

ISSN 2713-3192
DOI 10.15622/ia.2021.20.5
<http://ia.spcras.ru>

ТОМ 20 № 5

ИНФОРМАТИКА И АВТОМАТИЗАЦИЯ

INFORMATICS AND AUTOMATION



СПб ФИЦ РАН

Санкт-Петербург
2021

INFORMATICS AND AUTOMATION

Volume 20 № 5, 2021

Scientific and educational journal primarily specialized in computer science, automation, robotics, applied mathematics, interdisciplinary research

Founded in 2002

Founder and Publisher

St. Petersburg Federal Research Center of the Russian Academy of Sciences (SPC RAS)

Editor-in-Chief

R. M. Yusupov, Prof., Dr. Sci., Corr. Member of RAS, St. Petersburg, Russia

Editorial Council

A. A. Ashimov	Prof., Dr. Sci., Academician of the National Academy of Sciences of the Republic of Kazakhstan, Almaty, Kazakhstan
N. P. Veselkin	Prof., Dr. Sci., Academician of RAS, St. Petersburg, Russia
I. A. Kalyaev	Prof., Dr. Sci., Academician of RAS, Taganrog, Russia
Yu. A. Merkur'yev	Prof., Dr. Sci., Academician of the Latvian Academy of Sciences, Riga, Latvia
A. I. Rudskoi	Prof., Dr. Sci., Academician of RAS, St. Petersburg, Russia
V. Sgurev	Prof., Dr. Sci., Academician of the Bulgarian Academy of Sciences, Sofia, Bulgaria
B. Ya. Sovetov	Prof., Dr. Sci., Academician of RAE, St. Petersburg, Russia
V. A. Soyfer	Prof., Dr. Sci., Academician of RAS, Samara, Russia

Editorial Board

O. Yu. Gusikhin	Ph. D., Dearborn, USA
V. Delic	Prof., Dr. Sci., Novi Sad, Serbia
A. Dolgui	Prof., Dr. Sci., St. Etienne, France
M. N. Favorskaya	Prof., Dr. Sci., Krasnoyarsk, Russia
M. Zelezny	Assoc. Prof., Ph.D., Plzen, Czech Republic
H. Kaya	Assoc. Prof., Ph.D., Utrecht, Netherlands
A. A. Karpov	Assoc. Prof., Dr. Sci., St. Petersburg, Russia
S. V. Kuleshov	Dr. Sci., St. Petersburg, Russia
A. D. Khomonenko	Prof., Dr. Sci., St. Petersburg, Russia
D. A. Ivanov	Prof., Dr. Habil., Berlin, Germany
K. P. Markov	Assoc. Prof., Ph.D., Aizu, Japan
R. V. Meshcheryakov	Prof., Dr. Sci., Moscow, Russia
N. A. Moldovian	Prof., Dr. Sci., St. Petersburg, Russia
V. Yu. Osipov	Prof., Dr. Sci., St. Petersburg, Russia
V. K. Pshikhopov	Prof., Dr. Sci., Taganrog, Russia
A. L. Ronzhin	Prof., Dr. Sci., Deputy Editor-in-Chief, St. Petersburg, Russia
H. Samani	Assoc. Prof., Ph.D., Plymouth, UK
V. Skormin	Prof., Ph.D., Binghamton, USA
A. V. Smirnov	Prof., Dr. Sci., St. Petersburg, Russia
B. V. Sokolov	Prof., Dr. Sci., St. Petersburg, Russia

Editor: A.S. Lopotova

Interpreter: E.N. Mesheryakova

Art editor: N.A. Dormidontova

Editorial office address

14-th line V.O., 39, SPIIRAS, St. Petersburg, 199178, Russia,

e-mail: ia@spcras.ru, web: <http://ia.spcras.ru>

The journal is indexed in Scopus

The journal is published under the scientific-methodological supervision of Department for Nanotechnology and Information Technology of the Russian Academy of Sciences

© St. Petersburg Federal Research Center of the Russian Academy of Sciences, 2021

ИНФОРМАТИКА И АВТОМАТИЗАЦИЯ

Том 20 № 5, 2021

Научный, научно-образовательный журнал с базовой специализацией
в области информатики, автоматизации, робототехники, прикладной математики
и междисциплинарных исследований.

