированной ССП. В свою очередь, в работах [33, 34] предлагается модификация ССП для так называемой стейкхолдер-компания. В ее рамках предлагается, в том числе, метод построения карты стратегических целей, в котором устранено изначальное неравноправие интересов стейкхолдеров, заложенное в классической ССП. С точки зрения целей настоящей работы представляется перспективным совместить описанные модификации ССП. А именно применить модель рефлексивного отбора имплицитных показателей в рамках «стейкхолдерской» модификации ССП.

Важно отметить, что процесс создания различных видов ИК организации реализуется за счет его фундаментальной характеристики – когнитивной активности. Когнитивная активность обеспечивает необходимое разнообразие способов восприятия и переработки информации и их проявление через совокупность специфических психических процессов и состояний (познавательную деятельность) [35–37]. Определение возможных типов когнитивной активности (обучение, вовлечение, производственная рационализация, самосовершенствование, клиентоориентированная рационализация, инновационная деятельность) дает возможность осуществлять конкретные управленческие воздействия на них на различных уровнях. Типы когнитивной активности могут быть соотнесены со структурными компонентами ИК следующим образом: обучение и самосовершенствование способствуют развитию человеческого капитала, вовлечение и производственная рационализация развивают организационный капитал, инновационная деятельность и клиентоориентированная рационализация обеспечивают прирост отношенческого капитала.

Неявность имплицитных факторов, скрытость и опосредованность их влияния на развитие ИК (в свою очередь, являющегося имплицитным фак-

тором) приводит к необходимости использования нечетких инструментов при их выявлении. Значительным преимуществом использования нечетких моделей и методов является возможность формализации различного рода неопределенностей (прежде всего, лингвистической неопределенности). В научной литературе нам не удалось найти нечеткие модели выявления имплицитных факторов ИК. В то же время, существуют работы, в которых предлагаются нечеткие инструменты применительно к самым различным имплицитным факторам социально-экономических систем [38–41].

В работе [4] предложена нечеткая модель выявления имплицитных факторов в ССП организации. Выявление опосредованных влияний в рамках модели основано на технологии оценки нечетких бинарных отношений на некотором множестве. При этом элементами матриц нечетких бинарных отношений являются одноточечные нечеткие множества, что в определенной степени сужает возможности использования модели. Представляется перспективным развитие данной модели применительно к ИК в рамках новой модификации ССП в привязке к основным структурным компонентам ИК с учетом распределения показателей по типам когнитивной активности с изменением технологии оценки нечетких бинарных отношений.

Таким образом, целью настоящего исследования является разработка метода отбора эксплицитных и имплицитных факторов развития интеллектуального капитала организации в увязке с ее стратегией на основе модифицированной сбалансированной системы показателей с учетом распределения показателей по типам когнитивной активности в нечеткой постановке.

1. Метод формирования каузального поля показателей развития ИК

Формирование каузального поля показателей развития ИК организации предполагает выделение трех групп показателей:

1. ключевые показатели развития ИК;
2. эксплицитные факторы ИК (оказывающие очевидное непосредственное влияние на развитие ИК);
3. имплицитные факторы ИК (неявные факторы, оказывающие опосредованное влияние на развитие ИК).

Формирование каузального поля предлагается осуществлять в рамках модифицированной ССП. При этом предлагается сгруппировать стратегические цели организации, имеющие существенное отношение к развитию ее ИК, по трем группам, соответствующим основным структурным компонентам ИК. Заметим, что одна и та же цель может попасть в несколько групп. При этом интегральные показатели, соответствующие основным структурным компонентам ИК, могут быть рассмотрены как ключевые показатели ИК.

Поскольку, как было отмечено выше, с каждым структурным компонентом ИК могут быть соотнесены два типа когнитивной активности, то фактически происходит группировка стратегических целей по шести группам.

Предварительный отбор показателей, претендующих на включение в группы «эксплицитные факторы ИК» и «имплицитные факторы ИК» происходит среди показателей стратегических целей из шести групп. Концепция ССП предполагает соответствие каждой стратегической цели набора результирующих показателей, значения которых позволяют судить о степени достижения цели. Стратегические цели, которые способствуют, в той или иной мере, развитию ИК организации, могут быть направлены на развитие иных ключевых аспектов деятельности организации. Поэтому не все результирующие показатели этих целей будут показателями развития ИК.

Сформированный набор показателей необходимо разделить на три подгруппы: эксплицитные факторы ИК; имплицитные факторы ИК; показатели, влиянием которых на развитие ИК можно пренебречь (для конкретной организации в рамках ее стратегии на данном этапе ее развития).

Для этого на первом этапе необходимо оценить влияние всех отобранных показателей на ключевые показатели ИК. Под влиянием показателя на ИК будем понимать интегральную степень влияния данного показателя на ключевые показатели ИК. Все показатели, степень влияния которых на ИК превосходит определенную границу, будем относить к эксплицитным факторам ИК.

На втором этапе необходимо оценить влияние всех оставшихся показателей на уже отобранные эксплицитные факторы ИК. Здесь мы, вслед за Д.М. Назаровым [2], принимаем гипотезу о том, что имплицитные факторы влияют на основные показатели деятельности организации опосредованно. Причем, в роли опосредованных показателей выступают эксплицитные факторы. Соответствен-

но, влияние оставшихся показателей на развитие ИК может быть оценено как суперпозиция влияний этих показателей на эксплицитные факторы ИК и эксплицитных факторов ИК на ключевые показатели ИК. Все показатели, степень итогового (опосредованного) влияния которых на ИК превосходит определенную границу, будем относить к имплицитным факторам ИК. Заметим, что в общем случае «границы отсечения» при отборе эксплицитных и имплицитных факторов могут не совпадать. Будем считать, что влиянием на ИК показателей, оставшихся после отбора эксплицитных и имплицитных факторов, можно пренебречь.

В целом, базовая схема формирования каузального поля показателей развития ИК приведена на рисунке 1.



Рис. 1. Базовая схема формирования каузального поля показателей развития ИК организации.

2. Нечеткая модель

Пусть $C = \{c_1, c_2, \dots, c_k\}$ – множество ключевых показателей развития ИК,

$E = \{e_1, e_2, \dots, e_j\}$ – множество показателей стратегических целей, влияющих на развитие ИК,

$B = \{b_1, b_2, \dots, b_m\}$ – множество эксплицитных факторов ИК,

$A = \{a_1, a_2, \dots, a_n\}$ – множество имплицитных факторов ИК,

$D = \{d_1, d_2, \dots, d_s\}$ – множество факторов, влиянием которых на развитие ИК можно пренебречь.

Таким образом, $E = B \cup A \cup D$, причем $B \cap A \cap D = \emptyset$, то есть $t = m + n + s$. Степени влияния показателей множества E на показатели множества C определяются экспертно в заданной лингвистической шкале. В таблице 1 приведена возможная лингвистическая шкала и соответствующие лингвистическим переменным функции принадлежности нечетких множеств.

Таблица 1.

Терм-множество лингвистической переменной «влияние показателя e_i на показатель c_j »

Значение лингвистической переменной	Трапецевидная функция принадлежности
Очень слабое	$\langle 0; 0; 0,5; 1,5 \rangle$
Слабое	$\langle 0,25; 1,0; 1,5; 2,75 \rangle$
Среднее	$\langle 1,0; 2,0; 3,0; 4,0 \rangle$
Сильное	$\langle 2,25; 3,5; 4,0; 4,75 \rangle$
Очень сильное	$\langle 3,5; 4,5; 5,0; 5,0 \rangle$

Ответы экспертов следует проверить на согласованность и усреднить. При этом каждому эксперту может быть присвоен четкий или нечеткий весовой коэффициент, отражающий его уровень компетенции.

В результате имеем матрицу M_{EC} размерности $t \times k$, элементами которой являются нечеткие числа. Заметим, что элементами данной и последующих нечетких матриц могут быть нечеткие числа произвольного типа (не обязательно одноточечные).

Поставим в соответствие матрице M_{EC} вектор-столбец M_{EC}^* длины t следующим образом:

$$(M_{EC}^*)_i = \sum_{j=1}^k w_j (M_{EC})_{ij}, \quad (1)$$

где w_j – весовые коэффициенты ключевых показателей развития ИК. Заметим, что в общем случае коэффициенты w_j могут быть нечеткими (в частном случае можно считать $w_1 = w_2 = w_3 = 1/3$). Элементы вектор-столбца M_{EC}^* определяют влияние показателей множества E на ИК.

Тогда эксплицитными факторами ИК будем считать показатели e_i , для которых $(M_{EC}^*)_i$ превосходят экзогенно заданную «границу отсечения». «Граница отсечения» эксплицитных факторов в общем случае может быть задана нечетко. В этом случае необходимо воспользоваться одним из существующих методов сравнения нечетких множеств [42]. Если «граница отсечения» задана четко, то нечеткие элементы вектор-столбца M_{EC}^* могут быть дефазифицированы, после чего полученные четкие числа можно будет сравнивать с четкой «границей отсечения».

Заметим, что традиционно «граница отсечения» эксплицитных факторов задается вербально. Например, под эксплицитными факторами ключевого показателя деятельности организации обычно понимаются показатели, влияние которых «сильное» или «очень сильное». Иногда (реже) к ним добавляются также показатели со «средним» влиянием. В этом случае, под нечеткой «границей отсечения» следует понимать нечеткое множество с функцией принадлежности, соответствующей заданной вербальной оценке.

Обозначим через $F = \{f_1, f_2, \dots, f_{n+s}\}$ множество показателей стратегических целей, не являющихся эксплицитными факторами. То есть $F = E \setminus B = A \cup D$.

Определим экспертно в заданной лингвистической шкале степени влияния показателей множества F на показатели множества B . В результате имеем матрицу M_{FB} размерности $(t - m) \times m$, элементами которой являются нечеткие числа.

Рассмотрим матрицу M_{BC} размерности $m \times k$, полученную из матрицы M_{EC} удалением строк, соответствующих показателям множества F . Элементы матрицы M_{BC} отражают степени влияния эксплицитных факторов на ключевые показатели развития ИК.

Обозначим через M'_{FC} матрицу, полученную в результате произведения матриц M_{FB} и M_{BC} :

$$(M'_{FC})_{ij} = \sum_k (M_{FB})_{ik} \cdot (M_{BC})_{kj}. \quad (2)$$

Произведение и сложение элементов матриц в этом случае осуществляется по заданным правилам произведения и сложения нечетких чисел.

Существуют два основных подхода к реализации нечетких арифметических операций: подход на основе α -уровней с использованием интервальной арифметики, подход по принципу расширения с использованием различных t -норм. Для трапециевидных нечетких чисел в рамках первого подхода можно воспользоваться известными формулами сложения и произведения [43].

Существуют более сложные способы реализации нечеткой арифметики с использованием вычислительных методов, устраняющие недостатки двух основных подходов (завышение неопределенности в результирующих нечетких числах в рамках первого подхода и высокая чувствительность к изменениям входных нечетких чисел в рамках второго подхода). Однако в некоторых случаях сложность выполнения вычислительных операций в рамках данных методов может быть недопустимо высокой. В этой связи существуют упрощения процедуры арифметических операций над нечеткими числами определенных типов, в том числе, трапециевидными [44]. В работе [45] предложена унифицированная система правил выполнения арифметических операций над нечеткими числами (L - R)-типа.

Заметим, что при использовании простейших формул сложения и произведения трапециевидных нечетких чисел, средневзвешенные экспертные оценки будут также являться трапециевидными нечеткими числами. В то же время, при использовании упомянутой выше системы правил выполнения арифметических операций над нечеткими числами (L - R)-типа, средневзвешенные экспертные оценки могут иметь экспоненциальные (гауссовы) функции принадлежности (точнее, функции принадлежности получаемых нечетких множеств очень хорошо аппроксимируются гауссианами).

Нормируем при необходимости элементы матрицы M'_{FC} таким образом, чтобы универсальное множество получившихся нечетких чисел совпадало с исходным универсальным множеством (в нашем случае $[0; 5]$). Полученную матрицу обозначим M_{FC} .

Элементы матрицы M_{FC} отражают степени влияния показателей множества F на ключевые показатели развития ИК.

Поставим в соответствие матрице M_{FC} вектор-столбец M_{FC}^* длины t следующим образом:

$$(M_{FC}^*)_i = \sum_{j=1}^k w_j (M_{FC})_{ij}. \quad (3)$$

Элементы вектор-столбца M_{FC}^* определяют влияние показателей множества F на ИК.

Тогда имплицитными факторами ИК будем считать показатели f_i , для которых $(M_{FC}^*)_i$ превосходят экзогенно заданную «границу отсечения». «Граница отсечения» имплицитных факторов также может быть задана нечетко и в общем случае не совпадает с «границей отсечения» эксплицитных факторов.

3. Апробация модели

Предложенная модель была апробирована на примере крупного регионального университета (Владивостокского государственного университета, ВВГУ). В ВВГУ разработана стратегия развития университета, формализованная в виде стратегических карт в соответствии со «стейкхолдерской» модификацией ССП. Стратегические цели, имеющие существенное отношение к развитию ИК университета, были сгруппированы по шести группам в соответствии с типами когнитивной активности (табл. 2).

Среди результирующих показателей выбранных целей были отобраны показатели, значения которых позволяют судить о степенях достижения целей в аспекте развития ИК.

1. Использование электронного обучения (E_1).
2. Эффективность использования дистанционных образовательных технологий (E_2).
3. Стажировочная активность (E_3).
4. Эффективность стажировочной активности (E_4).
5. Степень соответствия квалификации персонала решаемым задачам (E_5).
6. Закрепление кадров (E_6).
7. Инфраструктурная обеспеченность (E_7).
8. Социально-психологическая удовлетворенность (E_8).
9. Уровень научной и научно-производственной кооперации с партнерами (E_9).
10. Степень соответствия системы мотивирования персонала решаемым задачам (E_{10}).
11. Цифровизация образовательного процесса (E_{11}).
12. Эффективность использования инфраструктуры (E_{12}).
13. Степень индивидуализации образовательных траекторий (E_{13}).
14. Эффективность сетевого взаимодействия с партнерами (E_{14}).

Таблица 2.

Стратегические цели университета в сфере развития ИК (фрагмент)

Группа стейкхолдеров	Перспектива ССП	Цель	Показатель	Когнитивная активность	Структурный компонент ИК
Сотрудники	Ресурсная	Внедрение процедур и критериев оценки качества и эффективности используемых электронных учебных курсов	Использование электронного обучения (E_1)	Обучение	Человеческий капитал
Сотрудники	Ресурсная	Внедрение процедур и критериев оценки качества и эффективности используемых электронных учебных курсов	Эффективность использования дистанционных образовательных технологий (E_2)	Обучение	Человеческий капитал
Сотрудники	Ресурсная	Создание центра взаимодействия университета с бизнесом	Стажировочная активность (E_3)	Обучение	Человеческий капитал
...					
Сотрудники	Ресурсная	Модернизация материально-технической базы университета	Инфраструктурная обеспеченность (E_7)	Вовлечение	Организационный капитал
Сотрудники	Стейк-холдерская	Формирование уникальной корпоративной среды, способствующей развитию и поддержанию корпоративной культуры	Социально-психологическая удовлетворенность (E_8)	Вовлечение	Организационный капитал
Бизнес-сообщество	Процессная	Формирование востребованного бизнесом «портфеля» проектов и научно-исследовательских тематик	Уровень научной и научно-производственной кооперации с партнерами (E_9)	Производственная рационализация	Организационный капитал
...					
Клиенты	Процессная	Создание системы оценки эффективности применения электронных учебных курсов в учебном процессе	Цифровизация образовательного процесса (E_{11})	Производственная рационализация	Организационный капитал
...					
Клиенты	Процессная	Включение российских и зарубежных стажировок в образовательные программы ВО и СПО	Эффективность сетевого взаимодействия с партнерами (E_{14})	Производственная рационализация	Организационный капитал
...					
Государство; Общество	Стейк-холдерская	Формирование научных школ	Публикационная активность (E_{21})	Самосовершенствование	Человеческий капитал
...					
Сотрудники	Стейк-холдерская	Создание системы мотивации персонала к достижению высоких показателей и карьерному росту	Личностный рост ППС (E_{27})	Самосовершенствование	Человеческий капитал
...					
Клиенты	Стейк-холдерская	Развитие бренда университета	Удовлетворенность студентов качеством обучения (E_{30})	Клиенто-ориентированная рационализация	Отношенческий капитал
Клиенты; Бизнес-сообщество; Общество	Стейк-холдерская	Развитие бренда университета	Эффективность управления брендом (E_{31})	Клиенто-ориентированная рационализация	Отношенческий капитал
Бизнес-сообщество; Общество; Государство	Стейк-холдерская	Создание комфортной среды и современной развитой инфраструктуры, необходимой для проведения крупных значимых мероприятий	Эффективность общественных и предпринимательских инициатив (E_{32})	Клиентоориентированная рационализация	Отношенческий капитал
...					
Сотрудники; Клиенты; Бизнес-сообщество; Государство	Стейк-холдерская	Развитие междисциплинарных научных исследований	Междисциплинарные научные проекты (E_{40})	Инновационная деятельность	Отношенческий капитал
Бизнес-сообщество; Государство	Стейк-холдерская	Создание системы R&D, потенциально востребованной реальным сектором экономики	Объем доходов от R&D (E_{41})	Инновационная деятельность	Отношенческий капитал
...					
Бизнес-сообщество; Государство; Общество	Стейк-холдерская	Способность выполнять научные проекты и, в частности, руководить работой студенческих команд по выполнению научных проектов, формированию компетенций STEMskills	Патентная активность (E_{44})	Инновационная деятельность	Отношенческий капитал

15. Уровень адаптации передовых технологий (E_{15}).
16. Уровень автоматизации управленческих процессов (E_{16}).
17. Уровень доступности цифровых образовательных ресурсов (E_{17}).
18. Уровень использования открытых образовательных платформ (E_{18}).
19. Уровень экспертной обеспеченности по стандартам WorldSkills (E_{19}).
20. Уровень освоения стандартов WorldSkills (E_{20}).
21. Публикационная активность (E_{21}).
22. Грантовая активность (E_{22}).
23. Защиты диссертаций (E_{23}).
24. Сформированность организационной культуры (E_{24}).
25. Международная научная острепененность (E_{25}).
26. Международная академическая мобильность (E_{26}).
27. Личностный рост ППС (E_{27}).
28. Инновационно-предпринимательская активность ППС (E_{28}).
29. Трудоустройство студентов (E_{29}).
30. Удовлетворенность студентов качеством обучения (E_{30}).
31. Эффективность управления брендом (E_{31}).
32. Эффективность общественных и предпринимательских инициатив (E_{32}).
33. Уровень поддержки предпринимательской активности студентов (E_{33}).
34. Уникальность портфеля образовательных программ университета (E_{34}).
35. Уровень использования цифрового маркетинга в процессе взаимодействия с абитуриентами (E_{35}).
36. Внутренняя востребованность ДОП университета (E_{36}).
37. Внешняя востребованность ДОП университета (E_{37}).
38. Международная образовательная активность (E_{38}).
39. Реализованные научные проекты (E_{39}).
40. Междисциплинарные научные проекты (E_{40}).
41. Объем доходов от R&D (E_{41}).
42. Квалификация персонала в сфере научных исследований и разработок (E_{42}).
43. Эффективность деятельности инновационного бизнес-инкубатора (E_{43}).
44. Патентная активность (E_{44}).

На следующем этапе был произведен опрос экспертов, в число которых вошли представители ППС и АУП университета, а также специально приглашенные внешние эксперты. Эксперты в рамках заданной лингвистической шкалы оценили степень влияния выбранных показателей на ключевые показатели развития ИК, соответствующие основным структурным компонентам ИК. Ответы экспертов были проверены на согласованность и усреднены с учетом экзогенно заданных уровней компетентности экспертов. Заметим, что каждый отдельный эксперт оценивал влияние на ключевые показатели ИК не всех 44 показателей, а лишь тех, в отношении которых обладал соответствующими экспертными знаниями (компетенциями). Результатами данного этапа экспертного опроса являются средневзвешенные экспертные оценки в виде нечетких чисел Гауссова типа. В *таблице 3* приведены параметры соответствующих аппроксимирующих гауссианов.

Поскольку «границы отсечения» эксплицитных и имплицитных показателей заранее не были известны, экспертам требовалось также оценить взаимное влияние всех 44 показателей друг на друга. В этом случае каждый эксперт также отвечал лишь на вопросы, относящиеся к сфере его экспертных знаний. Таким образом, каждому эксперту было необходимо ответить на разумное количество вопросов за приемлемое время. Такой подход оставляет лицу, принимающему решения, широкие возможности по варьированию «границ отсечения», не вызывая необходимости дополнительных экспертных вопросов. Результаты второго этапа экспертного опроса (в виде параметров соответствующих средневзвешенным нечетким экспертным оценкам аппроксимирующих гауссианов) частично приведены в *таблице 4*.

Для проведения экспертного опроса, обработки экспертных ответов и проведения необходимых расчетов на основе описанной выше нечеткой модели был разработан программный комплекс. Он позволяет, в том числе, формировать множества эксплицитных и имплицитных факторов ИК при заданных «границах отсечения» и выбранных методах дефаззификации (если «границы отсечения» заданы четко).

В *таблице 5* приведены множества эксплицитных и имплицитных факторов ИК университета при различных «границах отсечения», полученные с использованием трех методов дефаззификации (Center Of Gravity / Maximum Of Maximums / Median).

Таблица 3.

**Нечеткие оценки влияния показателей множества E
на ключевые показатели развития ИК (фрагмент)**

Показатель	Человеческий капитал (C_1)		Организационный капитал (C_2)		Отношенческий капитал (C_3)		Интеллектуальный капитал	
	μ	σ	μ	σ	μ	σ	μ	σ
E_1	2,4472	0,2675	1,2536	0,1401	3,7685	0,2108	2,4999	0,3247
E_2	3,8603	0,3763	1,2991	0,2477	0,1345	0,2980	1,8461	0,2185
E_3	2,3443	0,3999	2,4939	0,3249	3,6572	0,1670	2,8657	0,2251
...								
E_{42}	2,4143	0,3371	2,3756	0,4054	4,8342	0,2865	3,1524	0,1353
E_{43}	3,9457	0,1746	3,8787	0,3795	4,7375	0,1445	3,8849	0,1578
E_{44}	3,5705	0,1556	3,7692	0,1379	3,6417	0,2386	3,5971	0,3190

Таблица 4.

Нечеткие оценки взаимного влияния показателей множества E (фрагмент)

Показатель	E_1		E_2		E_3		...	E_{42}		E_{43}		E_{44}	
	μ	σ	μ	σ	μ	σ		μ	σ	μ	σ	μ	σ
E_1	*		0,26	0,14	4,72	0,14		4,93	0,14	4,85	0,28	3,55	0,34
E_2	1,26	0,27	*		1,07	0,35		2,55	0,29	1,23	0,26	2,31	0,33
E_3	3,83	0,33	4,57	0,20	*			4,64	0,20	3,74	0,35	3,83	0,14
...													
E_{42}	3,61	0,17	3,56	0,19	4,81	0,13		*		0,42	0,26	4,73	0,21
E_{43}	4,66	0,39	4,66	0,18	2,46	0,24		2,62	0,25	*		0,35	0,23
E_{44}	4,89	0,18	4,73	0,39	3,63	0,28		1,28	0,38	0,19	0,33	*	

Таблица 5.