Журнал основан в 2002 году

Учредитель и издатель

Федеральное государственное бюджетное учреждение науки
«Санкт-Петербургский Федеральный исследовательский центр Российской академии наук»
(СПб ФИЦ РАН)

Главный редактор

Р. М. Юсупов, чл.-корр. РАН, д-р техн. наук, проф., Санкт-Петербург, РФ

Редакционный совет

А. А. Ашимов	академик Национальной академии наук Республики Казахстан, д-р техн. наук, проф., Алматы, Казахстан
Н. П. Веселкин	академик РАН, д-р мед. наук, проф., Санкт-Петербург, РФ
И. А. Каляев	академик РАН, д-р техн. наук, проф., Таганрог, РФ
Ю. А. Меркурьев	академик Латвийской академии наук, д-р, проф., Рига, Латвия
А. И. Рудской	академик РАН, д-р техн. наук, проф., Санкт-Петербург, РФ
В. Сгурев	академик Болгарской академии наук, д-р техн. наук, проф., София, Болгария
Б. Я. Советов	академик РАН, д-р техн. наук, проф., Санкт-Петербург, РФ
В. А. Сойфер	академик РАН, д-р техн. наук, проф., Самара, РФ

Редакционная коллегия

О. Ю. Гусихин	д-р наук, Диаборн, США
В. Делич	д-р техн. наук, проф., Нови-Сад, Сербия
А. Б. Долгий	д-р наук, проф. Сент-Этьен, Франция
М. Железны	д-р наук, доцент, Пльзень, Чешская республика
Д. А. Иванов	д-р экон. наук, проф., Берлин, Германия
Х. Кайя	д-р наук, доцент, Утрехт, Нидерланды
А. А. Карпов	д-р техн. наук, доцент, Санкт-Петербург, РФ
С. В. Кулешов	д-р техн. наук, Санкт-Петербург, РФ
К. П. Марков	д-р наук, доцент, Аизу, Япония
Р. В. Мещеряков	д-р техн. наук, проф., Москва, РФ
Н. А. Молдовян	д-р техн. наук, проф., Санкт-Петербург, РФ
В.Ю. Осипов	д-р техн. наук, проф., Санкт-Петербург, РФ
В. Х. Пшихолов	д-р техн. наук, проф., Таганрог, РФ
А. Л. Ронжин	д-р техн. наук, проф., зам. главного редактора, Санкт-Петербург, РФ
Х. Самани	д-р наук, доцент, Плимут, Соединённое Королевство
В. А. Скормин	д-р наук, проф., Бингемптон, США
А. В. Смирнов	д-р техн. наук, проф., Санкт-Петербург, РФ
Б. В. Соколов	д-р техн. наук, проф., Санкт-Петербург, РФ
Л. В. Уткин	д-р техн. наук, проф., Санкт-Петербург, РФ
М. Н. Фаворская	д-р техн. наук, проф., Красноярск, РФ
А. Д. Хомоненко	д-р техн. наук, проф., Санкт-Петербург, РФ
Л. Б. Шереметов	д-р техн. наук, Мехико, Мексика

Выпускающий редактор: А.С. Лопотова

Переводчик: Е.Н. Мещерякова

Художественный редактор: Н.А. Дормидонтова

Адрес редакции

199178, г. Санкт-Петербург, 14-я линия В.О., д. 39

e-mail: ia@spcras.ru, сайт: <http://ia.spcras.ru>

Журнал индексируется в международной базе данных Scopus

Журнал входит в «Перечень ведущих рецензируемых научных журналов и изданий,
в которых должны быть опубликованы основные научные результаты диссертации
на соискание ученой степени доктора и кандидата наук»

Журнал выпускается при научно-методическом руководстве Отделения нанотехнологий
и информационных технологий Российской академии наук

© Федеральное государственное бюджетное учреждение науки

«Санкт-Петербургский Федеральный исследовательский центр Российской академии наук», 2021
Разрешается воспроизведение в прессе, а также сообщение в эфир или по кабелю опубликованных
в составе печатного периодического издания - журнала «ИНФОРМАТИКА И АВТОМАТИЗАЦИЯ»
статей по текущим экономическим, политическим, социальным и религиозным вопросам
с обязательным указанием имени автора статьи и печатного периодического издания
журнала «ИНФОРМАТИКА И АВТОМАТИЗАЦИЯ»