Наборы эксплицитных и имплицитных факторов ИК

Первая граница отсечения	Номера показателей ИК, принятых за эксплицитные	Вторая граница отсечения	Номера показателей ИК, принятых за имплицитные
2	1, 2, 4, 5, 6, 9, 11, 14, 21, 22, 23, 24, 31, 39, 40, 41, 42 / 1, 2, 4, 5, 6, 9, 11, 14, 21, 22, 23, 24, 31, 39, 40, 41, 42 / 1, 2, 4, 5, 6, 9, 11, 14, 21, 22, 23, 24, 31, 39, 40, 41, 42	1,5	3, 7, 8, 10, 12, 15, 26, 27, 28, 32, 33 / 3, 7, 8, 10, 12, 15, 26, 27, 28, 32, 33, 43, 44 / 3, 7, 8, 10, 12, 15, 26, 27, 28, 32, 33, 43, 44
		1,75	3, 7, 8, 12, 27, 33 / 3, 7, 8, 12, 27, 33, 43 / 3, 7, 8, 12, 27, 33, 43
		2	7, 8, 12 / 7, 8, 12 / 7, 8, 12
		2,25	Отсутствуют / Отсутствуют / Отсутствуют

Первая граница отсечения	Номера показателей ИК, принятых за эксплицитные	Вторая граница отсечения	Номера показателей ИК, принятых за имплицитные
2,5	1, 4, 6, 9, 11, 14, 21, 22, 23, 24, 31, 39, 40, 41, 42 / 1, 4, 6, 9, 11, 14, 21, 22, 23, 24, 31, 39, 40, 41, 42 / 1, 4, 6, 9, 11, 14, 21, 22, 23, 24, 31, 39, 40, 41, 42	1,5	2, 3, 5, 7, 8, 10, 12, 15, 26, 27, 28, 32, 33, 43, 44 / 2, 3, 5, 7, 8, 10, 12, 15, 26, 27, 28, 32, 33, 43, 44 / 2, 3, 5, 7, 8, 10, 12, 15, 26, 27, 28, 32, 33, 43, 44
		1,75	2, 3, 7, 8, 12, 27, 28, 33, 43 / 2, 3, 7, 8, 12, 27, 28, 33, 43 / 2, 3, 7, 8, 12, 27, 28, 33, 43
		2	2, 7, 8, 12 / 2, 7, 8, 12 / 2, 7, 8, 12
		2,25	2 / 2 / 2
3	1, 6, 9, 11, 14, 21, 22, 23, 39, 40, 41, 42 / 1, 6, 9, 11, 14, 21, 22, 23, 39, 40, 41, 42 / 1, 6, 9, 11, 14, 21, 22, 23, 39, 40, 41, 42	1,5	2, 3, 4, 5, 7, 8, 10, 12, 15, 24, 26, 27, 28, 32, 33, 43, 44 / 2, 3, 4, 5, 7, 8, 10, 12, 15, 24, 26, 27, 28, 32, 33, 43, 44 / 2, 3, 4, 5, 7, 8, 10, 12, 15, 24, 26, 27, 28, 32, 33, 43, 44
		1,75	2, 3, 4, 7, 8, 12, 24, 27, 28, 32, 33, 44 / 2, 3, 4, 7, 8, 12, 24, 27, 28, 32, 33, 44 / 2, 3, 4, 7, 8, 12, 24, 27, 28, 32, 33, 44
		2	2, 3, 4, 7, 8, 12, 24, 28 / 2, 3, 4, 7, 8, 12, 24 / 2, 3, 4, 7, 8, 12, 24, 28
		2,25	2, 4, 7, 12, 24 / 2, 4, 7, 12, 24 / 2, 4, 7, 12, 24
3,5	1, 6, 9, 11, 14, 22, 41, 42 / 1, 6, 9, 11, 14, 22, 41, 42 / 1, 6, 9, 11, 14, 22, 41, 42	1,5	2, 3, 4, 5, 7, 8, 10, 12, 13, 15, 17, 21, 23, 24, 28, 31, 32, 33, 39, 40, 43, 44 / 2, 3, 4, 5, 7, 8, 10, 12, 13, 15, 17, 21, 23, 24, 28, 31, 32, 33, 39, 40, 43, 44 / 2, 3, 4, 5, 7, 8, 10, 12, 13, 15, 17, 21, 23, 24, 28, 31, 32, 33, 39, 40, 43, 44
		1,75	2, 3, 4, 5, 7, 8, 12, 21, 23, 24, 28, 32, 33, 39, 40, 43, 44 / 2, 3, 4, 5, 7, 8, 12, 21, 23, 24, 28, 32, 33, 39, 40, 43, 44 / 2, 3, 4, 5, 7, 8, 12, 21, 23, 24, 28, 32, 33, 39, 40, 43, 44
		2	2, 3, 4, 7, 8, 12, 24, 32, 33, 39, 40 / 2, 3, 4, 7, 8, 12, 24, 32, 33, 39, 40 / 2, 3, 4, 7, 8, 12, 24, 32, 33, 39, 40
		2,25	2, 4, 7, 12, 24, 39, 40 / 2, 4, 7, 12, 24, 39, 40 / 2, 4, 7, 12, 24, 39, 40

Лицо, принимающее решение, имеет возможность задать первую («эксплицитную») «границу отсечения», исходя из требований к силе прямого влияния отбираемых факторов на результирующие показатели ИК. В результате будет сформирован набор эксплицитных факторов ИК. Затем, исходя из требований к силе опосредованного влияния отбираемых факторов на результирующие показатели, выбирается вторая («имплицитная») «граница отсечения». Тем самым, формируется набор имплицитных факторов ИК.

4. Обсуждение

Анализ полученных результатов показывает следующее.

1. Изменение первой («эксплицитной») «границы отсечения» приводит к изменению наборов фак-

торов ИК, принимаемых за эксплицитные. При этом, чем больше «граница отсечения» (что означает более строгие требования к силе прямого влияния отбираемых факторов на результирующие показатели), тем меньше количество эксплицитных факторов, и наоборот. Интересно, что при различных методах дефаззификации наборы эксплицитных факторов не меняются для фиксированной «границы отсечения». Это связано с тем, что полученные с помощью разных методов дефаззификации четкие оценки силы прямого влияния факторов отличаются недостаточно сильно для того, чтобы изменился состав эксплицитных факторов. Это, в свою очередь, по всей видимости связано с трапециевидным типом выбранных функций принадлежности.

2. Изменение второй («имплицитной») «границы отсечения» также приводит к изменению наборов факторов ИК, принимаемых за имплицитные. При

этом, также, чем больше «граница отсечения», тем меньше количество имплицитных факторов, и наоборот. При выборе имплицитных факторов начинает играть роль выбор метода дефаззификации, но только при малых значениях обеих «границ отсечения».

3. Изменения наборов имплицитных факторов при изменении второй «границы отсечения» в значительной степени зависят от выбранной первой «границы отсечения» вне зависимости от метода дефаззификации.

4. Некоторые факторы могут быть определены как эксплицитные (для одних границ отсечения) и как имплицитные (для других границ отсечения). Это связано, во-первых, с требованиями к силе прямого или опосредованного влияния фактора на результирующие показатели для отнесения к той или иной группе, во-вторых, с лингвистической неопределенностью при формулировании таких требований и экспертной оценке силы влияния. Именно поэтому и возникла необходимость разработки нечеткой модели.

5. Предложенный метод формирования каузального поля показателей ИК является универсальным в том смысле, что он применим к различным типам организаций разной отраслевой принадлежности. Стандартными будут ключевые показатели ИК, соответствующие основным структурным компонентам ИК (человеческий капитал, организационный капитал, отношенческий капитал), типы когнитивной активности (обучение, вовлечение, производственная рационализация, самосовершенствование, клиентоориентированная рационализация, инновационная деятельность) и соответствие между типами когнитивной активности и структурными компонентами ИК. Универсальными будут все этапы базовой схемы метода.

6. Вместе с тем наборы эксплицитных и имплицитных факторов ИК для различных организаций могут отличаться радикально по следующим причинам. Прежде всего, могут значительно различаться набор и состав групп заинтересованных сторон (стейкхолдеров) организаций и их запросов к организациям. Соответственно, будут существенно отличаться стратегические карты целей организаций, а, значит, и цели, относящиеся к развитию ИК, и их результирующие показатели (т.е. исходное множество показателей развития ИК, из которых выбираются эксплицитные и имплицитные факторы). Даже если исходные множества показателей оказались достаточно близки по составу, то могут существенно

отличаться степени влияния показателей на ключевые показатели ИК и друг на друга. Наконец, лица, принимающие решения, могут выбирать различные «границы отсечения» и методы дефаззификации.

Заключение

Предложена концептуальная схема формирования каузального поля показателей развития ИК в увязке со стратегией организации и типами когнитивной активности. Операционализация данной схемы осуществлена путем разработки нечеткой экономико-математической модели, позволяющей выявлять эксплицитные и имплицитные факторы ИК. Предложенная схема и модель имеют следующие отличительные особенности. Совокупность показателей ИК формируется на основе результирующих показателей стратегических целей, отобранных из карты целей модифицированной ССП с группировкой по шести типам когнитивной активности. Ключевыми показателями ИК являются интегральные показатели, соответствующие основным структурным компонентам ИК (человеческий капитал, организационный капитал, отношенческий капитал). Эксплицитные факторы ИК отбираются по результатам оценки прямого влияния на ключевые показатели ИК с помощью задания «границы отсечения». Имплицитные факторы ИК отбираются по результатам оценки опосредованного влияния на ключевые показатели ИК через эксплицитные факторы с помощью задания еще одной «границы отсечения». Оценки прямых влияний осуществляются экспертно в заданной лингвистической шкале с соответствующими функциями принадлежности нечетких множеств. Оценки опосредованных влияний вычисляются на основе операций с матрицами, элементами которых являются нечеткие числа. Представлены результаты апробации модели на примере крупного регионального университета. Показано, что наборы эксплицитных и имплицитных факторов ИК университета меняются в зависимости от заданных «границ отсечения» и выбранного метода дефаззификации. ■

Благодарности

Исследование выполнено за счет гранта Российского научного фонда № 23-28-01091 (<https://rscf.ru/project/23-28-01091/>) в ФГБОУ ВО «Владивостокский государственный университет», Приморский край.

Литература

1. Kucharska W. Tacit knowledge impact on intellectual capital and innovativeness in the healthcare sector: A cross-country study of Poland and the US // *Journal of Business Research*. 2022. Vol. 149. P. 869–883. <https://doi.org/10.1016/j.jbusres.2022.05.059>
2. Назаров Д.М. Методология нечетко-множественной оценки имплицитных факторов в деятельности организации. Екатеринбург: Изд-во Урал. гос. экон. ун-та, 2016.
3. Ding Y., Li G. Study on the management of intellectual capital // *International Journal of Business and Management*. 2010. Vol. 5. No. 2. P. 213–216. <https://doi.org/10.5539/ijbm.v5n2p213>
4. Назаров Д.М. Модель оценки имплицитных факторов на основе нечетко-множественных описаний // *Известия ДВФУ. Экономика и управление*. 2016. № 4(80). С. 3–17. <https://doi.org/10.5281/zenodo.220793>
5. Лурия А.Р. Язык и сознание. М.: Издательство Московского университета, 1979.
6. Приходько А.И. Имплицитность как способ скрытого оценивания // *Актуальные проблемы филологии и педагогической лингвистики*. 2012. № 14. С. 48–54.
7. Милосердова Е.В. Имплицитность как лингвопрагматическая категория и ее роль в разных типах дискурса // *Вестник Тамбовского университета. Серия: Гуманитарные науки*. 2011. Т. 104. № 12–2. С. 727–732.
8. del-Río-Ortega A., Resinas Arias de Reyna M., Durán Toro A., Ruiz-Cortés A. Defining process performance indicators by using templates and patterns // *Business Process Management*. 2012. Vol. 7481. P. 223–228. https://doi.org/10.1007/978-3-642-32885-5_18
9. Цюцюра С.В., Криворучко О.В., Цюцюра М.І. Теоретичні основи та сутність управлінських рішень. Моделі прийняття управлінських рішень // *Управління Розвитком Складних Систем*. 2012. No. 9. P. 50–58.
10. Назаров Д.М. Фундаментальные основы имплицитности в системе экономического развития организации // *Известия СПбГЭУ*. 2015. № 3(93). С. 7–14.
11. Markić M., Požega Ž., Crnković B. The impact of knowledge management on the economic indicators of the companies // *South East European Journal of Economics and Business*. 2022. Vol. 17. No. 2. P. 34–48. <https://doi.org/10.2478/jeb-2022-0013>
12. Van Looy A., Shafagatova A. Business process performance measurement: A structured literature review of indicators, measures and metrics // *SpringerPlus*. 2016. Vol. 5. P. 1–24. <https://doi.org/10.1186/s40064-016-3498-1>
13. Назаров Д.М. Модель рефлексивного отбора имплицитных показателей управленческой деятельности организации // *Вестник СПбГУ. Экономика*. 2017. Т. 33. № 3. С. 498–518. <https://doi.org/10.21638/11701/spbu05.2017.308>
14. Lawrie G.J.G., Cobbold I. Third-generation balanced scorecard: Evolution of an effective strategic control tool // *International Journal of Productivity and Performance Management*. 2004. Vol. 53. No. 7. P. 611–623. <https://doi.org/10.1108/17410400410561231>
15. Olve N.-G., Roy J., Wetter M. *Performance drivers: A practical guide to using the balanced scorecard*. Chichester: Wiley, 1999.
16. Kaplan R.S., Norton D.P. The balanced scorecard – measures that drive performance // *Harvard Business Review*. 1992. Vol. 70. No. 1. P. 71–79.
17. Tawse A., Tabesh P. Thirty years with the balanced scorecard: What we have learned // *Business Horizons*, 2022. Vol. 66. No. 1. P. 123–132. <https://doi.org/10.1016/j.bushor.2022.03.005>
18. Pierce E. A balanced scorecard for maximizing data performance // *Frontiers in Big Data*. 2022. Vol. 5. P. 821103. <https://doi.org/10.3389/fdata.2022.821103>
19. Kopecka N. The balanced scorecard implementation, integrated approach and the quality of its measurement // *Procedia Economics and Finance*. 2015. Vol. 25. P. 59–69. [https://doi.org/10.1016/S2212-5671\(15\)00713-3](https://doi.org/10.1016/S2212-5671(15)00713-3)
20. Mouritsen J., Larsen H., Bukh P. Dealing with the knowledge economy: intellectual capital versus balanced scorecard // *Journal of Intellectual Capital*. 2005. Vol. 6. No. 1. P. 8–27. <https://doi.org/10.1108/14691930510574636>
21. Johanson U., Martensson M., Skoog M. Mobilising change through the management control of intangibles // *Accounting, Organizations and Society*. 2001. Vol. 26. No. 7/8. P. 715–733. [https://doi.org/10.1016/S0361-3682\(01\)00024-1](https://doi.org/10.1016/S0361-3682(01)00024-1)
22. Petty R., Guthrie J. Intellectual capital literature review: measurement, reporting and management // *Journal of Intellectual Capital*. 2000. Vol. 1. No. 2/3. P. 155–176. <https://doi.org/10.1108/14691930010348731>
23. Bontis N., Dragonetti N.C., Jacobsen K., Roos G. The knowledge toolbox: A review of the tools available to measure and manage intangible resources // *European Management Journal*. 1999. Vol. 17. No. 4. P. 391–402.
24. Недосекин А.О., Шкатов М.Ю., Абдулаева З.И. Разработка системы сбалансированных показателей для морской нефтегазовой смешанной компании с использованием нечетко-множественных описаний // *Аудит и финансовый анализ*. 2013. № 4. С. 126–134.
25. Horváth & Partners (eds.). *Balanced Scorecard umsetzen (Balanced Scorecard implementation)*. (4th ed.). Stuttgart, Schäffer-Poeschel, 2007.

26. Ковалев Д.В. Система сбалансированных показателей (BSC) – большая правильная ошибка? // *Контроллинг*. 2006. № 2(18). С. 26–29.
27. Nørreklit H. The balanced scorecard: What is the score? A rhetorical analysis of the balanced scorecard // *Accounting, Organizations and Society*. 2003. Vol. 28. No. 6. P. 591–619. [https://doi.org/10.1016/S0361-3682\(02\)00097-1](https://doi.org/10.1016/S0361-3682(02)00097-1)
28. Sharabati A.A.A., Naji Jawad S., Bontis N. Intellectual capital and business performance in the pharmaceutical sector of Jordan // *Management Decision*. 2010. Vol. 48. No. 1. P. 105–131. <https://doi.org/10.1108/00251741011014481>
29. Lervik E. Relational capital: A study on its importance, quantification and its impact on business sectors and markets. The European Inter-University Association on Society, Science and Technology (ESST), 2006.
30. Montequin V.R., Fernandez F.O., Cabal V.A., Gutierrez N.R. An integrated framework for intellectual capital measurement and knowledge management implementation in small and medium-sized enterprises // *Journal of Information Science*. 2006. Vol. 32. P. 525–538. <https://doi.org/10.1177/0165551506067127>
31. Roos G., Pike S., Fernstrom L. Intellectual capital: Practice of management. Abingdon, Routledge, 2006.
32. Bontis N., Keow W.C.C., Richardson S. Intellectual capital and business performance in Malaysian industry // *Journal of Intellectual Capital*. 2000. Vol. 1. No. 1. P. 85–100.
33. Солодухин К.С., Дзина Г.А. Применение системы сбалансированных показателей в университете на основе теории заинтересованных сторон. *Контроллинг*. 2009. № 1(29). С. 12–23.
34. Солодухин К.С. Постановка системы сбалансированных показателей в стейкхолдер-компаниях. *Контроллинг*. 2009. № 2(30). С. 64–69.
35. Korobova E., Kardovich I., Konyshva M., Mironova D. Cognitive activity: Philosophical analysis, psychological and pedagogical characteristics. SHS Web of Conferences. 2018. Vol. 50. 01083. <https://doi.org/10.1051/shsconf/20185001083>
36. Коротаева Е.В., Нефедова А.Н. Развитие когнитивной активности у старших дошкольников // *Педагогическое образование в России*. 2012. № 3. С. 176–180.
37. Недолужко О.В. Новый подход к формированию понятийного пространства феномена интеллектуального капитала организации // *Вестник Удмуртского университета. Серия «Экономика и право»*. 2016. № 2(26). С. 42–49.
38. Bazil G.D., Adilova S.K., Abzhanova L.K., Sugurova L.A., Yerzhanova M.E. Fuzzy simulation of organizational adjustment processes management based on heat supply balanced scorecard // *Innovative Infrastructure Solutions*. 2021. Vol. 6. P. 77. <https://doi.org/10.1007/s41062-020-00435-3>
39. Биктимеров В.А. Методика анализа цифрового содержимого для выявления закономерностей. VI-технологии и корпоративные информационные системы в оптимизации бизнес-процессов цифровой экономики: материалы VI Международной научно-практической очно-заочной конференции. Екатеринбург: Издательство Уральского государственного экономического университета, 2019. С. 30–31.
40. Hushko S., Temchenko O., Kryshchyna I., Temchenko H., Maksymova I., Huk O. Modelling of management activity of the organization considering the impact of implicit factors in business processes // *Eastern-European Journal of Enterprise Technologies*. 2018. Vol. 3. No. 91. P. 13–21. <https://doi.org/10.15587/1729-4061.2018.121647>
41. Pakseresht M., Seyyedi M.A., Zade M.M., Gardesh H. Business process measurement model based on the fuzzy multi-agent systems // *AIKED Proceedings of WSEAS*. 2009. P. 501–506.
42. Chen S., Hwang C. Fuzzy multiple attribute decision making: Methods and applications. New York: Springer-Verlag, 1992.
43. Аньшин В.М., Демкин И.В., Царьков И.Н., Никонов И.М. Применение теории нечетких множеств к задаче формирования портфеля проектов // *Проблемы анализа рисков*. 2008. Т. 5. № 3. С. 8–21.
44. Vahidi J., Rezvani S. Arithmetic operations on trapezoidal fuzzy numbers // *Journal of Nonlinear Analysis and Application*. 2013. Vol. 2013. P. 1–8. <https://doi.org/10.5899/2013/jnaa-00111>
45. Raskin L., Sira O. Performing arithmetic operations over the (L-R)-type fuzzy numbers // *Eastern-European Journal of Enterprise Technologies*. 2020. Vol. 3. No. 4(105). P. 6–11. <https://doi.org/10.15587/1729-4061.2020.203590>

Об авторах

Завалин Георгий Сергеевич

начальник отдела интеллектуального анализа данных, стажер-исследователь, лаборатория стратегического планирования, Владивостокский государственный университет, 690014, г. Владивосток, ул. Гоголя, д. 41;

E-mail: georgiy.zavalin08@vvsu.ru

ORCID: 0000-0003-4519-0242

Недолужко Ольга Вячеславовна

кандидат экономических наук, доцент;

доцент, кафедра экономики и управления, доцент, кафедра математики и моделирования, научный сотрудник лаборатории стратегического планирования, Владивостокский государственный университет, 690014, г. Владивосток, ул. Гоголя, д. 41;

E-mail: olga.nedoluzhko25@gmail.com

ORCID: 0000-0002-3622-7501

Солодухин Константин Сергеевич

доктор экономических наук, профессор;

профессор, кафедра математики и моделирования, заведующий лабораторией стратегического планирования, Владивостокский государственный университет, 690014, г. Владивосток, ул. Гоголя, д. 41;

E-mail: k.solodukhin@mail.ru

ORCID: 0000-0003-3619-1219

Formation of the causal field of indicators for an organization's intellectual capital development: A concept and a fuzzy economic and mathematical model

Georgij S. Zavalin

E-mail: georgiy.zavalin08@vvsu.ru

Olga V. Nedoluzhko

E-mail: olga.nedoluzhko25@gmail.com

Konstantin S. Solodukhin

E-mail: k.solodukhin@mail.ru

Vladivostok State University

Address: 41, Gogolya Str., Vladivostok 690014, Russia

Abstract

The development of intellectual capital theory through the introduction of the concept of implicitness involves considering intellectual capital as an implicit factor, so that the process of its formation is largely determined by the impact of specific hidden factors whose impact is expressed implicitly and is difficult to formalize. Currently, the process of selecting explicit and implicit factors affecting intellectual capital is not formalized in domestic and foreign studies, and therein is the relevance of this work. The purpose of this study was to develop a scheme for selecting explicit and implicit factors in the development of the organization's intellectual capital in conjunction with its strategy based on a modified Balanced Scorecard,

taking into account the distribution of indicators by types of cognitive activity. The implementation of this scheme was carried out by developing a fuzzy economic and mathematical model suitable for practical use. The main feature of the model is the possibility of fuzzy setting of “cut-off boundaries” for explicit and implicit factors. We present the results of testing the model on the example of a large regional university. Sets of explicit and implicit factors of the university’s intellectual capital are given for various “cut-off boundaries” using various defuzzification methods.