CONTENTS

Preface	1005
Mathematical Modeling, Numerical Methods	
Y. Popkov, Y. Dubnov, A. Popkov FORECASTING DEVELOPMENT OF COVID-19 EPIDEMIC IN EUROPEAN UNION USING ENTROPY-RANDOMIZED APPROACH	1010
V. Zakharov, Y. Balykina BALANCE MODEL OF COVID-19 EPIDEMIC BASED ON PERCENTAGE GROWTH RATE	1034
V. Osipov, S. Kuleshov, A. Zaytseva, A. Aksenov APPROACH FOR THE COVID-19 EPIDEMIC SOURCE LOCALIZATION IN RUSSIA BASED ON MATHEMATICAL MODELING	1065
Artificial Intelligence, Knowledge and Data Engineering	
A. Smirnov, E. Moll, N. Teslya USE OF FUZZY COALITION GAMES IN SOCIALLY ORIENTED DECISION MAKING DURING HOSPITALIZATION IN PANDEMIC	1090
A. Dvoynikova, M. Markitantov, E. Ryumina, D. Ryumin, A. Karpov ANALYTICAL REVIEW OF AUDIOVISUAL SYSTEMS FOR DETERMINING PERSONAL PROTECTIVE EQUIPMENT ON A PERSON'S FACE	1115
A. Bogomolov INFORMATION TECHNOLOGIES OF DIGITAL ADAPTIVE MEDICINE	1153

СОДЕРЖАНИЕ

Предисловие	1005
Математическое моделирование и прикладная математика Ю.С. Попков, Ю.А. Дубнов, А.Ю. Попков ПРОГНОЗИРОВАНИЕ РАЗВИТИЯ ЭПИДЕМИИ COVID-19 В СТРАНАХ ЕВРОПЕЙСКОГО СОЮЗА С ИСПОЛЬЗОВАНИЕМ ЭНТРОПИЙНО- РАНДОМИЗИРОВАННОГО ПОДХОДА	1010
В.В. Захаров, Ю.Е. Балыкина БАЛАНСОВАЯ МОДЕЛЬ ЭПИДЕМИИ COVID-19 НА ОСНОВЕ ПРОЦЕНТНОГО ПРИРОСТА	1034
В.Ю. Осипов, С.В. Кулешов, А.А. Зайцева, А.Ю. Аксенов ПОДХОД К ЛОКАЛИЗАЦИИ ИСТОЧНИКА ЭПИДЕМИИ COVID-19 В РОССИИ НА ОСНОВЕ МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ	1065
Искусственный интеллект, инженерия данных и знаний А.В. Смирнов, Е.Г. Молл, Н.Н. Тесля ИСПОЛЬЗОВАНИЕ НЕЧЕТКИХ КОАЛИЦИОННЫХ ИГР ПРИ ПРИНЯТИИ СОЦИАЛЬНО ОРИЕНТИРОВАННЫХ РЕШЕНИЙ ПРИ ГОСПИТАЛИЗАЦИИ В УСЛОВИЯХ ПАНДЕМИИ	1090
А.А. Двойникова, М.В. Маркитантов, Е.В. Рюмина, Д.А. Рюмин, А.А. Карпов АНАЛИТИЧЕСКИЙ ОБЗОР АУДИОВИЗУАЛЬНЫХ СИСТЕМ ДЛЯ ОПРЕДЕЛЕНИЯ СРЕДСТВ ИНДИВИДУАЛЬНОЙ ЗАЩИТЫ НА ЛИЦЕ ЧЕЛОВЕКА	1115
А.В. Богомолов ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ ЦИФРОВОЙ АДАПТАЦИОННОЙ МЕДИЦИНЫ	1153

Предисловие

Данный тематический выпуск журнала посвящен рассмотрению некоторых проблем использования математического моделирования и информационных технологий в здравоохранении и медицине в условиях современных вызовов, стоящих перед обществом. В выпуске представлены большей частью статьи, описывающие первые результаты выполнения ряда проектов Программы Российского фонда фундаментальных исследований «Фундаментальные проблемы возникновения и распространения коронавирусных эпидемий» (2020-2022 гг.).