Keywords: cognitive activity, fuzzy model, implicit factor, intellectual capital, strategic management

Citation: Zavalin G.S., Nedoluzhko O.V., Solodukhin K.S. (2023) Formation of the causal field of indicators for an organization’s intellectual capital development: A concept and a fuzzy economic and mathematical model. *Business Informatics*, vol. 17, no. 3, pp. 53–69. DOI: 10.17323/2587-814X.2023.3.53.69

References

1. Kucharska W. (2022) Tacit knowledge impact on intellectual capital and innovativeness in the healthcare sector: A cross-country study of Poland and the US. *Journal of Business Research*, vol. 149, pp. 869–883. <https://doi.org/10.1016/j.jbusres.2022.05.059>
2. Nazarov D.M. (2016) *Methodology of fuzzy set evaluation of implicit factors in organizational activities*. Ekaterinburg: Ural State Economic University Press (in Russian).
3. Ding Y., Li G. (2010) Study on the management of intellectual capital. *International Journal of Business and Management*, vol. 5, no. 2, pp. 213–216. <https://doi.org/10.5539/ijbm.v5n2p213>
4. Nazarov D.M. (2016) The evaluation model of implicit factors on the basis of fuzzy-set descriptions. *Far Eastern Federal University News*, no. 4 (80), pp. 3–17 (in Russian). <https://doi.org/10.5281/zenodo.220793>
5. Luria A.R. (1979) *Language and consciousness*. Moscow: Moscow University Press (in Russian).
6. Prikhodko A.I. (2012) Implicitness as a means of latent evaluation. *Current Issues in Philology and Pedagogical Linguistics*, no. 14, pp. 48–54 (in Russian).
7. Miloserdova E.V. (2011) Implicitness as a linguo-pragmatic category and its role in different types of discourse. *Bulletin of Tomsk State University*, vol. 104, no. 12–2, pp. 727–732 (in Russian).
8. del-Río-Ortega A., Resinas Arias de Reyna M., Durán Toro A., Ruiz-Cortés A. (2012) Defining process performance indicators by using templates and patterns. *Business Process Management*, vol. 7481, pp. 223–228. https://doi.org/10.1007/978-3-642-32885-5_18
9. Tsiutsiura S.V., Kryvoruchko O.V., Tsiutsiura M.I. (2012) Theoretical basis and essence of managerial decisions. Models for accepting managerial decisions. *Development Management of Folding Systems*, no. 9, pp. 50–58 (in Ukrainian).
10. Nazarov D.M. (2015) The fundamental basis of implied in the system of economic development organizations. *Bulletin of Saint Petersburg State University of Economics*, no. 3(93), pp. 7–14 (in Russian).
11. Markić M., Požega Ž., Crnković B. (2022) The impact of knowledge management on the economic indicators of the companies. *South East European Journal of Economics and Business*, vol. 17, no. 2, pp. 34–48. <https://doi.org/10.2478/jeb-2022-0013>
12. Van Looy A., Shafagatova A. (2016) Business process performance measurement: A structured literature review of indicators, measures and metrics. *SpringerPlus*, vol. 5, pp. 1–24. <https://doi.org/10.1186/s40064-016-3498-1>
13. Nazarov D.M. (2017) Model of reflexive selection of implicit indicators of management activities of an organization. *Bulletin of Saint-Petersburg State University*, vol. 33, no. 3, pp. 498–518 (in Russian). <https://doi.org/10.21638/11701/spbu05.2017.308>
14. Lawrie G.J.G., Cobbold I. (2004) Third-generation balanced scorecard: Evolution of an effective strategic control tool. *International Journal of Productivity and Performance Management*, vol. 53, no. 7, pp. 611–623. <https://doi.org/10.1108/17410400410561231>
15. Olive N.-G., Roy J., Wetter M. (1999) *Performance drivers: A practical guide to using the balanced scorecard*. Chichester: Wiley.
16. Kaplan R.S., Norton D.P. (1992) The balanced scorecard – measures that drive performance. *Harvard Business Review*, vol. 70, no. 1, pp. 71–79.
17. Tawse A., Tabesh P. (2022) Thirty years with the balanced scorecard: What we have learned. *Business Horizons*, vol. 66, no. 1, pp. 123–132. <https://doi.org/10.1016/j.bushor.2022.03.005>
18. Pierce E. (2022) A balanced scorecard for maximizing data performance. *Frontiers in Big Data*, vol. 5, 821103. <https://doi.org/10.3389/fdata.2022.821103>
19. Kopecka N. (2015) The balanced scorecard implementation, integrated approach and the quality of its measurement. *Procedia Economics and Finance*, vol. 25, pp. 59–69. [https://doi.org/10.1016/S2212-5671\(15\)00713-3](https://doi.org/10.1016/S2212-5671(15)00713-3)

20. Mouritsen J., Larsen H., Bukh P. (2005) Dealing with the knowledge economy: intellectual capital versus balanced scorecard. *Journal of Intellectual Capital*, vol. 6, no. 1, pp. 8–27. <https://doi.org/10.1108/14691930510574636>
21. Johanson U., Martensson M., Skoog M. (2001) Mobilising change through the management control of intangibles. *Accounting, Organizations and Society*, vol. 26, no. 7/8, pp. 715–733. [https://doi.org/10.1016/S0361-3682\(01\)00024-1](https://doi.org/10.1016/S0361-3682(01)00024-1)
22. Petty R., Guthrie J. (2000) Intellectual capital literature review: measurement, reporting and management. *Journal of Intellectual Capital*, vol. 1, no. 2/3, pp. 155–176. <https://doi.org/10.1108/14691930010348731>
23. Bontis N., Dragonetti N.C., Jacobsen K., Roos G. (1999) The knowledge toolbox: A review of the tools available to measure and manage intangible resources. *European Management Journal*, vol. 17, no. 4, pp. 391–402.
24. Nedosekin A.O., Shkatov M.Yu., Abdulayeva Z.I. (2013) Developing a balanced scorecard system for a mixed marine oil and gas company using fuzzy sets. *Audit and Financial Analysis*, no. 4, pp. 126–134 (in Russian).
25. Horváth & Partners (eds.) (2007) *Balanced Scorecard umsetzen (Balanced Scorecard implementation)*, 4th ed. Stuttgart, Schäffer-Poeschel.
26. Kovalev D.V. (2006) The balanced scorecard (BSC) – a big right mistake? *Controlling*, vol. 2, no. 18, pp. 26–29 (in Russian).
27. Nørreklit H. (2003) The balanced scorecard: What is the score? A rhetorical analysis of the balanced scorecard. *Accounting, Organizations and Society*, vol. 28, no. 6, pp. 591–619. [https://doi.org/10.1016/S0361-3682\(02\)00097-1](https://doi.org/10.1016/S0361-3682(02)00097-1)
28. Sharabati A.A.A., Naji Jawad S., Bontis N. (2010) Intellectual capital and business performance in the pharmaceutical sector of Jordan. *Management Decision*, vol. 48, no. 1, pp. 105–131. <https://doi.org/10.1108/00251741011014481>
29. Lervik E. (2006). *Relational capital: A study on its importance, quantification and its impact on business sectors and markets*. The European Inter-University Association on Society, Science and Technology (ESST).
30. Montequin V.R., Fernandez F.O., Cabal V.A., Gutierrez N.R. (2006) An integrated framework for intellectual capital measurement and knowledge management implementation in small and medium-sized enterprises. *Journal of Information Science*, vol. 32, pp. 525–538. <https://doi.org/10.1177/0165551506067127>
31. Roos G., Pike S., Fernstrom L. (2006) *Intellectual capital: Practice of management*. Abingdon, Routledge.
32. Bontis N., Keow W.C.C., Richardson S. (2000) Intellectual capital and business performance in Malaysian industry. *Journal of Intellectual Capital*, vol. 1, no. 1, pp. 85–100.
33. Solodukhin K.S., Dzina G.A. (2009) Applying a balanced scorecard system at a university based on the stakeholder theory. *Controlling*, no. 1(29), pp. 12–23 (in Russian).
34. Solodukhin K.S. (2009) The implementation of a balanced scorecard system in a stakeholder company. *Controlling*, no. 2(30), pp. 64–69 (in Russian).
35. Korobova E., Kardovich I., Konysheva M., Mironova D. (2018) Cognitive activity: Philosophical analysis, psychological and pedagogical characteristics. *SHS Web of Conferences*, vol. 50, 01083. <https://doi.org/10.1051/shsconf/20185001083>
36. Korotaeva E.V., Nefedova A.N. (2012) Development of cognitive activity in senior preschoolers. *Pedagogical Education in Russia*, no. 3, pp. 176–180 (in Russian).
37. Nedoluzhko O.V. (2016) A new approach to forming the conceptual space of the intellectual capital phenomenon of an organization. *Bulletin of Udmurt University*, no. 2(26), pp. 42–49 (in Russian).
38. Bazil G.D., Adilova S.K., Abzhanova L.K., Sugurova L.A., Yerzhanova M.E. (2021) Fuzzy simulation of organizational adjustment processes management based on heat supply balanced scorecard. *Innovative Infrastructure Solutions*, vol. 6, pp. 77. <https://doi.org/10.1007/s41062-020-00435-3>
39. Biktimerov V.A. (2019) Methodology for analyzing digital content to identify patterns. *Proceedings of the BI-technologies and corporate information systems in optimizing business processes of the digital economy: VI International scientific-practical full-time and correspondence conference*. Yekaterinburg, pp. 30–31 (in Russian).
40. Hushko S., Temchenko O., Kryshchtopa I., Temchenko H., Maksymova I., Huk O. (2018) Modelling of management activity of the organization considering the impact of implicit factors in business processes. *Eastern-European Journal of Enterprise Technologies*, vol. 3, no. 91, pp. 13–21. <https://doi.org/10.15587/1729-4061.2018.121647>
41. Pakseresht M., Seyyedi M.A., Zade M.M., Gardesh H. (2009) Business process measurement model based on the fuzzy multi-agent systems. *AIKED Proceedings of WSEAS*, pp. 501–506.
42. Chen S., Hwang C. (1992) *Fuzzy multiple attribute decision making: Methods and applications*. New York: Springer-Verlag.
43. Anshin V.M., Dyomkin I.V., Tsarkov I.V., Nikonov I.M. (2008) On application of fuzzy set theory to the problem of project portfolio selection. *Issues of Risk Analysis*, vol. 5, no. 3, pp. 8–21 (in Russian).
44. Vahidi J., Rezvani S. (2013) Arithmetic operations on trapezoidal fuzzy numbers. *Journal of Nonlinear Analysis and Application*, vol. 2013, pp. 1–8. <https://doi.org/10.5899/2013/jnaa-00111>
45. Raskin L., Sira O. (2020) Performing arithmetic operations over the (L-R)-type fuzzy numbers. *Eastern-European Journal of Enterprise Technologies*, vol. 3, no. 4 (105), pp. 6–11. <https://doi.org/10.15587/1729-4061.2020.203590>

About the authors

Georgij S. Zavalin

Head of Data Mining Department, Intern Researcher at Strategic Planning Laboratory, Vladivostok State University, 41, Gogolya Str., Vladivostok, 690014, Russia;

E-mail: georgiy.zavalin08@vvsu.ru

ORCID: 0000-0003-4519-0242

Olga V. Nedoluzhko

Cand. Sci. (Econ.), Associate Professor;

Associate Professor, Economics and Management Department, Associate Professor, Mathematics and Modeling Department, Researcher, Strategic Planning Laboratory, Vladivostok State University, 41, Gogolya Str., Vladivostok, 690014, Russia;

E-mail: olga.nedoluzhko25@gmail.com

ORCID: 0000-0002-3622-7501

Konstantin S. Solodukhin

Dr. Sci. (Econ.), Professor;

Professor, Mathematics and Modeling Department, Head of Strategic Planning Laboratory, Vladivostok State University, 41, Gogolya Str., Vladivostok, 690014, Russia;

E-mail: k.solodukhin@mail.ru

ORCID: 0000-0003-3619-1219

DOI: 10.17323/2587-814X.2023.3.70.86

Имитационная модель интеллектуальной транспортной системы «умного города» с адаптивным управлением светофорами на основе нечеткой кластеризации

А.Л. Бекларян^{a, b} 

E-mail: abeklaryan@hse.ru

Л.А. Бекларян^b 

E-mail: beklar@cemi.rssi.ru

А.С. Акопов^b 

E-mail: akopovas@umail.ru

^a Национальный исследовательский университет «Высшая школа экономики»
Адрес: Россия, 101000, г. Москва, ул. Мясницкая, д. 20

^b Центральный экономико-математический институт, Российская академия наук
Адрес: Россия, 117418, г. Москва, Нахимовский проспект, д. 47

Аннотация

В данной статье представлена новая имитационная модель интеллектуальной транспортной системы (ИТС) «умного города» с адаптивным управлением светофорами. Предложенная транспортная модель, реализованная в AnyLogic, позволяет изучать поведение взаимодействующих агентов: транспортных средств (ТС) и пешеходов (П) в рамках многоагентной ИТС типа «Манхэттенской решетки». Пространственная динамика агентов в подобной ИТС описывается с использованием систем конечно-разностных уравнений с переменной структурой с учетом регулирующего воздействия «умных» светофоров. Исследованы различные способы управления светофорами, направленные на максимизацию суммарного трафика выходного потока ИТС, в частности, посредством формирования требуемой длительности фаз с использованием генетического оптимизационного алгоритма, с помощью локального («слабоадаптивного») управления переключениями и на основе предложенного алгоритма нечеткой кластеризации. Исследуются возможности оптимизации характеристик систем индивидуального управления

поведением светофоров при различных сценариях, в частности, для ИТС с пространственно-однородными и периодическими характеристиками. Для определения наилучших значений индивидуальных параметров систем управления светофорами, таких как, длительности фаз, радиусы наблюдений за транспортными и пешеходными потоками, пороговые коэффициенты, количество кластеров и др., применяется ранее предложенный параллельный генетический оптимизационный алгоритм с вещественным кодированием (класса RCGA). Предложенный метод адаптивного управления светофорами на основе нечеткой кластеризации демонстрирует большую эффективность в сравнении с известными методами коллективного воздействия и локального («слабоадаптивного») управления. Результаты работы могут быть рассмотрены как компонента системы принятия решений в управления городскими службами.

Ключевые слова: интеллектуальная транспортная система, «умный город», «умные светофоры», агентное моделирование, адаптивное управление, нечеткая кластеризация, AnyLogic

Цитирование: Бекларян А.Л., Бекларян Л.А., Акопов А.С. Имитационная модель интеллектуальной транспортной системы «умного города» с адаптивным управлением светофорами на основе нечеткой кластеризации // Бизнес-информатика. 2023. Т. 17. № 3. С. 70–86. DOI: 10.17323/2587-814X.2023.3.70.86

Введение

В настоящее время наблюдается рост потребностей в проектировании и внедрении интеллектуальных транспортных систем (ИТС) для «умного города», обусловленный постоянно увеличивающимся трафиком, обуславливающим формирование множественных дорожных заторов. При этом одним из наиболее перспективных направлений эволюционного развития ИТС является использование «умных» светофоров, анализирующих динамику и структуру транспортных и пешеходных потоков [1].

Известны различные подходы к рациональному управлению светофорами, в частности, на основе информационного обмена [2], с использованием методов машинного обучения с подкреплением [3], на основе смешанного целочисленного программирования [4, 5], с помощью генетических и роевых оптимизационных алгоритмов [6–8], а также искусственных нейронных сетей (ИНС), нечеткой логики, кластеризации и адаптивного управления в ИТС [11–13].

Для изучения поведения и оптимизации характеристик ИТС применяются различные комбинированные подходы, например, методы агентного и дискретно-событийного моделирования, поддерживаемые в AnyLogic [14, 15], совместное управление сигналами светофоров и траекториями

движения транспортных средств [16], адаптивное управление на основе прогнозной модели и обучения с подкреплением [17]. Вместе с тем, большинство подобных подходов применяется для ИТС с упрощенной конфигурацией, например, для двух последовательно расположенных перекрестков [15], одного перекрестка, состоящего из двух дорог и т.д. [17]. Различные сценарии, определяющие периодическую динамику взаимодействующих транспортных и пешеходных, при этом не учитываются.

Как правило, существенные трудности возникают при управлении характеристиками ИТС с более сложной геометрией типа «Манхэттенская решетка» [18, 19]. В подобных ИТС, несогласованное управление состояниями хотя бы одного светофора, как правило, приводит к изменению скорости ТС и плотности дорожного движения на всех связанных маршрутах. При этом, для максимизации трафика выходного потока необходимо осуществлять эффективное управление светофорами, обеспечивая периодическую приоритизацию между ТС и пешеходами. Так, например, при скоплении значительного числа людей на регулируемом пешеходном переходе, обосновано включение разрешающего сигнала светофора (подобный подход, в частности, уже успешно применяется в уличной дорожной сети некоторых городов в Австрии¹). При этом, основное назначение «умных светофо-

¹ <https://www.wien.gv.at/english/transportation-urbanplanning/diversity-themed-traffic-lights.html>

ров» состоит в мониторинге транспортных потоков и выборе оптимальных моментов времени для переключения регулирующих сигналов. Наибольшие трудности при управлении транспортными потоками обусловлены эффектом «волнового снижения скорости» [20], когда в результате торможения ТС у светофора, все последующие за ним водители непреднамеренно стремятся к увеличению безопасной дистанции, способствуя формированию дорожных заторов. Поэтому необходимо изучать неоднородную пространственную динамику агентов и использовать данные о структуре транспортных и пешеходных потоков для адаптивного управления светофорами.

В данной статье, предложена новая имитационная модель гетерогенных транспортных потоков в «умном городе» с адаптивным управлением поведением светофоров на основе нечеткой кластеризации. В рамках подобной модели, принятие индивидуальных решений о переключении регулирующих сигналов светофоров основано на нечеткой оценке дорожной ситуации, включающей эволюционную динамику как транспортных, так и пешеходных потоков (т.е. при равном приоритете по отношению к автомобилям и пешеходам). При этом, решается важная задача по максимизации суммарного трафика выходного потока при различных сценариях, в частности, для ИТС с пространственно-однородными и периодическими характеристиками потоков.

Представленные в работе сценарии, соответствующие оптимальные управления, а также, в целом, предложенная универсальная имитационная модель с возможностью дальнейшей модификации изучаемой геометрии перекрестков, как видится авторам, может быть рассмотрена как элемент комплексной системы принятия решений в управлении городскими службами.

1. Описание модели

Рассматривается ключевой фрагмент многоагентной транспортной системы типа «Манхэттенская решетка», состоящий из четырех взаимосвязанных узлов-перекрестков, допускающих произвольную смену направлений движения ТС, т.е. движение по прямой, повороты налево и направо, а также разворот и движение в обратном направлении (рис. 1).

Ранее, в [18, 19], исследована динамика транспортных потоков в системах класса «Манхэттенская решетка» и предложены различные способы

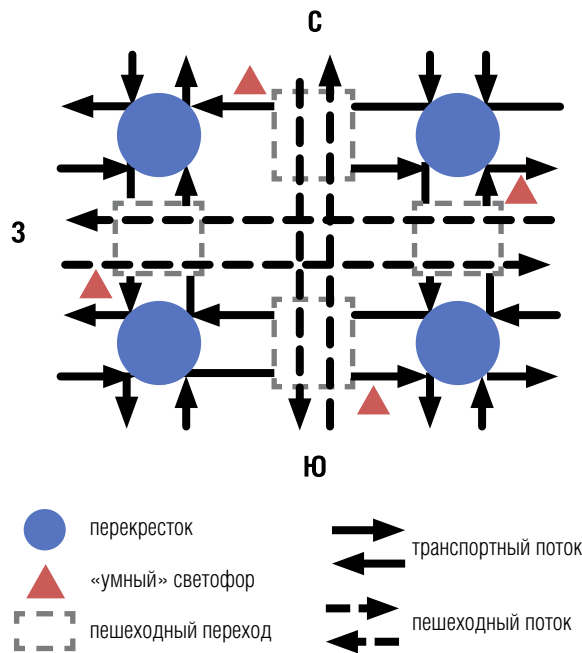


Рис. 1. Общая схема многоагентной транспортной системы типа «Манхэттенская решетка» с регулируемым движением на пешеходных переходах.

их оптимизации, преимущественно основанные на управлении маршрутами ТС, т.е. поиске и назначении оптимальных маршрутов для каждого агента-ТС [18], в том числе, с использованием генетических алгоритмов [19]. При этом, одним из важных способов снижения нагрузки на ИТС, является улучшение маневренности ТС, в том числе за счет выбора наименее загруженных полос движения, определение которых реализуется с использованием алгоритма нечеткой кластеризации [21, 22]. Существующие методы улучшения трафика дорожного потока реализуемы преимущественно для беспилотных транспортных средств (БТС), которым можно «назначить» оптимальный маршрут в зависимости от текущей ситуации. На пространственную динамику обычных транспортных средств (ОТС), наибольшее влияние оказывают «умные» светофоры, регулирующие движение транспортных и пешеходных потоков. В частности, они позволяют эффективно перераспределять потоки ТС на перекрестках и пешеходных переходах, препятствуя формированию дорожных заторов.

Подобные дорожные заторы, образуются в основном в результате «волнового снижения скорости»,

впервые исследованного в [20] и проиллюстрированного на *рисунке 2*. При торможении ТС, например, у стоп-линии регулируемого пешеходного перехода или перед ближайшим препятствием в виде другого ТС (*рис. 2*), следующий за ним автомобиль, из-за психологических особенностей реакции водителя, как правило, затормозит сильнее, чтобы сохранить безопасную дистанцию увеличив радиус своего личного пространства. Далее, эффект нарастающего увеличения «безопасной» дистанции («расширения» личного пространства) распространяется по цепочке, уменьшая скорость потока по мере удаления от первоначального источника затора (светофора), вплоть до полной остановки.

Для моделирования пространственной динамики агентов в ИТС (ТС и пешеходов) можно использовать системы конечно-разностных уравнений с переменной структурой [21, 23], что позволяет учесть различные сценарии взаимодействия ТС друг с другом и с внешней средой (типа V2V, V2P, V2I и т.д.) и влияние радиуса личного пространства каждого агента.

Приведем краткое формальное описание разработанной имитационной модели движения ТС с учетом влияния «умных» светофоров, регулирующих трафик транспортного и пешеходного потока в ИТС.

Пусть

$T = \{t_0, t_1, \dots, |T|\}$ – набор временных моментов (в мин.), где $|T|$ – общее количество временных моментов, $t_0 \in T, t_{|T|} \in T$ – начальные и конечные моменты времени;

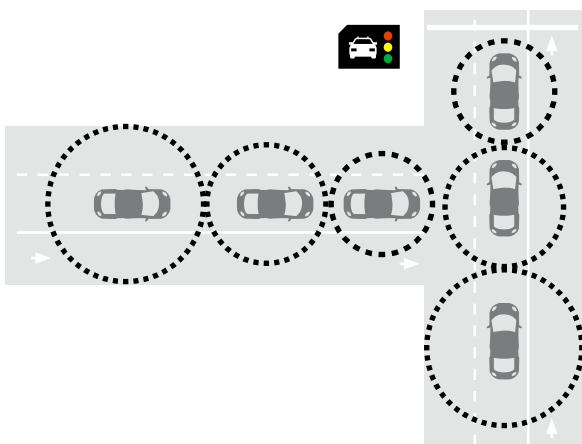


Рис. 2. Иллюстрация эффекта «волнового снижения скорости» в дорожной сети с регулируемым пешеходным переходом.