Выпуск включает в себя два раздела. В первом разделе, посвященном математическому моделированию, приведены оригинальные результаты, описывающие:

– Прогнозирование развития эпидемии COVID-19 в странах Европейского союза с использованием энтропийно-рандомизированного подхода. Основу предложенного метода прогнозирования составляет идея оценивания распределений вероятностей параметров модели по реальным данным вместе с распределением вероятностей измерительных шумов. Предлагаемый подход используется для прогнозирования общего количества инфицированных с помощью трехпараметрической логистической модели роста и основан на использовании реальных данных о распространении COVID-19 в нескольких странах Европейского союза. Предложенный подход позволяет более эффективно прогнозировать развитие эпидемии по сравнению со стандартным подходом, основанным на методе наименьших квадратов.

– Прогнозирование развития эпидемии COVID-19 с использованием балансовой модели эпидемии. Данная модель основывается на новом подходе, когда для моделирования эпидемии предлагается использовать вместо традиционной модели SIR дискретную стохастическую модели распространения эпидемии CIR, основанную на балансе показателей эпидемии в текущий и прошлые моменты времени. Данная модель описывает динамику общего количества заболевших, общего количества выздоровевших и умерших и числа активных случаев. Параметрами модели являются процентный прирост величины заболевших во времени и характеристика динамического баланса эпидемиологического процесса.

– Определение локализации источника эпидемии COVID-19 в России на основе математического моделирования. Предложенная модель развития эпидемии COVID-19 основана на сети из девяти крупных городов России: Москва, Санкт-Петербург, Нижний Новгород, Ростов-на-Дону, Краснодар, Екатеринбург, Новосибирск, Хабаровск, Владивосток и состоит из двадцати семи дифференциальных уравнений. На основе оригинального алгоритма обратного анализа этой модели эпидемии установлены наиболее вероятные города-источники начала эпидемии в России, а также момент ее начала.

Во втором разделе выпуска, посвященном использованию информационных технологий и моделей искусственного интеллекта, приведены оригинальные результаты, описывающие:

– Использование нечетких коалиционных игр для принятия социально-ориентированных решений при госпитализации в условиях пандемии. В основу предложенного подхода к совершенствованию систем принятия решений при госпитализации пациентов положены идеи ситуационного управления и необходимость учета влияния психологических факторов при формировании коалиций участников (игроков). Ими являются госпитали, бригады скорой помощи, пациенты и центры компьютерной томографии. Цель игры – сформировать наборы коалиций участников, обеспечивающие максимальную выгоду по времени и стоимости госпитализации в момент принятия решения. Показано, что время расчета коалиционной игры позволяет использовать предложенную интеллектуальную модель поддержки принятия решений при госпитализации в диспетчерской службе станций скорой помощи.

– Применение современных аудиовизуальных систем для определения средств индивидуальной защиты. В рамках аналитического обзора показано, что для контроля и своевременного выявления нарушителей общественных правил здравоохранения в условиях COVID-19 необходимо применять современные информационные технологии, которые будут детектировать защитные маски на лицах людей по видео- и аудиоинформации. В результате обзора существующих и разрабатываемых интеллектуальных информационных технологий бимодального анализа голосовых и лицевых характеристик человека в маске установлено, что необходимо интенсифицировать исследования и разработки, направленные на

детектирование средств индивидуальной защиты органов дыхания по акустическим характеристикам речи человека.

– Концептуальные основы и информационные технологии цифровой адаптационной медицины. В частности, дана комплексная характеристика информационных технологий применительно к разработке специализированных автоматизированных комплексов, программных моделей и систем изучения адаптационных возможностей человека к условиям внешней среды, разработки методов и средств для повышения этих возможностей, а также проведения прикладных системных исследований обеспечения жизнедеятельности, эффективности, надежности деятельности, сохранения профессионального здоровья и продления профессионального долголетия человека. Охарактеризованы шесть базовых концепций адаптационной медицины.

Руководитель лаборатории
интегрированных систем
автоматизации СПб ФИЦ РАН,
доктор технических наук, профессор
А.В. Смирнов

Preface

This special issue of the journal is devoted to the consideration of some of the problems of using mathematical modeling and information technologies in health care and medicine in the context of modern challenges facing society. The issue contains mostly papers describing the first results of the implementation of a number of projects of the Russian Foundation for Basic Research Program "Fundamental problems of the emergence and spread of coronavirus epidemics" (2020-2022 years).