$L = \{l_0, l_1, \dots, l_{|L|}\}$ – набор индексов «умных» светофоров, где $|L|$ – общее количество «умных» светофоров;

$s_l(t_{k-1}) \in \{1, 2, 3\}, l \in L$ – состояния фаз l -го «умного» светофора в момент $t_{k-1} (t_{k-1} \in T)$: $s_l(t_{k-1}) = 1$ – запрещающий (для агентов-ТС) сигнал светофора («красный»), $s_l(t_{k-1}) = 2$ – предупреждающий сигнал светофора («желтый»), $s_l(t_{k-1}) = 3$ – разрешающий (для агентов-ТС) сигнал светофора («зеленый»);

$\{\tau_{l1}, \tau_{l2}, \tau_{l3}\} \in T, l \in L$ – длительности фаз l -го «умного» светофора (в сек.) (*управляющий параметр модели*);

$\tilde{\tau}_l$ – минимально необходимая (для обеспечения безопасного трафика) длительность основной фазы («красного» или «зеленого») для «умных» светофоров (в сек.) (*управляющий параметр модели*);

$\{P_l(t_{k-1}), V_l(t_{k-1})\}$ – суммарное количество пешеходов и ТС, соответственно, находящихся в зонах мониторинга l -го «умного» светофора в момент $t_{k-1} (t_{k-1} \in T)$;

$\{\tilde{V}_l(t_{k-1}), \tilde{D}_l(t_{k-1})\}$ – суммарное количество ТС в кластерах и среднее межкластерное расстояние, вычисляемые с использованием алгоритма нечеткой кластеризации применительно к агентам, находящимся в зонах мониторинга l -го «умного» светофора в момент $t_{k-1} (t_{k-1} \in T)$;

$\delta_l, l \in L$ – пороговое соотношение между количеством пешеходов на переходе, регулируемым l -м «умным» светофором и суммарным числом агентов-ТС, планирующих переехать данный переход (в любом направлении), при котором необходимо включить разрешающий сигнал светофора (*управляющий параметр модели*);

$\alpha_l, l \in L$ – коэффициент значимости среднего межкластерного расстояния (для ТС), оцениваемого с использованием алгоритма нечеткой кластеризации при управлении транспортными потоками, регулируемые l -м «умным» светофором (*управляющий параметр модели*).

Состояние фазы l -го «умного» светофора ($l \in L$) в момент $t_k (t_k \in T)$ задается в соответствии со следующими правилами:

$$s_l(t_k) = \begin{cases} 1, & \text{если выполняется I, II, или III,} \\ 2, & \text{если выполняется IV,} \\ 3, & \text{если выполняется V, VI, или VII,} \end{cases} \quad (1)$$

при условиях:

I. $(t_k \leq t_{k-1} + \tau_{11}$ и $s_l(t_{k-1}) = 1$) или $(t_k > t_{k-1} + \tau_{12} + \tau_{13}$ и $s_l(t_{k-1}) = 3$) – при **первом способе** управления длительностью фаз, на основе **коллективного воздействия на светофоры**;

II. $((t_k \leq t_{k-1} + \tau_{11}$ и $s_l(t_{k-1}) = 1$) или $(t_k > t_{k-1} + \tau_{12} + \tau_{13}$ и $s_l(t_{k-1}) = 3$)) или

$$\left(\frac{P_l(t_{k-1})}{V_l(t_{k-1})} > \delta_l \text{ и } s_l(t_{k-1}) = 3 \text{ и } t_k > t_{k-1} + \tilde{\tau}_l \right) - \text{ при}$$

втором способе управления длительностью фаз на основе локального («слабоадаптивного») управления переключениями светофоров с учетом **приоритетизации трафика пешеходного потока**;

III. $((t_k \leq t_{k-1} + \tau_{11}$ и $s_l(t_{k-1}) = 1$) или

$$(t_k > t_{k-1} + \tau_{12} + \tau_{13} \text{ и } s_l(t_{k-1}) = 3)) \text{ или}$$

$$\left(\frac{P_l(t_{k-1})}{\tilde{V}_l(t_{k-1})(\tilde{D}_l(t_{k-1}))^{-\alpha_l}} > \delta_l \text{ и } s_l(t_{k-1}) = 3 \text{ и } t_k > t_{k-1} + \tilde{\tau}_l \right) -$$

при **третьем (адаптивном) способе** управления длительностью фаз на основе алгоритма нечеткой кластеризации с учетом **приоритетизации трафика пешеходного потока**;

IV. $((t_k > t_{k-1} + \tau_{11}$ и $s_l(t_{k-1}) = 1$) или $(t_k > t_{k-1} + \tau_{13}$ и $s_l(t_{k-1}) = 3$)) или $(t_k \leq t_{k-1} + \tau_{12}$ и $s_l(t_{k-1}) = 2)$ – что означает, что истекло время действия одного из основных сигналов светофора («красного» или «зеленого») или продолжает действовать, ранее включенный предупреждающий («желтый») сигнал;

V. $(t_k \leq t_{k-1} + \tau_{13}$ и $s_l(t_{k-1}) = 3$) или $(t_k > t_{k-1} + \tau_{12} + \tau_{11}$ и $s_l(t_{k-1}) = 1)$ – при **первом способе** управления длительностью фаз, на основе **коллективного воздействия на светофоры**;

VI. $((t_k \leq t_{k-1} + \tau_{13}$ и $s_l(t_{k-1}) = 3$) или

$$(t_k > t_{k-1} + \tau_{12} + \tau_{11} \text{ и } s_l(t_{k-1}) = 1)) \text{ или}$$

$$\left(\frac{P_l(t_{k-1})}{V_l(t_{k-1})} < \delta_l \text{ и } s_l(t_{k-1}) = 1 \text{ и } t_k > t_{k-1} + \tilde{\tau}_l \right) - \text{ при}$$

втором способе управления длительностью фаз на основе локального («слабоадаптивного») управления переключениями светофоров с учетом **приоритетизации трафика транспортного потока**;

VII. $((t_k \leq t_{k-1} + \tau_{11}$ и $s_l(t_{k-1}) = 3$) или

$$(t_k > t_{k-1} + \tau_{12} + \tau_{11} \text{ и } s_l(t_{k-1}) = 1)) \text{ или}$$

$\left(\frac{P_l(t_{k-1})}{\tilde{V}_l(t_{k-1})(\tilde{D}_l(t_{k-1}))^{-\alpha_l}} < \delta_l \text{ и } s_l(t_{k-1}) = 1 \text{ и } t_k > t_{k-1} + \tilde{\tau}_l \right) -$
при **третьем (адаптивном) способе** управления длительностью фаз, на основе алгоритма нечеткой кластеризации с учетом приоритетизации трафика транспортного потока.

Суммарное количество ТС и пешеходов, находящихся в зонах мониторинга l -го «умного» светофора ($l \in L$) в момент t_k ($t_k \in T$) вычисляемые при втором способе локального («слабоадаптивного») управления переключениями равны

$$V_l(t_k) = \sum_{i=1}^{|I|} v_{il}(t_k), \quad P_l(t_k) = \sum_{i=1}^{|\tilde{I}|} p_{i\tilde{l}}(t_k), \quad (2)$$

где

$$v_{il}(t_k) = \begin{cases} 1, & \text{если } d_{il}(t_k) \leq R_{1l}, \\ 0, & \text{если } d_{il}(t_k) > R_{1l}, \end{cases} \quad (3)$$

$$p_{i\tilde{l}}(t_k) = \begin{cases} 1, & \text{если } \tilde{d}_{i\tilde{l}}(t_k) \leq R_{1l}, \\ 0, & \text{если } \tilde{d}_{i\tilde{l}}(t_k) > R_{1l}, \end{cases}$$

где

$I = \{i_1, i_2, \dots, i_{|I|}\}$ – набор индексов агентов-транспортных средств (ТС), где $|I|$ – общее количество транспортных средств;

$\tilde{I} = \{\tilde{i}_1, \tilde{i}_2, \dots, \tilde{i}_{|\tilde{I}|}\}$ – набор индексов агентов-пешеходов, где $|\tilde{I}|$ – общее количество пешеходов;

$\{R_{1l}, R_{2l}\}, l \in L$ – радиусы зон мониторинга трафика дорожного и пешеходного потока, соответственно, для l -го «умного» светофора (*управляющий параметр модели*);

$\{d_{il}(t_k), \tilde{d}_{i\tilde{l}}(t_k)\}, i \in I, \tilde{i} \in \tilde{I}, l \in L$ – расстояния от i -го агента-ТС и \tilde{i} -го агента-пешехода до l -го «умного» светофора в момент t_k ($t_k \in T$).

Суммарное количество ТС в кластерах и среднее межкластерное расстояние для транспортных потоков, находящихся в зонах мониторинга l -го «умного» светофора ($l \in L$) в момент t_k ($t_k \in T$) вычисляемые при третьем способе адаптивного управления переключениями с использованием алгоритма нечеткой кластеризации равны

$$\tilde{V}_l(t_k) = \sum_{i=1}^{|I|} \sum_{c_j=1}^{|C_j|} \tilde{v}_{ic_j}(t_k), \quad \tilde{D}_l(t_k) = \frac{1}{|C_l|} \sum_{c_j=1}^{|C_l|} \sum_{\tilde{c}_j=1}^{\hat{c}_{c_j}} \hat{d}_{c_j\tilde{c}_j}(t_k), \quad (4)$$

где

$C_l = \{c_{1l}, c_{2l}, \dots, c_{|C_l|l}\}, l \in L$ – набор индексов кластеров определяемых для анализа дорожной ситуации в районе расположения l -го «умного»

светофора с использованием алгоритма нечеткой кластеризации, где $|C_l|$ – общее количество кластеров (*управляющий параметр модели*);

$\tilde{v}_{c_i}(t_k)$, $i \in I$, $\tilde{i} \in \tilde{I}$, $l \in L$ – суммарное количество ТС, принадлежащих c_l -му кластеру в момент t_k ($t_k \in T$);

$\hat{d}_{c_i \tilde{c}_i}(t_k)$, $c_l, \tilde{c}_l \in C_l$, $c_l \neq \tilde{c}_l$, $l \in L$ – попарные расстояния между центрами кластеров, относящихся к l -му «умному» светофору в момент t_k ($t_k \in T$).

Пространственная динамика агентов-ТС и пешеходов может быть смоделирована с использованием систем конечно-разностных уравнений с переменной структурой с учетом регулирующего воздействия «умных» светофоров.

Пусть

$\{x_{il}(t_k), y_{il}(t_k)\}$, $\{\tilde{x}_{\tilde{i}l}(t_k), \tilde{y}_{\tilde{i}l}(t_k)\}$, $i \in I$, $\tilde{i} \in \tilde{I}$, $l \in L$ – координаты i -го агента-ТС и \tilde{i} -го агента-пешехода находящегося в зоне мониторинга l -го «умного» светофора в момент t_k ($t_k \in T$);

$\{v_i(t_{k-1}), \tilde{v}_{\tilde{i}}(t_{k-1})\}$, $i \in I$, $\tilde{i} \in \tilde{I}$ – предпочитаемая скорость i -го агента-ТС и \tilde{i} -го агента-пешехода в момент t_{k-1} ($t_{k-1} \in T$);

$\{r_i(t_{k-1}), \tilde{r}_{\tilde{i}}(t_{k-1})\}$, $i \in I$, $\tilde{i} \in \tilde{I}$, $l \in L$ – радиусы личных пространств i -го агента-ТС и \tilde{i} -го агента-пешехода, значения которых зависят от плотности транспортного (пешеходного) потока, состоящего из агентов, снижающих свою скорость и расположенных по ходу движения (см. *рис. 2*) в момент t_{k-1} ($t_{k-1} \in T$);

$\{m_{ib}(t_k), \tilde{m}_{\tilde{i}b}(t_k)\}$, $i \in I$, $\tilde{i} \in \tilde{I}$, $b \in I \cup \tilde{I}$ – расстояния от i -го агента-ТС и \tilde{i} -го агента-пешехода до ближайшего b -го агента-препятствия в момент t_{k-1} ($t_{k-1} \in T$);

$\{w_i(t_{k-1}), \tilde{w}_{\tilde{i}}(t_{k-1})\}$, $\{q_i(t_{k-1}), \tilde{q}_{\tilde{i}}(t_{k-1})\} \in \{-1, 0, 1\}$, $i \in I$, $\tilde{i} \in \tilde{I}$ – параметры, определяющие направления движения i -го агента-ТС и \tilde{i} -го агента-пешехода в момент t_{k-1} ($t_{k-1} \in T$);

$w_i(t_{k-1}), \tilde{w}_{\tilde{i}}(t_{k-1}) = -1$ – при движении в направлении **ВЗ** (см. *рис. 1*),

$w_i(t_{k-1}), \tilde{w}_{\tilde{i}}(t_{k-1}) = 0$ – при движении в направлении **СЮ** или **ЮС**,

$w_i(t_{k-1}), \tilde{w}_{\tilde{i}}(t_{k-1}) = 1$ – при движении в направлении **ЗВ**, $q_i(t_{k-1}), \tilde{q}_{\tilde{i}}(t_{k-1}) = -1$ – при движении в направлении **ЮС**,

$q_i(t_{k-1}), \tilde{q}_{\tilde{i}}(t_{k-1}) = 0$ – при движении в направлении **ЗВ** или **ВЗ**,

$q_i(t_{k-1}), \tilde{q}_{\tilde{i}}(t_{k-1}) = 1$ – при движении в направлении **ЮС**;

λ – коэффициент задающий соотношение масштабов реального и модельного времени.

Пространственная динамика i -го агента-ТС ($i \in I$) и \tilde{i} -го агента-пешехода ($\tilde{i} \in \tilde{I}$), находящихся в зоне мониторинга l -го «умного» светофора ($l \in L$) в момент времени t_k ($t_k \in T$) без учета внутреннего маневрирования (связанного с обгоном, сменой полос движения и т.д.) задается следующей системой конечно-разностных уравнений с переменной структурой:

$$x_{il}(t_k) = \begin{cases} x_{il}(t_{k-1}) + w_i(t_{k-1})\lambda v_i(t_{k-1}), & \text{если выполняется VIII,} \\ x_{il}(t_{k-1}), & \text{если выполняется IX,} \end{cases} \quad (5)$$

$$y_{il}(t_k) = \begin{cases} y_{il}(t_{k-1}) + q_i(t_{k-1})\lambda v_i(t_{k-1}), & \text{если выполняется VIII,} \\ y_{il}(t_{k-1}), & \text{если выполняется IX,} \end{cases} \quad (6)$$

$$\tilde{x}_{\tilde{i}l}(t_k) = \begin{cases} \tilde{x}_{\tilde{i}l}(t_{k-1}), & \text{если выполняется X,} \\ \tilde{x}_{\tilde{i}l}(t_{k-1}) + \tilde{w}_{\tilde{i}}(t_{k-1})\lambda \tilde{v}_{\tilde{i}}(t_{k-1}), & \text{если выполняется XI,} \end{cases} \quad (7)$$

$$\tilde{y}_{\tilde{i}l}(t_k) = \begin{cases} \tilde{y}_{\tilde{i}l}(t_{k-1}), & \text{если выполняется X,} \\ \tilde{y}_{\tilde{i}l}(t_{k-1}) + \tilde{q}_{\tilde{i}}(t_{k-1})\lambda \tilde{v}_{\tilde{i}}(t_{k-1}), & \text{если выполняется XI,} \end{cases} \quad (8)$$

$i \in I, \tilde{i} \in \tilde{I}, b \in I \cup \tilde{I}, l \in L$,
где

VIII. $s_i(t_{k-1}) = 3$ и $m_{ib}(t_{k-1}) > (r_i(t_{k-1}) + r_b(t_{k-1}))$ для ближайшего агента ($b \in I \cup \tilde{I}$), что означает, что действует разрешающий (для агентов-ТС) сигнал светофора («зеленый») и у i -го агента-ТС ($i \in I$) на пути нет препятствий в виде других ТС или пешеходов;

IX. $s_i(t_{k-1}) = 1$ или $m_{ib}(t_{k-1}) \leq (r_i(t_{k-1}) + r_b(t_{k-1}))$ для ближайшего агента ($b \in I \cup \tilde{I}$), что означает, что действует запрещающий (для агентов-ТС) сигнал светофора («красный») или у i -го агента-ТС ($i \in I$) на пути нет есть препятствие в виде другого ТС или пешехода;

X. $s_i(t_{k-1}) = 1$ и $\tilde{m}_{\tilde{i}b}(t_{k-1}) > (\tilde{r}_{\tilde{i}}(t_{k-1}) + r_b(t_{k-1}))$ для ближайшего агента ($b \in I \cup \tilde{I}$), что означает, что действует запрещающий (для агентов-ТС) сигнал светофора («красный») и у \tilde{i} -го агента-пешехода ($\tilde{i} \in \tilde{I}$) на пути нет препятствий в виде других пешеходов или ТС;

XI. $s_i(t_{k-1}) = 1$ или $\tilde{m}_{\tilde{i}b}(t_{k-1}) \leq (\tilde{r}_{\tilde{i}}(t_{k-1}) + r_b(t_{k-1}))$ для ближайшего агента ($b \in I \cup \tilde{I}$), что означает, что действует разрешающий (для агентов-ТС) сигнал светофора («зеленый») или у \tilde{i} -го агента-

пешехода ($\tilde{i} \in \tilde{I}$) на пути есть препятствие в виде другого пешехода или ТС.

Суммарный трафик выходного потока, который должен быть максимизирован равен

$$N = \sum_{t_k=0}^{|T|} \left(\sum_{i=1}^{|I|} n_i + \sum_{\tilde{i}=1}^{|\tilde{I}|} \tilde{n}_{\tilde{i}} \right), \quad (9)$$

где

$$n_i(t_k) = \begin{cases} 1, & \text{если } \{x_i(t_{k-1}), y_i(t_{k-1})\} \notin \{X, Y\}, \\ 0, & \text{если } \{x_i(t_{k-1}), y_i(t_{k-1})\} \in \{X, Y\}, \end{cases} \quad (10)$$

$$\tilde{n}_{\tilde{i}}(t_k) = \begin{cases} 1, & \text{если } \{\tilde{x}_{\tilde{i}}(t_{k-1}), \tilde{y}_{\tilde{i}}(t_{k-1})\} \notin \{X, Y\}, \\ 0, & \text{если } \{\tilde{x}_{\tilde{i}}(t_{k-1}), \tilde{y}_{\tilde{i}}(t_{k-1})\} \in \{X, Y\}, \end{cases} \quad (11)$$

где

$\{x_i(t_{k-1}), y_i(t_{k-1})\}, \{\tilde{x}_{\tilde{i}}(t_{k-1}), \tilde{y}_{\tilde{i}}(t_{k-1})\}, i \in I, \tilde{i} \in \tilde{I}$ – координаты i -го агента-ТС и \tilde{i} -го агента-пешехода в ИТС в момент t_{k-1} ($t_{k-1} \in T$);

$\{X, Y\}$ – множество всех координат цифровой дорожной сети ИТС.

Тогда, можно сформулировать следующую оптимизационную задачу, решаемую с учетом выбранного способа управления «умными» светофорами.

Задача А. Требуется максимизировать суммарный трафик выходного потока по набору управляющих параметров $\{\tau_{11}, \tau_{12}, \tau_{13}, \tilde{\tau}_l, \delta_l, R_{1l}, R_{2l}, |C_l|, \alpha_l\}$:

$$\max_{\{\tau_{11}, \tau_{12}, \tau_{13}, \tilde{\tau}_l, \delta_l, R_{1l}, R_{2l}, |C_l|, \alpha_l\}} N \quad (12)$$

при ограничениях

$$\tau_{11}, \tau_{12}, \tau_{13}, \tilde{\tau}_l \in [\underline{\tau}, \bar{\tau}], \delta_l \in [\underline{\delta}, \bar{\delta}], R_{1l} \in [\underline{R}_1, \bar{R}_1], R_{2l} \in [\underline{R}_2, \bar{R}_2], |C_l| \in [\underline{C}, \bar{C}], \alpha_l \in [\underline{\alpha}, \bar{\alpha}],$$

где $\{\underline{\tau}, \underline{\delta}, \underline{R}_1, \underline{R}_2, \underline{C}, \underline{\alpha}\}, \{\bar{\tau}, \bar{\delta}, \bar{R}_1, \bar{R}_2, \bar{C}, \bar{\alpha}\}$ – нижние и верхние граничные значения управляющих параметров модели.

Для решения **Задачи А** используется ранее предложенный генетический оптимизационный алгоритм вещественного кодирования (класса RCGA) [20, 25], агрегированный по целевому функционалу с разработанной имитационной моделью транспортной системы, реализованной в AnyLogic.

2. Алгоритм нечеткой кластеризации

Для оценки структуры транспортного потока и адаптивного управления «умными «светофорами» предлагается использовать алгоритм нечеткой кластеризации (Fuzzy C-means) [21, 22, 26, 27]. Вы-

бор данного алгоритма обусловлен, прежде всего, возможностью принятия во внимание различных характеристик движущихся ТС формировании кластеров, в частности, плотности, скорости, удаленности от светофора, регулирующего движение на переходном переходе и др. Включение подобных характеристик в кластерный анализ, позволяет достичь максимального «правдоподобия» при оценке структуры транспортного потока. В отличие от классических алгоритмов, Fuzzy C-means не относит объект однозначно к какому-нибудь кластеру, а сопоставляет каждому кластеру вероятность отнесения к нему наблюдаемых объектов, формируя так называемую матрицу принадлежности.

Укрупненная схема предложенного алгоритма нечеткой кластеризации представлена на *рисунке 3*. Важным отличием разработанного алгоритма от ранее известных, является то, что его ключевые характеристики (например, количество кластеров, радиус зоны мониторинга трафика дорожного потока и др.), вычисляются с использованием генетического оптимизационного алгоритма (класса RCGA) в рамках решения основной задачи по максимизации трафика выходного потока. В результате, результаты нечеткой кластеризации непосредственно влияют на возможности поиска оптимальных решений для изучаемой ИТС.

На *рисунке 3* используются следующие обозначения:

- ♦ $z \in [0, 1]$ – мера нечеткости;
- ♦ $M(k)$ – матрица принадлежности на k -м шаге алгоритма, $k = 1, 2, \dots, |K|$, где $|K|$ – максимальное число итераций;
- ♦ ε – малый параметр, являющийся критерием остановки алгоритма.

Таким образом, предложенный алгоритм нечеткой кластеризации агрегирован по целевому функционалу (суммарному трафику выходного потока ИТС), с генетическим алгоритмом вещественного кодирования (*рис. 3*). RCGA использует эвристические операторы кроссинговера и мутации (например, LX, SBX, SNUM, см. [21, 24, 25]) для формирования новых потенциальных решений с лучшими характеристиками. Алгоритм Fuzzy C-means был встроен в имитационную модель ИТС, реализованную в AnyLogic и выполняется на каждом шаге модельного времени, обеспечивая оценку структуры транспортного потока, находящегося в зоне мониторинга каждого «умного» светофора.