The issue includes two sections. The first section on mathematical modeling contains original results describing:

– Forecasting the development of the COVID-19 epidemic in the countries of the European Union use of the entropy-randomized approach. The proposed forecasting method is based on the idea of estimating the probability distributions of the model parameters and noises on real data. The proposed approach is used to predict the total number of infected people using a three-parameter logistic growth model and is based on real COVID-19 epidemic data in several countries of the European Union. This approach makes it possible to predict the development of the epidemic more efficiently in comparison with the standard approach based on the least-squares method.

– Forecasting the development of the COVID-19 epidemic using the epidemic balance model. This model is based on a new approach, when for modeling epidemic, instead of the traditional SIR model, it is proposed to use a discrete stochastic model of the CIR epidemic spread, based on the balance of the epidemic indicators in the current and past time points. This model describes the dynamics of the total number of cases, the total number of recovered and deaths, and the number of active cases. The parameters of the model are the percentage increase in the number of cases over time and the characteristic of the dynamic balance of the epidemiological process.

– Determination of the localization of the source of the COVID-19 epidemic in Russia based on mathematical modeling. The proposed model for the development of the COVID-19 epidemic is based on a network of nine large cities in Russia: Moscow, St. Petersburg, Nizhny Novgorod, Rostov-on-Don, Krasnodar, Yekaterinburg, Novosibirsk, Khabarovsk, Vladivostok and consists of twenty-seven differential equations. On the basis of the original algorithm for reverse analysis of this epidemic model, the most likely source cities of the epidemic beginning in Russia, as well as the moment of its beginning, have been identified.

The second section of the issue, devoted to the use of information technology and artificial intelligence models, presents original results describing:

– The use of fuzzy coalition games for making socially-oriented decisions during hospitalization in a pandemic. The proposed approach to improving decision-making systems during hospitalization of patients is based on the ideas of situational management and the need to take into account the influence of psychological factors when forming coalitions of participants (players). These are hospitals, ambulance teams, patients and computed tomography centers. The goal of the game is to form sets of coalitions of participants that provide the maximum benefit in terms of time and cost of hospitalization at the time of decision making. It is shown that the calculation time of the coalition game makes it possible to use the proposed intelligent model of decision support during hospitalization in the dispatch service of ambulance stations.

– The use of modern audiovisual systems for the identification of personal protective equipment. As part of the analytical review, it was shown that in order to control and timely identify violators of public health rules in the context of COVID-19, it is necessary to use modern information technologies that will detect protective masks on people's faces using video and audio information. As a result of a review of existing and developing intelligent information technologies for bimodal analysis of the voice and facial characteristics of a masked person, it was found that it is necessary to intensify research and development aimed at detecting personal respiratory protection equipment based on the acoustic characteristics of a person's speech.

– Conceptual foundations and information technologies of digital adaptive medicine. In particular, a comprehensive characteristic of information technologies is given in relation to the development of specialized automated complexes, program models and systems for studying the adaptive capabilities of a person to the conditions of the external environment, the development of methods and means to increase these capabilities, as well as conducting applied systemic studies of life support, efficiency, reliability of activities, preservation of professional health and extension of professional longevity of a person. Six basic concepts of adaptive medicine are characterized.

Head of Computer-Aided Integrated Systems Laboratory,
St.Petersburg Federal Research Center of the Russian Academy of Sciences,
Doctor of Technical Sciences, Professor
Alexander V. Smirnov

Ю.С. Попков, Ю.А. ДУБНОВ, А.Ю. ПОПКОВ
**ПРОГНОЗИРОВАНИЕ РАЗВИТИЯ ЭПИДЕМИИ COVID-19 В
СТРАНАХ ЕВРОПЕЙСКОГО СОЮЗА С ИСПОЛЬЗОВАНИЕМ
ЭНТРОПИЙНО-РАНДОМИЗИРОВАННОГО ПОДХОДА**

Попков Ю.С., Дубнов Ю.А., Попков А.Ю. Прогнозирование развития эпидемии COVID-19 в странах Европейского союза с использованием энтропийно-рандомизированного подхода.