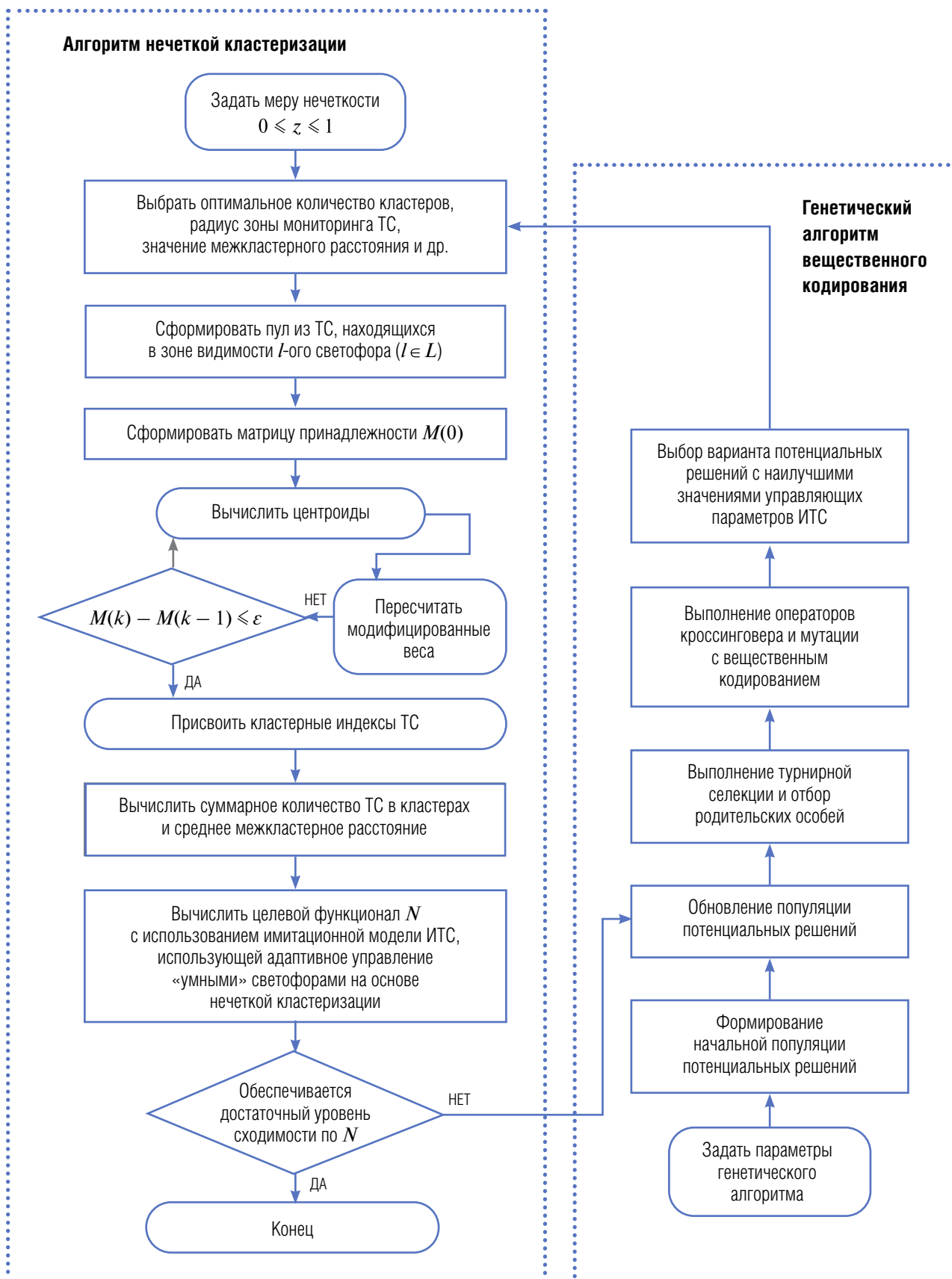


Рис. 3. Алгоритм нечеткой кластеризации для адаптивного управления светофорами.

3. Программная реализация модели

Ключевой фрагмент программной реализации приложенной имитационной модели ИТС, выполненной в среде AnyLogic, представлен на *рисунке 4*.

Важной особенностью программной реализации модели (*рис. 4*) является комбинированное использование методов дискретно-событийного и агентного моделирования, в том числе, поддерживаемых в библиотеке дорожного движения AnyLogic [28, 29]. В частности, элементы типа **carSource** и **pedSource** обеспечивают генерацию новых аген-

тов их добавление в соответствующие популяции ТС и пешеходов, элементы типа **SelectOutput** (s1, s2 на *рис. 4б*) используются для распределения транспортного потока по возможным маршрутам при достижении ТС перекрестков, элементы **carMoveTo** и **pedGoTo** перемещают агентов-ТС и пешеходов к заданной цели, в соответствии с их предопределенными характеристиками (предпочитаемой скоростью, интенсивностью прибытия и др.), **carDispose** и **pedSink** обеспечивают удаление агентов из соответствующих популяций и расчет трафика выходного потока.

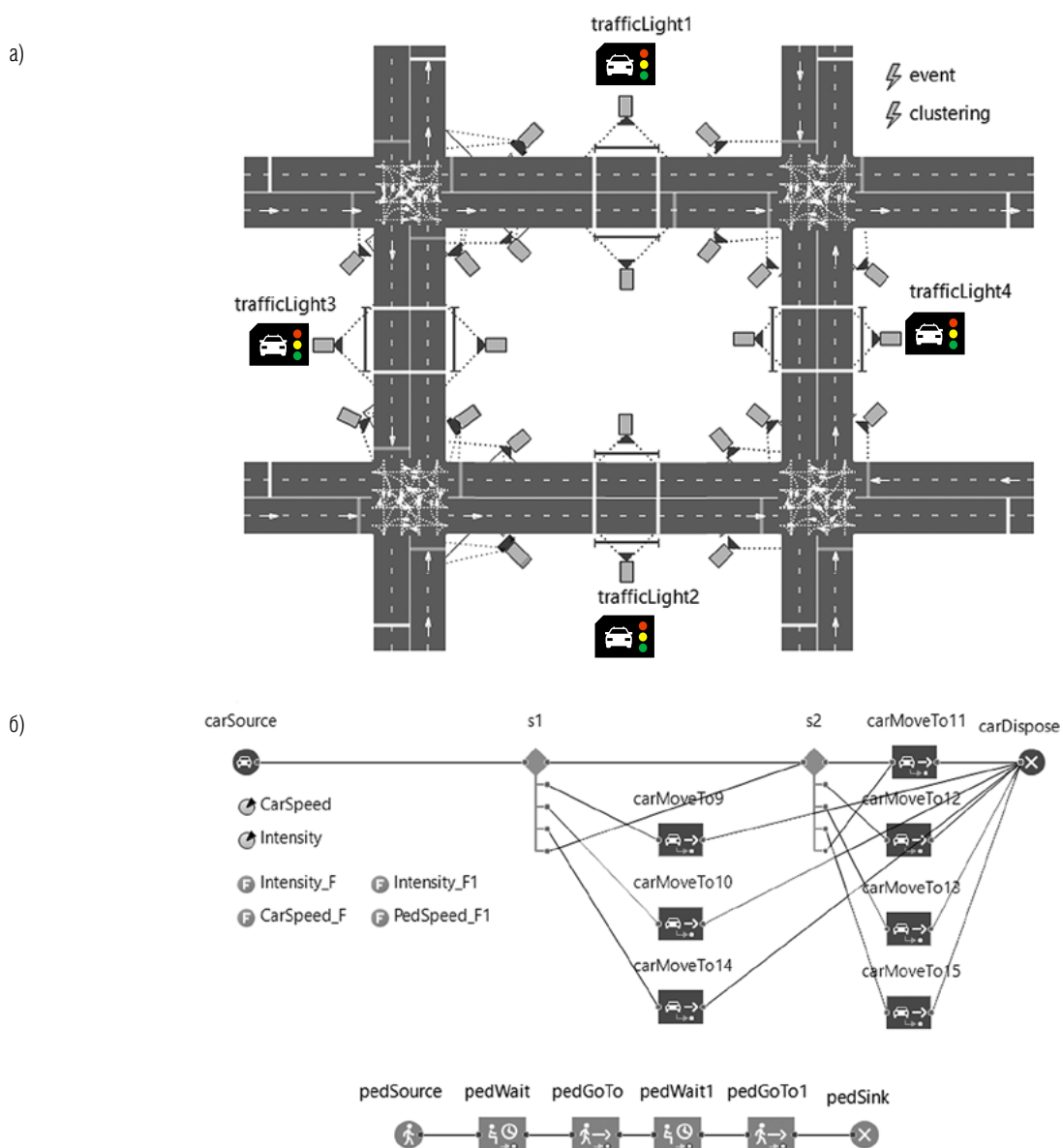


Рис. 4. Программная реализация имитационной модели ИТС в AnyLogic:
 а) схема цифровой дорожной сети с «умными» светофорами
 б) фрагмент дискретно-событийной модели движения агентов-ТС и пешеходов по заданным маршрутам.

4. Результаты оптимизационных экспериментов

Оптимизационные эксперименты были проведены для ИТС с пространственно-однородными и периодическими характеристиками при трех способах управления «умными светофорами»:

- ◆ посредством формирования требуемой длительности фаз с использованием генетического оптимизационного алгоритма;
- ◆ с помощью локального («слабоадаптивного») управления переключениями;
- ◆ на основе предложенного алгоритма нечеткой кластеризации.

Для системы с пространственно-однородными характеристиками потоков, интенсивности прибытия агентов и их предпочтительные скорости являются константами.

Интенсивность прибытия агентов в ИТС с периодическими характеристиками, имитирующей наличие и отсутствие пиковых нагрузок, вычисляется в каждый момент времени t_k ($t_k \in T$):

$$\theta(t_k) = \begin{cases} \text{truncnormal}(\hat{\theta}, \sigma, \underline{\theta}, \bar{\theta}), & \text{если } t_k \leq \frac{1}{4}|T|, \\ \tilde{\theta}, & \text{если } \frac{1}{4}|T| < t_k < \frac{3}{4}|T|, \\ \text{truncnormal}(\hat{\theta}, \sigma, \underline{\theta}, \bar{\theta}), & \text{если } t_k \geq \frac{3}{4}|T|, \end{cases} \quad (13)$$

где

$\text{truncnormal}(\hat{\theta}, \sigma, \underline{\theta}, \bar{\theta})$ – случайное значение интенсивности прибытия агентов, задаваемое с использованием усеченного нормального распределения с математическим ожиданием $\hat{\theta}$, стандартным отклонением σ , нижними и верхними граничными значениями $\underline{\theta}$, $\bar{\theta}$, соответствующее условиям экстремального трафика;

$\tilde{\theta}$ – интенсивность прибытия, соответствующее условиям нормального трафика.

Аналогичным способом задаются средние скорости движения агентов.

Основные модельные допущения (исходные данные) представлены в *таблице 1*.

На первом этапе, с использованием метода класса Монте-Карло [30], были проведены численные эксперименты для оценки чувствительности целевого функционала (суммарного трафика выходного потока) по отношению к значениям управляющих параметров ИТС с пространственно-однородными и периодическими характеристиками потоков (*рис. 5*).

Таблица 1.

Исходные данные имитационной модели

№	Параметры модели	Значения	
		ТС	пешеходы
1	Длина и ширина дорог, м.	155	
2	Количество перекрестков	4	
3	Расстояние между соседними перекрестками, м.	65	
4	Количество полос дорожного движения для каждой дороги	2	
5	Ширина разделительной полосы, м.	2	
6	Количество пешеходных переходов, регулируемых «умными» светофорами	4	
7	Период моделирования, мин.	20	
8	Интенсивность прибытия ТС и пешеходов в ИТС с пространственно-однородными характеристиками на каждый вход дорожной сети (агентов в час)	ТС	1000
		пешеходы	500
9	Предпочитаемая скорость ТС и пешеходов в ИТС с пространственно-однородными характеристиками (км/час для ТС и м/с для пешеходов)	100	0,75
10	Параметры для вычисления интенсивности прибытия ТС и пешеходов в ИТС с периодическими характеристиками потоков (агентов в час)	$\hat{\theta}$	500
		σ	100
		$\underline{\theta}$	100
		$\bar{\theta}$	1500
		$\tilde{\theta}$	500
11	Параметры для вычисления предпочитаемой скорости ТС и пешеходов в ИТС с периодическими характеристиками потоков (км/час для ТС и м/с для пешеходов)	$\hat{\theta}$	1,3
		σ	0,5
		$\underline{\theta}$	1
		$\bar{\theta}$	1,5
		$\tilde{\theta}$	1

Из *рисунка 5* следует, что суммарный трафик выходного потока чувствителен по отношению к значениям управляющих параметров ИТС, как при пространственно-однородных, так и периодических характеристиках потоков. При этом, наиболее вероятные диапазоны значений суммарного трафика выходного потока: 1800–1900 агентов (ТС и пешеходов).

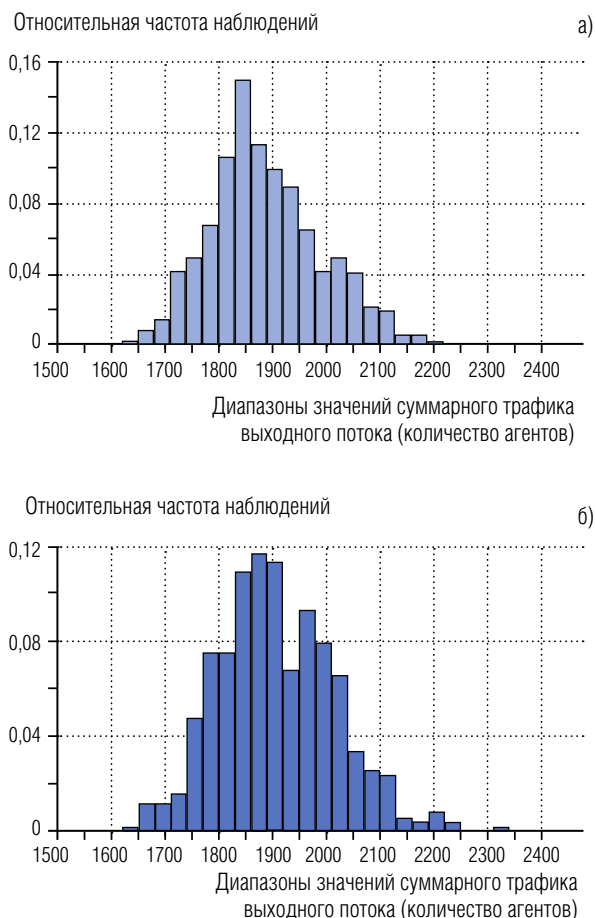


Рис. 5. Оценка чувствительности (плотность вероятности) суммарного трафика выходного потока для ИТС: а) с пространственно-однородными и б) периодическими характеристиками потоков.

На рисунке 6 представлена динамика сходимости целевой функции, полученная с использованием разработанной имитационной модели, агрегированной по целевому функционалу с генетическим алгоритмом (RCGA).

Максимально возможные значения суммарного выходного потока для ИТС с периодическими характеристиками потоков, в среднем, меньше, чем для ИТС с пространственно-однородными характеристиками (рис. 5). Полученные субоптимальные значения управляющих параметров модели, соответствующие рассмотренным выше сценариям реализации ИТС в укрупненном виде, представлены в таблице 2.

Из рисунка 6 и таблицы 2 следует, что наиболее перспективным способом управления «умными» светофорами, является адаптивное управление переключениями на основе нечеткой кластеризации.

Предложенный подход демонстрирует свою эффективность даже для ИТС с периодическими характеристиками потоков, обеспечивая лучшее конечное значение целевой функции.

Заключение

В данной статье представлена новая имитационная модель интеллектуальной транспортной системы (ИТС) «умного города» с адаптивным управлением светофорами. Предложена модель движения ансамбля ТС с использованием систем конечно-разностных уравнений с переменной структурой с учетом регулирующего воздействия «умных» светофоров. Для оценки структуры транспортного потока и адаптивного управления «умными светофорами» предложен алгоритм нечеткой кластеризации, ключевые характеристики которого, вычисляются с использованием генетического оптимизационного алгоритма (класса RCGA) в рамках решения основной

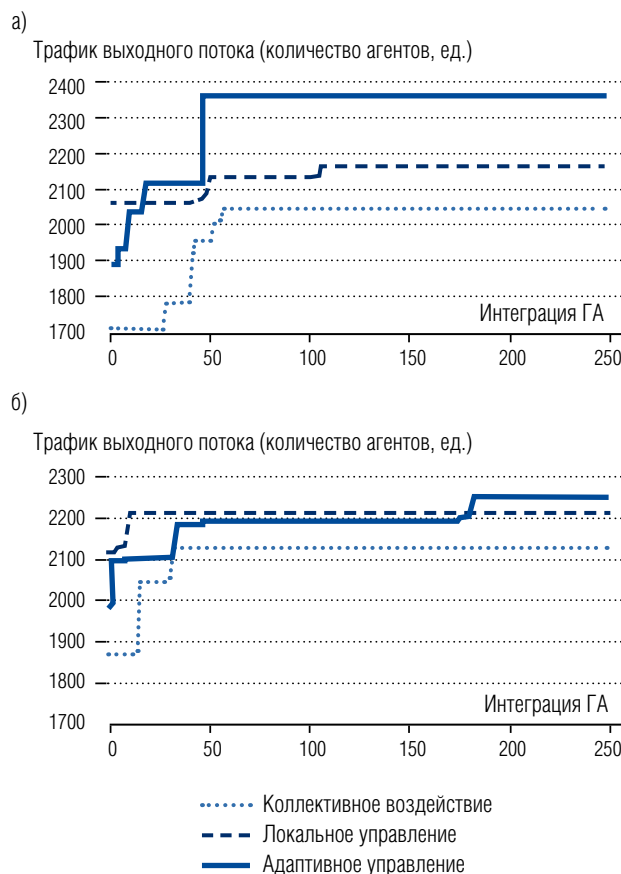


Рис. 6. Динамика сходимости целевой функции (суммарного трафика выходного потока) для ИТС: а) с пространственно-однородными и б) периодическими характеристиками потоков.

Таблица 2.

Полученные субоптимальные значения управляющих параметров модели

			Система с пространственно-однородными характеристиками потоков	Система с периодическими характеристиками потоков	
Коллективное воздействие					
Суммарный трафик, агентов			2392	2125	
Длительность фаз светофоров, сек.	первый и второй светофоры	«красный»	11,327	22,675	
		«зеленый»	25,816	18,061	
	третий и четвертый светофоры	«красный»	10,136	18,768	
		«зеленый»	11,933	103,071	
	все светофоры		«желтый»	1,098	1,874
	Локальное управление				
Суммарный трафик, агентов			2166	2209	
Минимально необходимая длительность основной фазы, мин.			1,839	4,918	
Радиус зоны мониторинга трафика дорожного потока, м.			31,59	12,25	
Радиус зоны мониторинга трафика пешеходного потока, м.			19,13	2,56	
Пороговое соотношение между количеством пешеходов на переходе и суммарным числом ТС			88,5	241,7	
Адаптивное управление на основе нечеткой кластеризации					
Суммарный трафик, агентов			2357	2246	
Минимально необходимая длительность основной фазы, мин.			1,514	3,555	
Радиус зоны мониторинга трафика дорожного потока, м.			26,11	17,89	
Радиус зоны мониторинга трафика пешеходного потока, м.			20,21	25,01	
Пороговое соотношение между количеством пешеходов на переходе и суммарным числом ТС (с поправкой на межкластерное расстояние)			162,5	255,1	
Коэффициент значимости среднего межкластерного расстояния			0,602	1	
Количество кластеров (ТС)			3	3	

задачи по максимизации трафика выходного потока. С помощью разработанной имитационной модели исследованы возможности рационального управления «умными» светофорами, в частности, для ИТС с пространственно-однородными и периодическими характеристиками. В результате, на модельном примере продемонстрирована большая эффективность адаптивного управления переключениями на основе нечеткой кластеризации.

Дальнейшие исследования будут направлены на проектирование крупномасштабной агентной модели ИТС «умного города» с использованием FLAME GPU. ■

Благодарности

Исследование выполнено за счет гранта Российского научного фонда (проект № 23-11-00080).

Литература

1. Li L., Wen D., Yao D. A survey of traffic control with vehicular communications // *IEEE Transactions on Intelligent Transportation Systems*. 2014. Vol. 15. No. 1. P. 425–432. <https://doi.org/10.1109/TITS.2013.2277737>
2. Tang C., Wei X., Hao M., Zhu C., Wang R., Chen W. Traffic signal phase scheduling based on device-to-device communication // *IEEE Access*. 2018. Vol. 6. P. 47636–47645. <https://doi.org/10.1109/ACCESS.2018.2867553>
3. Tong W., Pan Z., Liu K., Yali Y., Xiumin W., Huawei H., Wu D.O. Multi-agent deep reinforcement learning for urban traffic light control in vehicular networks // *IEEE Transactions on Vehicular Technology*. 2020. Vol. 69. No. 8. P. 8243–8256. <https://doi.org/10.1109/TVT.2020.2997896>
4. Lin W-H., Wang C. An enhanced 0-1 mixed-integer LP formulation for traffic signal control // *IEEE Transactions on Intelligent Transportation Systems*. 2004. Vol. 5. No. 4. P. 238–245. <https://doi.org/10.1109/TITS.2004.838217>
5. Gartner N.H., John D.C., Gabbay L.H. Optimization of traffic signal settings by mixed-integer linear programming // *Transportation Science*. 1975. Vol. 9. No. 4. P. 321–343. <https://doi.org/10.1287/trsc.9.4.321>
6. Davydov I., Tolstykh D., Kononova P., Legkih I. Genetic based approach for Novosibirsk traffic light scheduling // *Proceedings of the 15th International Asian School-Seminar Optimization Problems of Complex Systems (OPCS), Novosibirsk, Russia, 2019*. P. 31–36. <https://doi.org/10.1109/OPCS.2019.8880158>
7. Turkey A.M., Ahmad M.S., Yusoff M.Z.M., Sabar N.R. Genetic algorithm application for traffic light control // *Information Systems: Modeling, Development, and Integration. UNISCON 2009. Lecture Notes in Business Information Processing*. Vol. 20. (eds. J. Yang, A. Ginige, H.C. Mayr, R.D. Kutsche). Berlin: Springer, Heidelberg, 2009. P. 115–120. https://doi.org/10.1007/978-3-642-01112-2_12
8. García-Nieto J., Alba E., Carolina Olivera A. Swarm intelligence for traffic light scheduling: Application to real urban areas // *Engineering Applications of Artificial Intelligence*. 2012. Vol. 25. No. 2. P. 274–283. <https://doi.org/10.1016/j.engappai.2011.04.011>
9. Royani T., Haddadnia J., Alipoor M. Control of traffic light in isolated intersections using fuzzy neural network and genetic algorithm // *International Journal of Computer and Electrical Engineering*. 2013. Vol. 5. No. 1. P. 142–146. <https://doi.org/10.7763/IJCEE.2013.V5.682>
10. Agramelal F., Sadik M., Hannani A.E., Moubarak Y. A traffic-aware street lighting system based on fuzzy logic controller // *Proceedings of the 2022 IEEE 18th International Colloquium on Signal Processing & Applications (CSPA), Selangor, Malaysia, 2022*. P. 132–137. <https://doi.org/10.1109/CSPA55076.2022.9782045>
11. Rashid H., Ashrafi M.J.F., Azizi M., Heydarinezhad M.R. Intelligent traffic light control based on clustering using Vehicular Ad-hoc Networks // *Proceedings of the 2015 7th Conference on Information and Knowledge Technology (IKT), Urmia, Iran, 2015*. P. 1–6. <https://doi.org/10.1109/IKT.2015.7288801>
12. Sachan A., Kumar N. SDN control-enabled and time-quantum-driven max-pressure approach for intersection management in smart city // *IEEE Systems Journal*. 2023. Vol. 17. No. 1. P. 1694–1702, <https://doi.org/10.1109/JSYST.2022.3211933>
13. Marino F., Leccese F., Pizzuti S. Adaptive street lighting predictive control // *Energy Procedia*. 2017. Vol. 111. P. 790–799. <https://doi.org/10.1016/j.egypro.2017.03.241>
14. Makarov V.L., Bakhtizin A.R., Epstein J.M. Agent-based modeling for a complex world. Moscow: Scientific publications department, GAUGN, 2022. <https://doi.org/10.18254/978-5-604-5843-5-4>
15. Веремчук Н.С. Элементы имитационного моделирования в вопросах оптимизации дорожного движения // *Вестник кибернетики*. 2022. Т. 4. № 48. С. 23–28. <https://doi.org/10.34822/1999-7604-2022-4-23-28>
16. Agafonov A., Yumaganov A., Myasnikov V. Cooperative control for signalized intersections in intelligent connected vehicle environments // *Mathematics*. 2023. Vol. 11. No. 6. 1540. <https://doi.org/10.3390/math11061540>
17. Мясников В.В., Агафонов А.А., Юмаганов А.С. Детерминированная прогнозная модель управления сигналами светофоров в интеллектуальных транспортных и геоинформационных системах // *Компьютерная оптика*. 2021. Т. 45. № 6. С. 917–925. <https://doi.org/10.18287/2412-6179-CO-1031>
18. Sarma S.S., Sinha K., Sub-r-pa C., Chakraborty G., Sinha B.P. Optimal distribution of traffic in Manhattan road networks for minimizing routing-time // *IEEE Transactions on Intelligent Transportation Systems*. 2021. Vol. 22. No. 11. P. 6799–6820. <https://doi.org/10.1109/TITS.2020.2994836>
19. Akopov A.S., Beklaryan L.A., Beklaryan A.L. Simulation-based optimisation for autonomous transportation systems using a parallel real-coded genetic algorithm with scalable nonuniform mutation // *Cybernetics and Information Technologies*. 2021. Vol. 21. No. 3. P. 127–144. <https://doi.org/10.2478/cait-2021-0034>
20. Sugiyama Y., Fukui M., Kikuchi M., Hasebe K., Nakayama A., Nishinari K., Tadaki S., Yukawa S. Traffic jams without bottlenecks – experimental evidence for the physical mechanism of the formation of a jam // *New Journal of Physics*. 2008. Vol. 10. No. 3. 033001. <https://doi.org/10.1088/1367-2630/10/3/033001>

21. Akopov A.S., Beklaryan L.A., Thakur M. Improvement of maneuverability within a multiagent fuzzy transportation system with the use of parallel biobjective real-coded genetic algorithm // *IEEE Transactions on Intelligent Transportation Systems*. 2022. Vol. 23. No. 8. P. 12648–12664. <https://doi.org/10.1109/TITS.2021.3115827>
22. Акопов А.С., Хачатрян Н.К., Бекларян Л.А., Бекларян А.Л. Система управления беспилотными транспортными средствами на основе нечеткой кластеризации. Часть 2. Нечеткая кластеризация и программная реализация // *Вестник компьютерных и информационных технологий*. 2020. Т. 17. № 10(196). С. 21–29. <https://doi.org/10.14489/vkit.2020.10.pp.021-029>
23. Акопов А.С., Бекларян Л.А. Моделирование динамики дорожно-транспортных происшествий с участием беспилотных автомобилей в транспортной системе «умного города» // *Бизнес-информатика*. 2022. Т. 16. № 4. С. 19–35. <https://doi.org/10.17323/2587-814X.2022.4.19.35>
24. Herrera F., Lozano M., Verdegay J.L. Tackling real-coded genetic algorithms: Operators and tools for behavioural analysis // *Artificial Intelligence Review*. 1998. Vol. 12. No. 4. P. 265–319. <https://doi.org/10.1023/A:1006504901164>
25. Akopov A.S., Beklaryan L.A., Thakur M., Verma B.D. Parallel multi-agent real-coded genetic algorithm for large-scale black-box single-objective optimisation // *Knowledge-Based Systems*. 2019. Vol. 174. P. 103–122. <https://doi.org/10.1016/j.knsys.2019.03.003>
26. Bezdek J.C. Cluster validity with fuzzy sets // *Journal of Cybernetics*. 1974. Vol. 3. No. 3. P. 58–73. <https://doi.org/10.1080/01969727308546047>
27. Beklaryan A.L., Akopov A.S. Simulation of agent-rescuer behaviour in emergency based on modified fuzzy clustering // *Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems (AAMAS '16)*. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 2016. P. 1275–1276.
28. Kuklová J. Highway modeling in anylogic for multi-agent approach to smart city management // *Proceedings of the 2021 Smart City Symposium Prague (SCSP)*. Prague, Czech Republic, 2021. P. 1–6, <https://doi.org/10.1109/SCSP52043.2021.9447402>
29. Оо Т.М., Lupin S., Khaing M.T., Thu A. Using hybrid modeling for estimation of the efficiency of service systems // *Proceedings of the 2021 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus)*, St. Petersburg, Moscow, Russia, 2021. P. 2277–2280. <https://doi.org/10.1109/ElConRus51938.2021.9396596>
30. Metropolis N. The beginning of the Monte Carlo method // *Los Alamos Science Special Issue*. 1987. Vol. 15. P. 125–130.

Об авторах

Бекларян Армен Левонович

кандидат технических наук;

доцент, департамент бизнес-информатики, Национальный исследовательский университет «Высшая школа экономики», 101000, г. Москва, ул. Мясницкая, д. 20;

старший научный сотрудник, лаборатория социального моделирования, Центральный экономико-математический институт, Российская академия наук, 117418, г. Москва, Нахимовский проспект, д. 47;

E-mail: abeklaryan@hse.ru

ORCID: 0000-0001-5141-5140

Бекларян Левон Андреевич

доктор физико-математических наук, профессор;

главный научный сотрудник, лаборатория динамических моделей экономики и оптимизации, Центральный экономико-математический институт, Российская академия наук, 117418, г. Москва, Нахимовский проспект, д. 47;

E-mail: beklar@cemi.rssi.ru

ORCID: 0000-0001-5695-4450

Акопов Андраник Сумбатович

доктор технических наук, профессор; профессор Российской Академии Наук;

главный научный сотрудник, лаборатория динамических моделей экономики и оптимизации, Центральный экономико-математический институт, Российская академия наук, 117418, г. Москва, Нахимовский проспект, д. 47;

E-mail: akopovas@umail.ru

ORCID: 0000-0003-0627-3037

Simulation model of an intelligent transportation system for the “smart city” with adaptive control of traffic lights based on fuzzy clustering

Armen L. Beklaryan^{a, b}

E-mail: abeklaryan@hse.ru

Levon A. Beklaryan^b

E-mail: beklar@cemi.rssi.ru

Andranik S. Akopov^b

E-mail: akopovas@umail.ru

^a HSE University
Address: 20, Myasnitskaya Street, Moscow 101000, Russia

^b Central Economics and Mathematics Institute, Russian Academy of Sciences
Address: 47, Nakhimovsky Prospect, Moscow 117418, Russia

Abstract

This article presents a new simulation model of an intelligent transportation system (ITS) for the “smart city” with adaptive traffic light control. The proposed transportation model, implemented in the AnyLogic, allows us to study the behavior of interacting agents: vehicles (V) and pedestrians (P) within the framework of a multi-agent ITS of the “Manhattan Lattice” type. The spatial dynamics of agents in such an ITS is described using the systems of finite-difference equations with the variable structure, considering the controlling impact of the “smart traffic lights.” Various methods of traffic light control aimed at maximizing the total traffic of the ITS output flow have been studied, in particular, by forming the required duration phases with the use of a genetic optimization algorithm, with a local (“weakly adaptive”) switching control and based on the proposed fuzzy clustering algorithm. The possibilities of optimizing the characteristics of systems for individual control of the behavior of traffic lights under various scenarios, in particular, for the ITS with spatially homogeneous and periodic characteristics, are investigated. To determine the best values of individual parameters of traffic light control systems, such as the phases’ durations, the radius of observation of traffic and pedestrian flows, threshold coefficients, the number of clusters, etc., the previously proposed parallel real-coded genetic optimization algorithm (RCGA type) is used. The proposed method of adaptive control of traffic lights based on fuzzy clustering demonstrates greater efficiency in comparison with the known methods of collective impact and local (“weakly adaptive”) control. The results of the work can be considered a component of the decision-making system in the management of urban services.

Keywords: intelligent transportation system, “smart city”, “smart traffic lights”, agent-based modeling, adaptive control, fuzzy clustering, AnyLogic

Citation: Beklaryan A.L., Beklaryan L.A., Akopov A.S. (2023) Simulation model of an intelligent transportation system for the “smart city” with adaptive control of traffic lights based on fuzzy clustering. *Business Informatics*, vol. 17, no. 3, pp. 70–86. DOI: 10.17323/2587-814X.2023.3.70.86

References

1. Li L., Wen D., Yao D. (2014) A survey of traffic control with vehicular communications. *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 1, pp. 425–432. <https://doi.org/10.1109/TITS.2013.2277737>
2. Tang C., Wei X., Hao M., Zhu C., Wang R., Chen W. (2018) Traffic signal phase scheduling based on device-to-device communication. *IEEE Access*, vol. 6, pp. 47636–47645. <https://doi.org/10.1109/ACCESS.2018.2867553>
3. Tong W., Pan Z., Liu K., Yali Y., Xiumin W., Huawei H., Wu D.O. (2020) Multi-agent deep reinforcement learning for urban traffic light control in vehicular networks. *IEEE Transactions on Vehicular Technology*, vol. 69, no. 8, pp. 8243–8256. <https://doi.org/10.1109/TVT.2020.2997896>
4. Lin W-H., Wang C. (2004) An enhanced 0-1 mixed-integer LP formulation for traffic signal control. *IEEE Transactions on Intelligent Transportation Systems*, vol. 5, no. 4, pp. 238–245. <https://doi.org/10.1109/TITS.2004.838217>
5. Gartner N.H., John D.C., Gabbay L.H. (1975) Optimization of traffic signal settings by mixed-integer linear programming. *Transportation Science*, vol. 9, no. 4, pp. 321–343. <https://doi.org/10.1287/trsc.9.4.321>
6. Davydov I., Tolstykh D., Kononova P., Legkih I. (2019) Genetic based approach for Novosibirsk traffic light scheduling. Proceedings of the *15th International Asian School-Seminar Optimization Problems of Complex Systems (OPCS)*, Novosibirsk, Russia, pp. 31–36. <https://doi.org/10.1109/OPCS.2019.8880158>
7. Turky A.M., Ahmad M.S., Yusoff M.Z.M., Sabar N.R. (2009) Genetic algorithm application for traffic light control. *Information Systems: Modeling, Development, and Integration. UNISCON 2009. Lecture Notes in Business Information Processing*, vol. 20. (eds. J. Yang, A. Ginige, H.C. Mayr, R.D. Kutsche). Berlin: Springer, Heidelberg, pp. 115–120. https://doi.org/10.1007/978-3-642-01112-2_12
8. García-Nieto J., Alba E., Carolina Olivera A. (2012) Swarm intelligence for traffic light scheduling: Application to real urban areas. *Engineering Applications of Artificial Intelligence*, vol. 25, no. 2, pp. 274–283. <https://doi.org/10.1016/j.engappai.2011.04.011>
9. Royani T., Haddadnia J., Alipoor M. (2013) Control of traffic light in isolated intersections using fuzzy neural network and genetic algorithm. *International Journal of Computer and Electrical Engineering*, vol. 5, no. 1, pp. 142–146. <https://doi.org/10.7763/IJCEE.2013.V5.682>
10. Agramamel F., Sadik M., Hannani A.E., Moubarak Y. (2022) A traffic-aware street lighting system based on fuzzy logic controller. Proceedings of the 2022 *IEEE 18th International Colloquium on Signal Processing & Applications (CSPA)*, Selangor, Malaysia, pp. 132–137. <https://doi.org/10.1109/CSPA55076.2022.9782045>
11. Rashid H., Ashrafi M.J.F., Azizi M., Heydarinezhad M.R. (2015) Intelligent traffic light control based on clustering using vehicular ad-hoc networks. Proceedings of the 2015 *7th Conference on Information and Knowledge Technology (IKT)*, Urmia, Iran, pp. 1–6. <https://doi.org/10.1109/IKT.2015.7288801>
12. Sachan A., Kumar N. (2023) SDN control-enabled and time-quantum-driven max-pressure approach for intersection management in smart city. *IEEE Systems Journal*, vol. 17, no. 1, pp. 1694–1702. <https://doi.org/10.1109/JSYST.2022.3211933>
13. Marino F., Leccese F., Pizzuti S. (2017) Adaptive street lighting predictive control. *Energy Procedia*, vol. 111, pp. 790–799. <https://doi.org/10.1016/j.egypro.2017.03.241>
14. Makarov V.L., Bakhtizin A.R., Epstein J.M. (2022) *Agent-based modeling for a complex world*. Moscow: Scientific publications department, GAUGN. <https://doi.org/10.18254/978-5-604-5843-5-4>
15. Veremchuk N.S. (2022) Elements of simulation modeling in issues of traffic optimization. *Vestnik kibernetiki*, vol. 4, no. 48, pp. 23–28 (in Russian). <https://doi.org/10.34822/1999-7604-2022-4-23-28>
16. Agafonov A., Yumaganov A., Myasnikov V. (2023) Cooperative control for signalized intersections in intelligent connected vehicle environments. *Mathematics*, vol. 11, no. 6, 1540. <https://doi.org/10.3390/math11061540>
17. Myasnikov V.V., Agafonov A.A., Yumaganov A.S. A deterministic predictive traffic signal control model in intelligent transportation and geoinformation systems. *Computer Optics*, vol. 45, no. 6, pp. 917–925 (in Russian). <https://doi.org/10.18287/2412-6179-CO-1031>
18. Sarma S.S., Sinha K., Sub-r-pa C., Chakraborty G., Sinha B.P. (2021) Optimal distribution of traffic in Manhattan road networks for minimizing routing-time. *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 11, pp. 6799–6820. <https://doi.org/10.1109/TITS.2020.2994836>
19. Akopov A.S., Beklaryan L.A., Beklaryan A.L. Simulation-based optimisation for autonomous transportation systems using a parallel real-coded genetic algorithm with scalable nonuniform mutation. *Cybernetics and Information Technologies*, 2021, vol. 21, no. 3, pp. 127–144. <https://doi.org/10.2478/cait-2021-0034>
20. Sugiyama Y., Fukui M., Kikuchi M., Hasebe K., Nakayama A., Nishinari K., Tadaki S., Yukawa S. (2008) Traffic jams without bottlenecks – experimental evidence for the physical mechanism of the formation of a jam. *New Journal of Physics*, vol. 10, no. 3, 033001. <https://doi.org/10.1088/1367-2630/10/3/033001>
21. Akopov A.S., Beklaryan L.A., Thakur M. (2022) Improvement of maneuverability within a multiagent fuzzy transportation system with the use of parallel biobjective real-coded genetic algorithm. *IEEE Transactions on Intelligent Transportation Systems*. vol. 23, no. 8. pp. 12648–12664. <https://doi.org/10.1109/TITS.2021.3115827>

22. Akopov A.S., Khachatryan N.K., Beklaryan L.A., Beklaryan A.L. (2020) Unmanned vehicle control system based on fuzzy clustering. Part 2. Fuzzy clustering and software implementation. *Vestnik komp'uternykh i informatsionnykh tekhnologii*, vol. 17, no. 10(196), pp. 21–29. <https://doi.org/10.14489/vkit.2020.10.pp.021-029>
23. Akopov A.S., Beklaryan L.A. (2022) Simulation of rates of traffic accidents involving unmanned ground vehicles within a transportation system for the 'smart city'. *Business Informatics*, vol. 16, no. 4, pp. 19–35. <https://doi.org/10.17323/2587-814X.2022.4.19.35>
24. Herrera F., Lozano M., Verdegay J.L. (1998) Tackling real-coded genetic algorithms: Operators and tools for behavioural analysis. *Artificial Intelligence Review*, vol. 12, no. 4, pp. 265–319. <https://doi.org/10.1023/A:1006504901164>
25. Akopov A.S., Beklaryan L.A., Thakur M., Verma B.D. (2019) Parallel multi-agent real-coded genetic algorithm for large-scale black-box single-objective optimisation. *Knowledge-Based Systems*, vol. 174, pp. 103–122. <https://doi.org/10.1016/j.knosys.2019.03.003>
26. Bezdek J.C. (1974) Cluster validity with fuzzy sets. *Journal of Cybernetics*, vol. 3, no. 3, pp. 58–73. <https://doi.org/10.1080/01969727308546047>
27. Beklaryan A.L., Akopov A.S. (2016) Simulation of agent-rescuer behaviour in emergency based on modified fuzzy clustering. Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems (AAMAS'16). *International Foundation for Autonomous Agents and Multiagent Systems*, Richland, SC, pp. 1275–1276.
28. Kuklová J. (2021) Highway modeling in anylogic for multi-agent approach to smart city management. Proceedings of the 2021 Smart City Symposium Prague (SCSP), Prague, Czech Republic, pp. 1–6. <https://doi.org/10.1109/SCSP52043.2021.9447402>
29. Oo T.M., Lupin S., Khaing M.T., Thu A. (2021) Using hybrid modeling for estimation of the efficiency of service systems. Proceedings of the IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus), St. Petersburg, Moscow, Russia, pp. 2277–2280. <https://doi.org/10.1109/ElConRus51938.2021.9396596>
30. Metropolis N. (1987) The beginning of the Monte Carlo method. *Los Alamos Science Special Issue*, vol. 15, pp. 125–130.

About the authors

Armen L. Beklaryan

Cand. Sci. (Tech.);

Associate Professor, Department of Business Informatics, HSE University, 20, Myasnitskaya Street, Moscow 101000, Russia;

Senior Researcher, Laboratory of Social Modeling, Central Economics and Mathematics Institute, Russian Academy of Sciences, 47, Nakhimovsky Prospect, Moscow 117418, Russia;

Email: abeklaryan@hse.ru

ORCID: 0000-0001-5141-5140

Levon A. Beklaryan

Dr. Sci. (Phys.-Math.), Professor;

Chief Researcher, Laboratory of Dynamic Models of Economy and Optimization, Central Economics and Mathematics Institute, Russian Academy of Sciences, 47, Nakhimovsky Prospect, Moscow 117418, Russia;

E-mail: beklar@cemi.rssi.ru

ORCID: 0000-0001-5695-4450

Andranik S. Akopov

Dr. Sci. (Tech.), Professor; Professor of the Russian Academy of Sciences;

Chief Researcher, Laboratory of Dynamic Models of Economy and Optimization, Central Economics and Mathematics Institute, Russian Academy of Sciences, 47, Nakhimovsky Prospect, Moscow 117418, Russia;

E-mail: akopovas@umail.ru

ORCID: 0000-0003-0627-3037

DOI: 10.17323/2587-814X.2023.3.87.100

Прогнозирование финансовых временных рядов с использованием сингулярного спектрального анализа*

А.В. Зиненко

E-mail: anna-z@mail.ru

Сибирский федеральный университет

Адрес: Россия, 660041, г. Красноярск, Свободный проспект, д. 79

Аннотация

Финансовые временные ряды представляют собой объемные массивы информации по котировкам и объемам торгов акций, валют и других биржевых и внебиржевых инструментов. Анализ и прогнозирование таких рядов всегда представляли особый интерес как для исследователей-аналитиков, так и для инвесторов-практиков. Однако, финансовые временные ряды имеют свою специфику, не позволяющую найти единственно верный и работающий метод прогнозирования. В настоящее время алгоритмы машинного обучения позволяют анализировать большие объемы данных и производить тестирование полученных моделей. Современные технологии позволяют тестировать и применять сложные методы прогнозирования, требующие объемных вычислений. Они дают возможность развивать математическую базу прогнозирования, комбинировать различные подходы в одном методе. Примером такого современного подхода является метод сингулярного спектрального анализа (SSA), который сочетает в себе разложение временного ряда в сумму временных рядов, метод главных компонент и рекуррентное прогнозирование. Целью настоящей работы является анализ возможности применения SSA метода к финансовым временным рядам. Метод SSA был рассмотрен в сравнении с другими распространенными методами прогнозирования финансовых временных рядов: ARIMA, разложение Фурье и рекуррентная нейронная сеть. Для реализации методов был разработан программный алгоритм на языке Python. Также была осуществлена апробация метода на временных рядах котировок российских и американских акций, валют и криптовалют.

Ключевые слова: нестационарные временные ряды, прогнозирование, сингулярный спектральный анализ, метрики ошибок

Цитирование: Зиненко А.В. Прогнозирование финансовых временных рядов с использованием сингулярного спектрального анализа // Бизнес-информатика. 2023. Т. 17. № 3. С. 87–100. DOI: 10.17323/2587-814X.2023.3.87.100

* Статья опубликована при поддержке Программы НИУ ВШЭ «Университетское партнерство»

Введение

К финансовым временным рядам можно отнести наборы данных по рыночным ценам различных финансовых инструментов, таких как ценные бумаги, валюты, производные контракты и драгоценные металлы. Такого рода временные ряды имеют некоторые особенности, которые следует учитывать при применении к ним тех или иных методов прогнозирования. Базовые инвестиционные теории, такие как модели Марковитца и Шарпа или модель ценообразования опционов Блэка-Шоулза [1], предполагали, что рыночные временные ряды являются случайными и подчиняются закону нормального распределения. Однако параллельно проводимые исследования, к примеру, Бенуа Мандельбротом [1, 2] и Эдгаром Петерсом [3, 4] доказали, что ценовые временные ряды неслучайны и, соответственно, нестационарны, поэтому применение к их прогнозу методов, основанных на предположении о случайности процессов, не даст адекватного прогноза.

Можно выделить следующие особенности финансовых временных рядов.

1. Как было указано выше, временные ряды рыночных цен нестационарны. Это значит, что среднее и дисперсия непостоянны на протяжении исследуемого промежутка. Поэтому прежде, чем применять к таким временным рядам методы, которые хорошо работают на случайных процессах, следует привести их к стационарному виду. Это можно сделать, к примеру, взятием разниц, как предлагает модель ARIMA.
2. Финансовые временные ряды персистентны. Это означает, что последующие показатели сильно зависят от предыдущих. К этому близка характерная для хаоса чувствительность к начальным условиям. Определить, персистентен ли временной ряд, позволяет показатель Херста [3]. Если данный показатель колеблется в пределах от 0,5 до 1, то временной ряд является персистентным. Херст называл такого рода процессы «процессами с долговременной памятью». Расчет показателя Херста для много численных временных рядов подтверждает их персистентный характер [4, 5].
3. По рыночным временным рядам имеется в открытом доступе большой объем информации. Это делает возможным применять к ним методы машинного обучения и анализа данных, в том числе и глубокое обучение, которое требует больших объемов тренировочной и тестовой выборки. Аналитики и трейдеры с успехом используют алгоритмы торгов, основанные на искусственных нейронных сетях и других методах машинного обучения. Финансовые данные доступны в разных временных промежутках, от минутных до недельных. С одной стороны, если брать минутные данные за несколько лет, то можно получить очень большой объем для обучения нейросети. Но с другой стороны, такая выборка может быть сильно зашумленной [6].
4. Финансовые временные ряды недифференцируемы, но при этом непрерывны. Если посмотреть на кривую биржевых котировок, то можно увидеть, что она не является гладкой. Это означает невозможность провести к ней касательную и, следовательно, невозможность посчитать производную в любой точке. Примером недифференцируемой непрерывной функции является функция Вейерштрасса [1], которая напоминает нетрендовый биржевой ряд. Данную специфику следует учитывать при применении определенных методов, например, анализа Фурье, которые разлагает временной ряд в сумму тригонометрических функций. В случае применения анализа Фурье к недифференцируемым рядам следует сгладить исходный ряд, например, скользящим средним.

Все перечисленные особенности следует учитывать при выборе методов прогнозирования финансовых временных рядов. К примеру, долговременную память могут учитывать некоторые виды рекуррентных нейросетей, такие как LSTM нейросети, способные в процессе обучения выбирать из прошлых данных те, которые наиболее существенно влияют на прогнозируемые значения.

Для проверки адекватности прогноза и сравнения методов прогнозирования следует разделять временной ряд на тренировочную и тестовую составляющую. На тренировочной составляющей алгоритм обучается (например, подбираются параметры модели Arima, коэффициенты Фурье, коэффициенты регрессии и т.д.). Затем полученная на тренировочном временном ряду модель применяется для прогноза, и прогнозный временной ряд сравнивается с тестовым с использованием метрик [7].

Все внутренние метрики основаны на оценке разницы между фактическими и прогнозными значениями. Средняя абсолютная ошибка (MAE) показывает среднее арифметическое разниц между фактическими и прогнозными значениями, средняя квадратичная ошибка (MSE) – среднее

арифметическое квадратов разниц, а средняя абсолютная ошибка в процентах (MAPE) – среднее арифметическое относительных разниц. MAPE вычисляется по следующей формуле

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \cdot 100 \right|, \quad (1)$$

где n – число наблюдений в тестовой выборке;

y_i – фактическое значение показателя в тестовой выборке;

\hat{y}_i – предсказанное значение показателя.

В тестировании предложенного метода прогнозирования была использована метрика MAPE, так как она позволяет оценить отклонение прогнозных данных от фактических в относительном выражении, то есть показывает отклонение прогноза от факта в процентах.

1. Материалы и методы

В данной работе прогноз финансовых временных рядов был осуществлен с использованием сингулярного спектрального анализа (SSA) или метода гусеницы. Как и анализ Фурье, SSA раскладывает исходный временной ряд на сумму компонент. Однако, в анализе Фурье временные ряды – слагаемые представляют собой периодические функции разной частоты и амплитуды, тогда как SSA разделяет исходный ряд на трендовую, периодическую и шумовую составляющие [8, 9]. Таким образом, SSA учитывает такие особенности финансовых временных рядов, как нестационарность и недифференцируемость.

Метод гусеницы – это вариант SSA анализа, разработанный в СССР в конце 80-х годов, к которому отечественные ученые пришли независимо от зарубежных коллег. В настоящее время наиболее глубоко раскрывает данный метод работа Н.Э. Голяндиной «Метод «Гусеница» – SSA анализ временных рядов» [10]. Отметим, что и зарубежные авторы ссылаются на работу Голяндиной как на первоисточник [11, 12]. Другие работы, посвященные SSA анализу, также принадлежат российским авторам: Л.Н. Леонтьевой [13], Д.Л. Данилову [14]. Авторы работ и пособий выделяют две проблемы при реализации метода: выбор длины гусеницы и выбор главных компонент. Если вторую проблему можно решить, рассмотрев вклад каждой компоненты в общую дисперсию, то по выбору длины гусеницы рекомендации носят скорее эвристический характер [15].

Суть метода SSA состоит в том, что исходный временной ряд трансформируется в матрицу, затем матрица с помощью сингулярного разложения разделяется на компоненты (здесь используется метод главных компонент). Следующий шаг зависит от целей анализа: либо компоненты разделяются на трендовую, периодическую и шумовую составляющие, и по ним производится анализ и прогноз (шумовые компоненты убираются), либо выбираются главные компоненты, и по ним строится продолжение ряда на длину заданного при формировании исходной матрицы шага. В данной работе использован второй подход. Алгоритм метода SSA выглядит следующим образом.

Анализируемый временной ряд имеет длину n . Выбирается длина гусеницы L , $2 \leq L \leq n/2$ и строится матрица X , полученная сдвигом каждой следующей колонки на одно значение размерности $L \times (n - L + 1)$.

$$\begin{matrix} x_1 & x_2 & \cdots & \cdots & x_{n-L+1} \\ x_2 & x_3 & \cdots & \cdots & x_{n-L+2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ x_L & x_{L+1} & \cdots & \cdots & x_n \end{matrix} \quad (2)$$

Затем строится сингулярное разложение матрицы X :

$$X = U \sigma V^T, \quad (3)$$

где U – левый сингулярный вектор;

σ – диагональная матрица значений;

V^T – правый сингулярный вектор.

Полученные элементы векторов сортируются по убыванию собственных значений. Исходная матрица X разлагается на L элементарных матриц:

$$X_j = \sum_{j=1}^d \sigma_j U_j V_j^T, \quad j = \overline{1, L}, \quad (4)$$

где σ_j – j -е собственное значение;

$U_j V_j^T$ – соответствующие ему левый и правый сингулярный вектор;

d – ранг матрицы X .

Матрицы X_j имеют размерность $L \times (n - L + 1)$. Для того, чтобы привести данные матрицы к одномерному виду используется диагональное усреднение. Каждая строка матрицы сдвигается на $i = 1 \dots n - L + 1$ значение, затем рассчитываются средние значения по столбцу, которые и составляют расчетные значения j -й компоненты исходного временного ряда. Процедура диагонального усреднения показана формулой (5).

$$\begin{matrix}
 x_{11} & x_{12} & \dots & \dots & x_{1(n-L+1)} \\
 & x_{21} & x_{22} & \dots & \dots & x_{2(n-L+1)} \\
 & & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\
 & & & & & & & & & x_{L1} & x_{L2} & \dots & \dots & x_{L(n-L+1)} \\
 \bar{x}_1 & \bar{x}_2 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \bar{x}_n
 \end{matrix} \quad (5)$$

Таким образом, исходный временной ряд разлагается на сумму L рядов. Далее полученные ряды анализируются на предмет трендовости, периодичности или шума, либо среди них выделяются главные компоненты, то есть те составляющие исходного временного ряда, которые оказывают наиболее существенное влияние на его динамику. Проанализировать характер компонент можно, к примеру, с помощью тепловых карт, однако задачей данного исследования является не анализ, а прогноз, поэтому мы, как это было указано выше, использовали только метод главных компонент. Вклад каждой полученной компоненты в общую дисперсию считается по формуле

$$\frac{\sigma_j^2}{\sum_{j=1}^d \sigma_j^2}, \quad (6)$$

где σ_j^2 – квадрат j -го собственного значения.

На рисунке 1 показаны дисперсии 10 компонент. Как видно из рисунка, наиболее существенный вклад в общую дисперсию дает первое собственное значение. Напомним, что собственные значения и соответствующие им левые и правые сингулярные векторы сортируются по убыванию собственных значений. Таким образом, на примере рисунка 1 можно выбрать только ряд X_1 для прогноза.

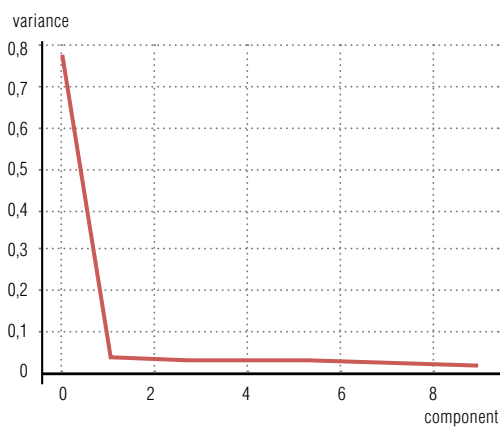


Рис. 1. Вклад дисперсий компонент.

Последним шагом анализа является непосредственно прогноз по выделенным главным компонентам. Отметим, что если компонент, оказывающий существенное влияние на дисперсию, больше одной, то следует их просуммировать. В работе использовался метод рекуррентного прогнозирования. Для этого были взяты последние $2L + 1$ элементов полученного временного ряда и была построена следующая система линейных уравнений:

$$\begin{aligned}
 a_1 X_{n-2L+1} + a_2 X_{n-2L+2} + \dots + a_L X_{n-L-1} &= X_{n-L} \\
 a_1 X_{n-2L+2} + a_2 X_{n-2L+3} + \dots + a_L X_{n-L} &= X_{n-L+1} \\
 \dots & \\
 a_1 X_{n-L+1} + a_2 X_{n-L+2} + \dots + a_L X_{n-1} &= X_n.
 \end{aligned} \quad (7)$$

Решая эту систему относительно коэффициентов a , получаем матрицу коэффициентов, которую подставляем в матрицу, составленную аналогичным с формулой (7) способом из последних L значений временного ряда и получаем прогнозные L значений.

$$\begin{aligned}
 a_1 X_{n-L+1} + a_2 X_{n-L+2} + \dots + a_L X_n &= X_{n+1} \\
 a_1 X_{n-L+2} + a_2 X_{n-L+3} + \dots + a_L X_{n+1} &= X_{n+2} \\
 \dots & \\
 a_1 X_{L+1} + a_2 X_{L+2} + \dots + a_L X_{n+L-1} &= X_{n+L}.
 \end{aligned} \quad (8)$$

Также для сравнения результатов, полученных с использованием SSA, в работе использовались прогнозирование другими методами: модель ARIMA (модель авторегрессионного скользящего среднего), разложение Фурье и рекуррентная нейронная сеть. Модель ARIMA была выбрана как самая распространенная для прогнозирования финансовых временных рядов. Разложение Фурье было выбрано, так как использует несколько схожий с SSA анализом принцип: исходный временной ряд также разлагается в сумму нескольких рядов (только в случае разложения Фурье все эти ряды – периодические). Что касается рекуррентной нейронной сети, то данный метод был выбран по причине перспективности и динамичного развития методов машинного и, в частности, глубокого обучения.

Модель ARIMA имеет три параметра – порядок авторегрессии p , порядок взятия разностей d и порядок скользящего среднего q . Порядок взятия разностей определяется тестом Дики-Фуллера. Порядок авторегрессии определяется графиком ав-

токорреляции уровней ряда, в котором по оси X называются временные лаги, а по оси Y — значения коэффициента корреляции между соответствующими лагу уровнями. Порядок авторегрессии выбирается равным такому временному лагу, при котором коэффициент корреляции принимает последнее максимальное отличное от нуля значение. Аналогичным образом выбирается порядок скользящего среднего, только вместо коэффициентов автокорреляции рассчитываются коэффициенты частичной автокорреляции. Частичная автокорреляция отличается от автокорреляции тем, что не учитывает влияние уровней, находящихся между текущим уровнем и уровнем, отстающим на временной лаг. Очевидно, что при единичном лаге автокорреляция и частичная автокорреляция совпадают.

Для осуществления прогноза по модели ARIMA на Python достаточно определить параметры p , d , q модели. Формально модель ARIMA описывается следующим образом

$$\Delta^d y_t = \sum_{i=1}^p \alpha_i \Delta^d y_{t-i} - \sum_{j=1}^q \beta_j \varepsilon_{t-j}, \quad (9)$$

где Δ^d — разность d -го порядка, необходимая для достижения стационарности;

α_i — коэффициенты авторегрессии p -го порядка;

β_j — коэффициенты скользящего среднего q -го порядка;

ε_{t-j} — ошибки предсказания скользящего среднего.

Коэффициенты α_i и β_j оцениваются и подставляются в прогноз.

Преобразование Фурье разлагает периодическую функцию на сумму синусоид и косинусоид с известными частотами, амплитудами и фазами. В общем виде разложение Фурье показано формулой (9).

$$y_t = \bar{y}_t + \sum_{i=1}^{N/2} (\alpha_i \cos \omega_i t + b_i \sin \omega_i t), \quad (10)$$

где y_t — преобразованное значение временного ряда;

\bar{y}_t — среднее значение исходного временного ряда;

ω_i — частота i -й гармоники (первая частота соответствует периоду функции, остальные кратны ей);

α_i, β_i — коэффициенты, которые необходимо определить.

Коэффициенты ряда Фурье вычисляются по следующим формулам:

$$a_1 = \frac{2}{N} \sum_{t=0}^{2\pi(N-1)/N} y_t \cos t, \quad b_1 = \frac{2}{N} \sum_{t=0}^{2\pi(N-1)/N} y_t \sin t \quad (11)$$

для первой гармоники,

$$a_2 = \frac{2}{N} \sum_{t=0}^{2\pi(N-1)/N} y_t \cos 2t, \quad b_2 = \frac{2}{N} \sum_{t=0}^{2\pi(N-1)/N} y_t \sin 2t \quad (12)$$

для второй гармоники и так далее. Обычно для прогноза достаточно первых двух–трех гармоник. Коэффициенты Фурье, полученные по формулам (11) и (12), подставляются в уравнение (10), и по данному уравнению осуществляется прогноз. В данной работе для прогнозирования методом разложения Фурье был использован код на языке программирования Python, разработанный автором.

Алгоритм прогнозирования временных рядов с использованием **рекуррентной нейросети** можно представить следующим образом.

1. Исходный временной ряд длиной n преобразовывается в матрицу, в которой строки соответствуют шагу L для построения прогноза. Массив размерностью $(L \times (n - L + 1))$ подается на вход в рекуррентный слой. Также задается m — шаг, на сколько периодов вперед строится прогноз.
2. Внутри рекуррентного слоя поданный массив обрабатывается активационной функцией, на выход подается массив размерности, заданной пользователем. Если данный слой последний, то размерность выходного массива равна $(m \times 1)$.
3. Готовое решение сравнивается с фактическими данными. Задается функция потерь (разница между фактическими и предсказанными нейросетью данными), и с помощью оптимизационной функции реализуется алгоритм обратного распространения ошибки. Алгоритм обратного распространения ошибки изменяет заданные случайным образом на этапе прямой работы нейросети веса таким образом, чтобы минимизировать функцию потерь. Большие массивы данных для «подгонки» весов разделяются на пакеты (батчи), то есть оптимизатор меняет веса не после подачи каждого сигнала, а после подачи пакета. Размер пакета задается при конструировании нейросети и обычно является степенью двойки. Число эпох определяет количество «прогонок» нейросети для успешного обучения. Качество обучения определяет функция потерь на тренировочных данных и метрика ошибки на тестовых.

2. Результаты и обсуждение

Для практического применения SSA был разработан алгоритм на языке Python с использованием библиотек numpy, pandas, matplotlib, sklearn. Данные по зарубежным финансовым временным рядам

были получены с сайта yahoo.finance при помощи библиотеки Python yfinance, а данные по котировкам отечественных акций – при помощи библиотеки Python arimoex.

В целях анализа были выбраны дневные котировки топовых российских и американских акций за период с июня 2022 по март 2023. Всего было проанализировано 30 компаний – 15 российских и 15 американских. Данные за 110 дней были разделены на тренировочную и тестовую выборку. Тренировочная выборка включала в себя 100 значений, тестовая – 10. Поскольку прогноз осуществлялся на период L – длину гусеницы, то, соответственно, параметр L был выбран равный 10. Таким образом, исходная матрица X имела размерность 10×91 , и в результате сингулярного разложения была разложена на 10 матриц такой же размерности. Данные матрицы были преобразованы в одномерные массивы размерностью 100, и из них были выбраны те, собственные значения матриц которых дают наибольший вклад в дисперсию. Затем полученные прогнозные значения сравнивались с тестовой выборкой с помощью метрики MAPE. В *таблице 1* показаны значения MAPE SSA-прогноза анализируемых акций.

В среднем ошибка прогноза составила 5,54%, в том числе по отечественным акциям – 4,62%, а по акциям американских компаний – 6,46%. Отметим,

что присутствует сильный выброс – по компании JPMorgan Chase & Co. ошибка оказалось значительной – 26%. По российским компаниям достаточно большая ошибка только по компании Русал – 11% (это вполне приемлемый для прогноза результат). По акциям компаний США существенная ошибка также по компании Mastercard – 17%. Также достаточно велики ошибки по компаниям United Health (9,5%) и Advanced Micro Devices (11%). Самые лучшие прогнозы были осуществлены по акциям Норильский никель (ошибка – 0,5%), МТС (ошибка – 1,7%) и McDonald's (ошибка – 1,2%). На *рисунке 2* показаны несколько графиков, иллюстрирующих SSA метод, при этом визуализированы как «плохие», так и «хорошие» прогнозы.

Для сравнения SSA анализа с методами ARIMA, разложение Фурье и рекуррентная нейронная сеть был осуществлен прогноз по тем же группам акций за тот же временной промежуток. Параметры p и d модели ARIMA для всех акций составили соответственно 2 и 1, тогда как параметр q менялся в зависимости от графика автокорреляции. Разложение Фурье осуществлялось по трем гармоникам. Рекуррентная нейронная сеть была сконструирована с использованием библиотек Keras и Tensor Flow Python. Поскольку для нейросети анализируемый промежуток весьма мал, во избежание переобучения был выбран один

Таблица 1.

MAPE метода SSA по акциям

Россия	MAPE	США	MAPE
Полиметал	7,00%	Amazon	5,00%
Полюс Золото	5,30%	Apple	4,70%
Мечел	7,80%	American Express Company	6,20%
ГМК «Норильский никель»	0,50%	Tesla	2,60%
Яндекс	5,20%	Advanced Micro Devices	11,00%
Аэрофлот	2,90%	Pfizer	2,10%
ВТБ	4,60%	Netflix	2,12%
Магнит	2,90%	Microsoft	2,00%
Алроса	5,20%	Mastercard	17,00%
Тинькофф Груп	6,10%	Visa	2,10%
Русал	11,00%	Starbucks	2,30%
Новатэк	2,40%	JPMorgan Chase & Co.	26,00%
Сургутнефтегаз	3,20%	McDonald's	1,20%
МТС	1,70%	Boeing	3,10%
Северсталь	3,50%	United Health	9,50%

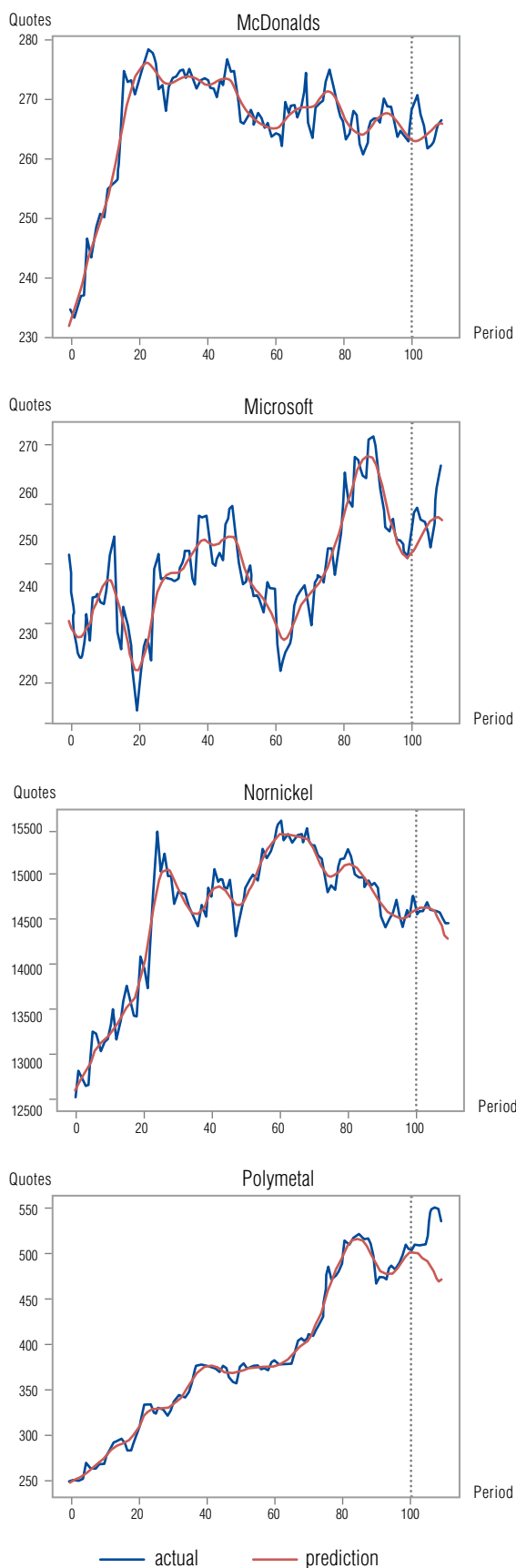


Рис. 2. Визуализация прогноза SSA методом.

Таблица 2.

**MAPE методов ARIMA,
разложение Фурье, RNN**

Россия	MAPE, ARIMA	MAPE, Фурье	MAPE, RNN
Полиметал	1,70%	27,60%	95,77%
Полус Золото	1,50%	17,00%	99,79%
Мечел	1,80%	22,20%	19,50%
ГМК «Норильский никель»	0,50%	1,50%	54,83%
Яндекс	1,20%	3,20%	95,51%
Аэрофлот	0,70%	8,80%	5,86%
ВТБ	0,80%	1,50%	3,60%
Магнит	0,40%	3,30%	99,05%
Алроса	0,70%	2,30%	29,19%
Тинькофф Груп	1,00%	3,90%	98,34%
Русал	0,70%	5,00%	2,65%
Новатэк	0,80%	2,60%	95,85%
Сургутнефтегаз	0,80%	3,50%	2,19%
МТС	1,00%	8,40%	81,83%
Северсталь	1,10%	18,40%	95,12%
Среднее по акциям РФ	0,98%	8,61%	58,61%
Amazon	1,50%	3,10%	56,39%
Apple	1,60%	6,40%	69,55%
American Express Company	2,30%	8,50%	70,46%
Tesla	3,10%	5,90%	76,90%
Advanced Micro Devices	2,00%	17,00%	34,24%
Pfizer	1,00%	17,10%	8,81%
Netflix	1,70%	3,00%	83,86%
Microsoft	1,30%	5,90%	81,80%
Mastercard	1,40%	3,80%	86,55%
Visa	1,30%	4,80%	78,09%
Starbucks	1,10%	3,50%	53,21%
JPMorgan Chase & Co.	2,70%	4,10%	64,73%
McDonald's	0,80%	2,00%	83,24%
Boeing	2,10%	10,80%	72,31%
United Health	0,70%	9,40%	91,70%
Среднее по акциям США	1,64%	7,18%	67,46%
Общее среднее	1,31%	7,90%	63,03%

RNN слой. Активационная функция, функция потерь и функция-оптимизатор были выбраны, исходя из рекомендованных в источнике [6] параметров для прогнозирующих рекуррентных нейронных сетей. Размер пакета был выбран минимальный – 2, а количество эпох было выбрано такое, при котором функция потерь и ошибка (средняя абсолютная ошибка и средняя абсолютная ошибка в процентах соответственно) перестали меняться. В большей части экспериментов использовалось 20 эпох. Прогноз, так же, как и SSA методом, осуществлялся на 10 дней вперед.

Как видно из *таблицы 2*, метод ARIMA по точности прогноза значительно превзошел прочие методы, включая SSA. Однако SSA анализ показал более точные результаты, чем разложение Фурье и рекуррентная нейронная сеть, а по акциям Норильский

Никель, МТС и McDonald's вполне сопоставим с ARIMA. Разложение Фурье в целом показало на акциях также неплохие результаты, выбросы ошибки присутствовали только по акциям Мечел, Advanced Micro Devices и Pfizer. Что касается рекуррентной нейронной сети, то на коротком промежутке по акциям она не работает, что подтвердили неприемлемые размеры ошибок.

Далее сингулярный спектральный анализ в сравнении с прочими методами был применен к валютному рынку и рынку криптовалют, которые являются более динамичными, чем рынок акций. Временной промежуток остался тем же – с июня 2022 по март 2023. Были взяты десять валютных пар и десять котировок криптовалют к доллару США. Средняя абсолютная ошибка в процентах по сравниваемым методам показана в *таблице 3*.

Таблица 3.

MAPE методов SSA, ARIMA, преобразования Фурье и RNN по валютам

Валюта	MAPE, SSA	MAPE, ARIMA	MAPE, Фурье	MAPE, RNN
Евро / Доллар США	59,00%	0,50%	2,70%	1,89%
Фунт / Доллар США	0,70%	0,60%	2,10%	2,22%
Доллар США / Юань	3,70%	0,50%	1,70%	1,47%
Доллар США / Рубль	37,90%	0,30%	12,20%	30,46%
Доллар США / Иена	40,20%	0,60%	1,80%	69,06%
Доллар США / Гонконгский доллар	43,60%	0,00%	1,80%	0,33
Доллар США / Южноафриканский Рэнд	8,30%	0,80%	5,00%	2,27%
Австралийский Доллар / Доллар США	0,90%	0,60%	1,90%	0,33%
Доллар США / Мексиканское Песо	1,70%	1,10%	5,20%	1,36%
Новозеландский Доллар / Доллар США	2,00%	0,70%	1,90%	2,42%
Среднее по фиатным валютам	16,38%	0,57%	3,63%	15,90%
Bitcoin	6,80%	3,10%	9,80%	99,90%
Ethereum	8,20%	3,10%	8,30%	98,75%
Binance Coin	47,3%	1,80%	4,00%	6,57%
Polygon	8,10%	4,00%	8,00%	6,25%
Lightcoin	12,30%	5,70%	8,40%	74,24%
Ripple	1,90%	1,90%	2,50%	6,29%
Polkadot	6,40%	3,50%	4,10%	5,70%
Chainlink	7,40%	3,30%	4,60%	9,44%
Avalanche	17,40%	4,70%	4,8%	11,13%
Tron	45,50%	4,00%	6,70%	4,52%
Среднее по криптовалютам	16,13%	3,51%	6,08%	34,82%
Итоговое среднее	16,26%	2,04%	4,85%	25,36%

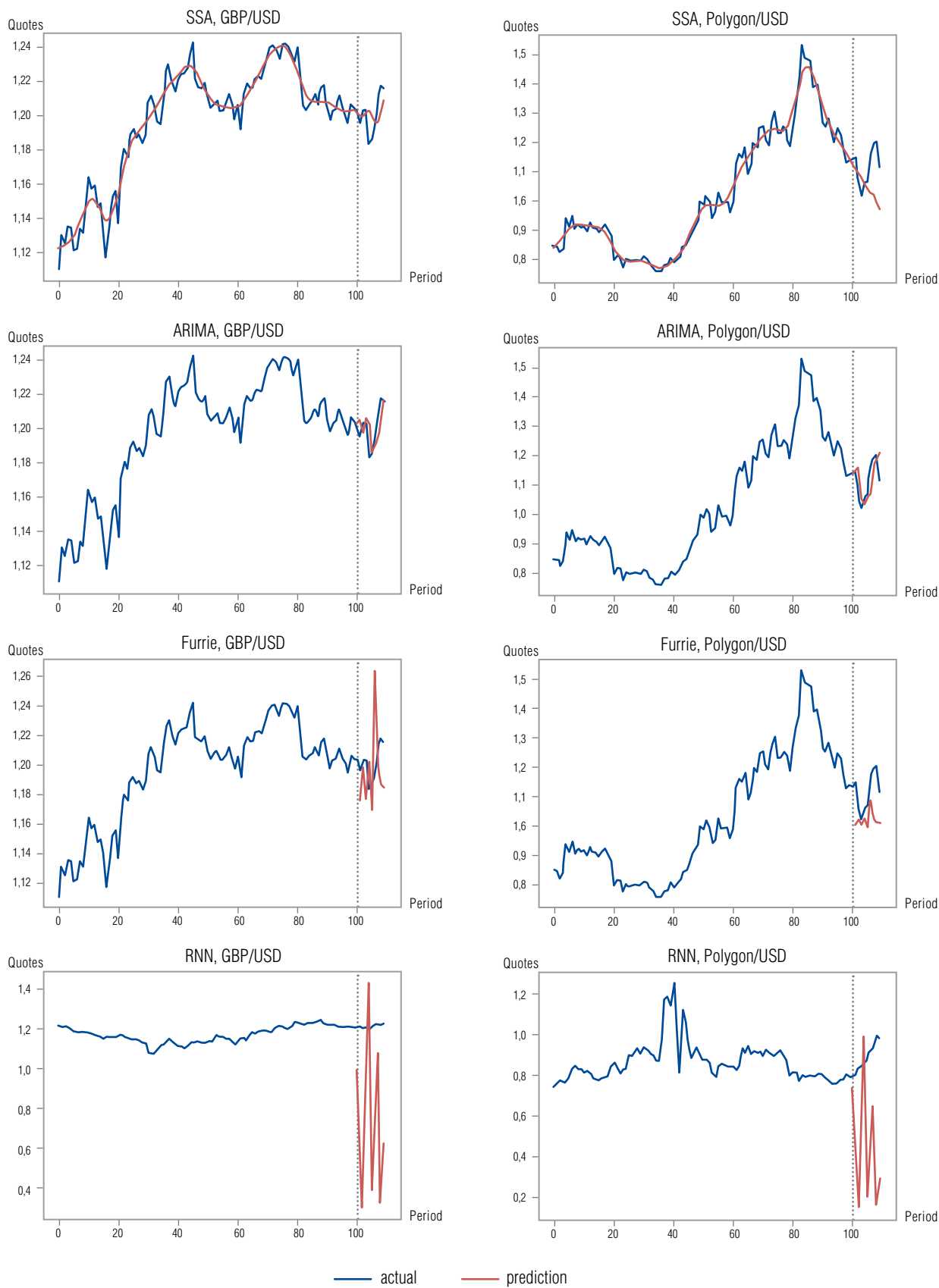


Рис. 3. Визуализация прогнозов по валютам методами SSA, ARIMA, разложение Фурье и RNN.

С валютами, как фиатными, там и крипто, метод SSA показал не очень хорошие результаты. Если по криптовалютам выброс произошел только по валютам BNB и Tron, то по фиатным валютным парам метод показал крупную ошибку в четырех случаях из десяти. Метод ARIMA на валютах работал так же хорошо, как и на акциях, а методы разложения Фурье и рекуррентная нейросеть существенно улучшили точность прогноза. При этом прогноз рекуррентной нейросети остался неудовлетворительным. Отметим, что все рассмотренные методы на криптовалютах работают хуже, чем на фиатных. Это может быть связано с тем, что в силу молодости крипторынка на нем все еще остаются лазейки для арбитражеров. На *рисунке 3* показаны графики прогнозов некоторых фиатных и криптовалют рассмотренными методами.

В качестве решения проблемы плохого прогноза SSA как на фиатных, так и на криптовалютах, мы увеличили временной промежуток вдвое. Новые тренировочная и тестовая выборки по тем же валютам составили 200 и 20 (длина гусеницы соответственно была так же выбрана 20). Средняя абсолютная ошибка в процентах при новых условиях показана в *таблице 4*.

При увеличении временного промежутка результаты прогноза методом SSA значительно улучшились, тогда как точность прогноза методом ARIMA практически не изменилась, а методом Фурье даже ухудшилась. Рекуррентная нейронная сеть существенно улучшила точность прогноза, однако по двум фиатным валютам и трем криптовалютам остались значительные выбросы. Лучшая точность осталась за моделью ARIMA. На *рисунке 4* показаны прогнозы по валютам на длинном временном промежутке.

Таблица 4.

MAPE методов SSA и рекуррентной нейросети по валютам (длинный промежуток)

Валюта	MAPE, SSA	MAPE, ARIMA	MAPE, Фурье	MAPE, RNN
Евро / Доллар США	2,60%	0,40%	3,00%	2,44%
Фунт / Доллар США	8,20%	0,50%	1,20%	2,62%
Доллар США / Юань	0,90%	0,40%	1,10%	2,40%
Доллар США / Рубль	1,50%	1,10%	16,30%	50,53%
Доллар США / Иена	6,50%	0,50%	1,50%	78,31%
Доллар США / Гонконгский доллар	1,00%	0,00%	0,90%	0,22%
Доллар США / Южноафриканский Рэнд	1,80%	0,60%	6,60%	2,94%
Австралийский Доллар / Доллар США	3,10%	0,50%	1,70%	0,38%
Доллар США / Мексиканское Песо	1,10%	0,70%	8,1%	1,53%
Новозеландский Доллар / Доллар США	3,20%	0,50%	1,10%	2,99%
Среднее по фиатным валютам	3,16%	0,52%	4,48%	14,44%
Bitcoin	9,90%	2,30%	13,30%	99,83%
Ethereum	48,80%	2,40%	9,90%	97,76%
Binance Coin	3,20%	1,50%	3,80%	7,22%
Polygon	16,50%	3,70%	20,50%	7,23%
Lightcoin	12,40%	3,90%	19,40%	50,07%
Ripple	4,40%	1,50%	6,10%	8,88%
Polkadot	19,70%	3,20%	5,80%	6,32%
Chainlink	10,80%	2,90%	5,20%	8,55%
Avalanche	7,40%	4,00%	6,2%	18,19%
Tron	7,00%	2,70%	9,80%	4,73%
Среднее по криптовалютам	14,01%	2,81%	10,02%	30,88%
Итоговое среднее	8,58%	1,67%	7,25%	22,66%

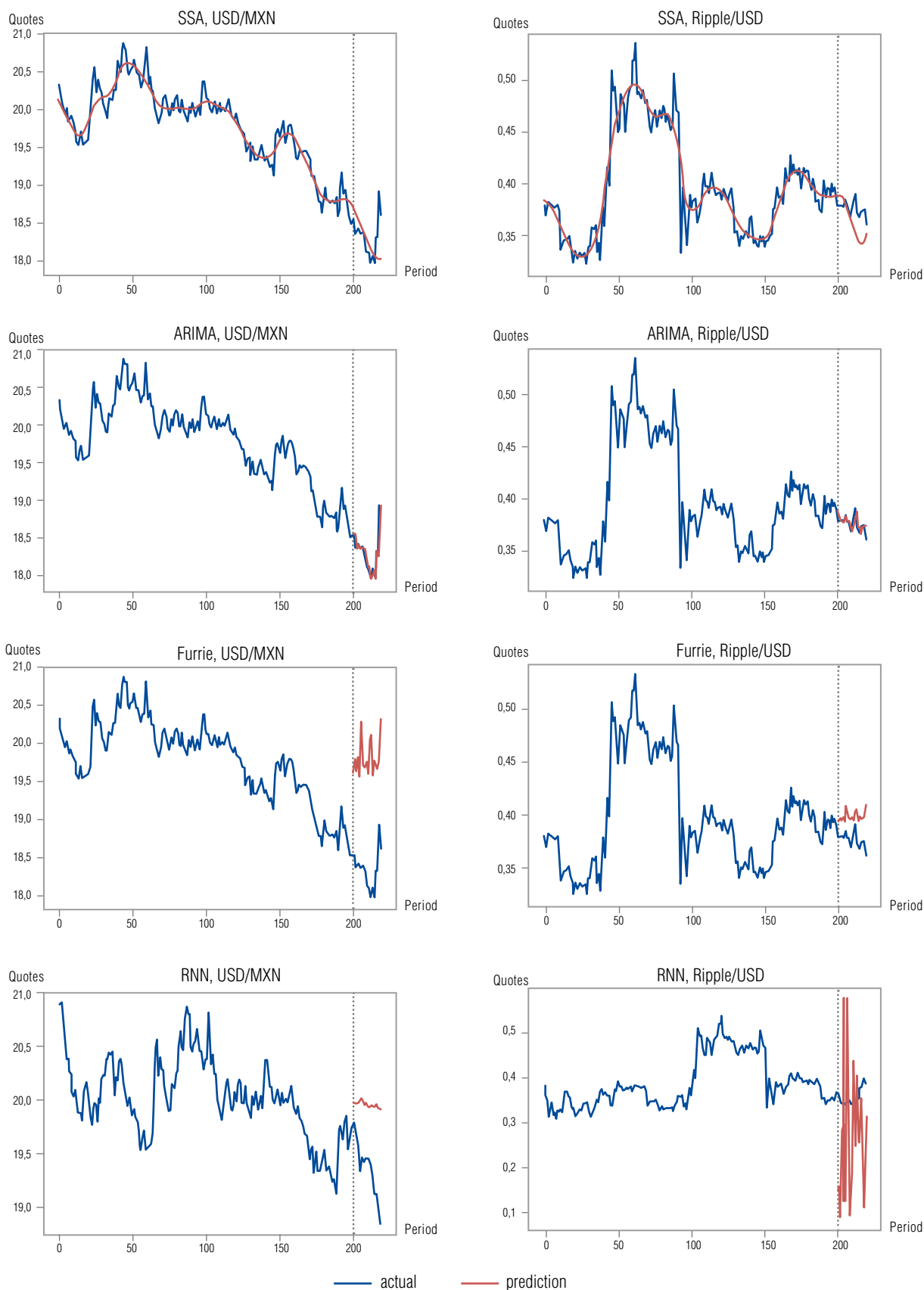


Рис. 4. Визуализация прогнозов по валютам методами SSA, ARIMA, разложение Фурье и RNN (длинный временной промежуток).

Заключение

В работе был рассмотрен метод прогнозирования временных рядов SSA анализ (метод гусеницы). Метод был реализован путем разработки алгоритма на языке Python, затем протестирован на 30 временных рядах котировок акций РФ и США, а также на 20 фиатных и криптовалютах по отношению к доллару США. Для сравнения были взяты три метода прогнозирования: ARIMA, разложение Фурье и рекуррентная нейронная сеть. Метод SSA анализа во всех случаях, кроме короткого временного промежутка по валютам, показал второй по точности после метода ARIMA результат. По некоторым бумагам и акциям ошибка SSA метода сравнима с ошибкой метода ARIMA. При этом увеличение временного промежутка значительно улучшило результаты применения SSA метода, тогда как результаты метода ARIMA остались неизменными.

Можно сделать вывод о том, что хоть метод SSA анализа показывает меньшую точность прогноза, чем общепринятый в анализе финансовых временных рядов метод ARIMA, он может быть применен как к акциям, так и к другим финансовым инструментам, даже к таким волатильным, как криптовалюты. Его можно использовать как подтверждение результатов метода ARIMA, а также отдельно как метод прогнозирования. Отметим, что для анализа большого массива акций метод SSA удобнее, чем ARIMA, так как ARIMA требует пересчета по крайней мере порядка скользящего среднего для каждого временного ряда, тогда как для прогноза методом SSA достаточно одной главной компоненты. Кроме того, если рассматривать другие компоненты сингулярного разложения, то можно сделать выводы о соотношении трендовости, периодичности и шума в анализируемом временном ряде, что не позволяет ни один из прочих рассмотренных методов. ■

Литература

1. Mandelbrot B., Fisher A., Calvet L. A multifractal model of asset returns // Yale Cowles Foundation for Research in Economics. 1997. Discussion Paper No. 1164.
2. Мандельброт Б. Фракталы, случай и финансы. Москва, Ижевск: НИЦ «Регулярная и хаотическая динамика», 2004.
3. Петерс Э. Фрактальный анализ финансовых рынков. Применение хаоса в инвестициях и экономике. М.: Интернет-трейдинг, 2004.
4. Петерс Э. Хаос и порядок на рынках капитала. Новый аналитический взгляд на циклы, цены и изменчивость рынка. М.: Мир, 2000.
5. Зиненко А.В. R/S анализ на фондовом рынке // Бизнес-информатика. 2012. № 3(21). С. 21–27.
6. Шоле Ф. Глубокое обучение на Python. СПб.: Питер, 2018.
7. Kong Q., Han J., Jin X., Li C., Wang T., Bai Q., Chen Y. Polar motion prediction using the combination of SSA and ARMA // Geodesy and Geodynamics. 2023. Vol. 14. No. 4. P. 368–376. <https://doi.org/10.1016/j.geog.2022.12.004>
8. Li K., Zhang Z., Guo H., Li W., Yan Y. Prediction method of pipe joint opening-closing deformation of immersed tunnel based on singular spectrum analysis and SSA-SVR // Applied Ocean Research. 2023. Vol. 135. 103526. <https://doi.org/10.1016/j.apor.2023.103526>
9. Montalvo C., Pantera L., Lipcsei S., Torres L.A. Signal processing applied in cortex project: From noise analysis to OMA and SSA methods. Annals of Nuclear Energy. 2022. Vol. 175. 109193. <https://doi.org/10.1016/j.anucene.2022.109193>
10. Голяндина Н.Э. Метод «Гусеница» – SSA: анализ временных рядов. СПб.: СПбГУ, 2004.
11. Coussin M. Singular spectrum analysis for real-time financial cycles measurement // Journal of International Money and Finance. 2022. Vol. 120. 102532. <https://doi.org/10.1016/j.jimonfin.2021.102532>
12. Lahmiri S., Bekiros S., Bezzina F. Evidence of the fractal market hypothesis in European industry sectors with the use of bootstrapped wavelet leaders singularity spectrum analysis // Chaos, Solitons & Fractals. 2022. Vol. 165. Part 1. 112813. <https://doi.org/10.1016/j.chaos.2022.112813>
13. Леонтьева Л.Н. Многомерная гусеница, выбор длины и числа компонент // Машинное обучение и анализ данных. 2011. № 1. С. 5–15.
14. Солнцев В.Н., Данилов Д.Л., Жиглявский А.А. Главные компоненты временных рядов: Метод «Гусеница». СПб.: СПбГУ, 1997.
15. Baharanchi S.A., Vali M., Modares M. Noise reduction of lung sounds based on singular spectrum analysis combined with discrete cosine transform // Applied Acoustics. 2022. Vol. 199. 109005. <https://doi.org/10.1016/j.apacoust.2022.109005>

Об авторе

Зиненко Анна Викторовна

к.т.н., доцент;

доцент, кафедра экономической и финансовой безопасности, Сибирский федеральный университет, 660041, г. Красноярск, Свободный проспект, д. 79;

E-mail: anna-z@mail.ru

Forecasting financial time series using singular spectrum analysis

Anna V. Zinenko

E-mail: anna-z@mail.ru

Siberian Federal University

Address: 79, Svobodny Prospect, Krasnoyarsk 660041, Russia

Abstract

Financial time series are big arrays of information on quotes and trading volumes of shares, currencies and other exchange and over-the-counter instruments. The analysis and forecasting of such series has always been of particular interest for both research analysts and practicing investors. However, financial time series have their own features, which do not allow one to choose the only correct and well-functioning forecasting method. Currently, machine-learning algorithms allow one to analyze large amounts of data and test the resulting models. Modern technologies enable testing and applying complex forecasting methods that require volumetric calculations. They make it possible to develop the mathematical basis of forecasting, to combine different approaches into a single method. An example of such a modern approach is the Singular Spectrum Analysis (SSA), which combines the decomposition of a time series into a sum of time series, principal component analysis and recurrent forecasting. The purpose of this work is to analyze the possibility of applying SSA to financial time series. The SSA method was considered in comparison with other common methods for forecasting financial time series: ARIMA, Fourier transform and recurrent neural network. To implement the methods, a software algorithm in the Python language was developed. The method was also tested on the time series of quotes of Russian and American stocks, currencies and cryptocurrencies.

Keywords: non-stationary time series, forecasting, singular spectrum analysis, error metrics

Citation: Zinenko A.V. (2023) Forecasting financial time series using singular spectrum analysis. *Business Informatics*, vol. 17, no. 3, pp. 87–100. DOI: 10.17323/2587-814X.2023.3.87.100

References

1. Mandelbrot B., Fisher A., Calvet L. (1997) A multifractal model of asset returns. *Yale Cowles Foundation for Research in Economics*, Discussion Paper No. 1164.
2. Mandelbrot B. (2004) *Fractals, case and finance*. Moscow, Izhevsk: Research Center “Regular and Chaotic Dynamics” (in Russian).
3. Peters E. (2004) *Fractal analysis of financial markets. Application of chaos in investment and economics*. Moscow: Internet Trading (in Russian).
4. Peters E. (2000) *Chaos and order in capital markets. A new analytical perspective on cycles, prices and market volatility*. Moscow: Mir (in Russian).
5. Zinenko A.V. (2012) R/S analysis in the stock market. *Business Informatics*, no. 3(21), pp. 21–27 (in Russian).
6. Cholet F. (2018) *Deep learning with Python*. Saint Petersburg: Peter (in Russian).
7. Kong Q., Han J., Jin X., Li C., Wang T., Bai Q., Chen Y. (2023) Polar motion prediction using the combination of SSA and ARMA. *Geodesy and Geodynamics*, vol. 14, no. 4, pp. 368–376. <https://doi.org/10.1016/j.geog.2022.12.004>
8. Li K., Zhang Z., Guo H., Li W., Yan Y. (2023) Prediction method of pipe joint opening-closing deformation of immersed tunnel based on singular spectrum analysis and SSA-SVR. *Applied Ocean Research*, vol. 135, 103526. <https://doi.org/10.1016/j.apor.2023.103526>
9. Montalvo C., Pantera L., Lipcsei S., Torres L.A. (2022) Signal processing applied in cortex project: From noise analysis to OMA and SSA methods. *Annals of Nuclear Energy*, vol. 175, 109193. <https://doi.org/10.1016/j.anucene.2022.109193>
10. Golyandina N.E. (2004) *Method “Caterpillar”-SSA: analysis of time series*. Saint Petersburg: St. Petersburg State University (in Russian).
11. Coussin M. (2022) Singular spectrum analysis for real-time financial cycles measurement. *Journal of International Money and Finance*, vol. 120, 102532. <https://doi.org/10.1016/j.jimonfin.2021.102532>
12. Lahmiri S., Bekiros S., Bezzina F. (2022) Evidence of the fractal market hypothesis in European industry sectors with the use of bootstrapped wavelet leaders singularity spectrum analysis. *Chaos, Solitons & Fractals*, vol. 165, part 1, 112813. <https://doi.org/10.1016/j.chaos.2022.112813>
13. Leontyeva L.N. (2011) Multidimensional caterpillar, choice of length and number of components. *Machine Learning and Data Analysis*, no. 1, pp. 5–15 (in Russian).
14. Solntsev V.N., Danilov D.L., Zhiglyavsky A.A. (1997) *Principal components of time series: Method “Caterpillar”*. Saint Petersburg: St. Petersburg State University (in Russian).
15. Baharanchi S.A., Vali M., Modares M. (2022) Noise reduction of lung sounds based on singular spectrum analysis combined with discrete cosine transform. *Applied Acoustics*, vol. 199, 109005. <https://doi.org/10.1016/j.apacoust.2022.109005>

About the author

Anna V. Zinenko

Cand. Sci. (Tech.), Associate Professor;

Associate Professor, Department of Economic and Financial Security, Siberian Federal University, 79, Svobodny Prospect, Krasnoyarsk 660041, Russia;

E-mail: anna-z@mail.ru