Аннотация. Работа посвящена прогнозированию развития эпидемии COVID-19 с помощью нового метода рандомизированного машинного обучения. Основу метода составляет идея оценивания распределений вероятностей параметров модели по реальным данным вместе с распределением вероятностей измерительных шумов. Энтропийно-оптимальные распределения соответствуют состоянию максимальной неопределенности, что позволяет использовать получаемые в итоге прогнозы, как прогнозы наиболее “негативного” сценария исследуемого процесса. Полученные оценки параметров и шумов, которые представляют собой распределения вероятностей, необходимо генерировать, получая таким образом ансамбль траекторий, который требуется анализировать статистическими методами. Для целей такого анализа проводится вычисление средней и медианной по ансамблю траектории, а также траектории, соответствующей средним по распределению значениям параметров модели. Предлагаемый подход используется для прогнозирования общего количества инфицированных с помощью трехпараметрической логистической модели роста. Проведенный эксперимент основан на реальных данных о распространении COVID-19 в нескольких странах Европейского союза. Основной целью эксперимента является демонстрация энтропийно-рандомизированного подхода для прогнозирования эпидемического процесса на основе реальных данных вблизи пика. Существенная неопределенность, содержащаяся в доступных реальных данных моделируется аддитивным шумом в пределах 30%, который используется как на этапе обучения модели, так и при прогнозировании. Для настройки гиперпараметров модели используется схема их настройки по тестовой выборке с последующим переобучением. Показано, что при одинаковых наборах данных, предлагаемый подход позволяет более эффективно прогнозировать развитие эпидемии по сравнению со стандартным подходом, основанным на методе наименьших квадратов.

Ключевые слова: моделирование эпидемий, SARS-CoV-2, COVID-19, рандомизированное машинное обучение, энтропия, энтропийное оценивание, прогнозирование, рандомизированное прогнозирование.

1. Введение. Эпидемия новой коронавирусной инфекции началась предположительно в ноябре-декабре 2019 года в Китае, и уже к февралю 2020 года Всемирная организация здравоохранения присвоила ей статус пандемии, тем самым подтвердив ее глобальное распространение. Быстрый рост числа инфицированных и высокая смертность от вызываемой вирусом SARS-CoV-2 болезни COVID-19 привели к перегрузке систем здравоохранения почти во всех странах, включая страны Европейского союза и США. К марту 2020 года правительствам всех этих стран, в том числе и России, пришлось прибегать к крайним противоэпидемиче-

ским мерам, включая тотальный локдаун, запрет на перемещения людей, закрытие границ между регионами и государствами.

С самого начала пандемии начали проводиться научные исследования, направленные на прогнозирование развития эпидемии, основными целями которых являлась оценка масштабов эпидемии в кратко- и среднесрочной перспективе. Понимание эпидемических процессов и их прогнозирование необходимо для оценки готовности системы здравоохранения, принимаемых мер сдерживания эпидемии, состояния экономики.

Основными подходами к моделированию эпидемического процесса на сегодняшний день являются подходы, основанные на применении статических моделей подгонки кривой под имеющиеся данные; динамические модели, основанные на системах дифференциальных уравнений, описывающих динамику основных показателей эпидемии и сетевые модели, ориентированные на моделировании динамики эпидемических процессов с учетом неоднородной структуры общества, что отражается на неравномерном развитии эпидемии в разных группах населения.

Большинство подходов первого типа базируются на идеях, восходящих к моделям популяций в биологических системах, начиная с моделей экспоненциального роста конца 19 – начала 20 веков [1]. Модели экспоненциального роста показывают свою эффективность на начальном этапе развития эпидемии, когда происходит резкий рост числа инфицированных. Со временем, когда появляются естественные (например, увеличение доли иммунных членов популяции) или искусственные ограничения этого процесса (например, введение физических барьеров для ограничения контактов между членами популяции), происходит снижение количества инфицированных. Этот переломный момент в развитии эпидемического процесса, а также его дальнейшую динамику, простые модели экспоненциального роста не могут предсказать. По этой причине, в современном мире, в котором вводятся различные меры по сдерживанию эпидемии, применяются в основном модели логистического типа, которые используются для моделирования общего количества инфицированных. В начале 2020 года появилось большое количество исследований, основанных на логистических моделях, которые описывали эпидемический процесс с хорошим уровнем качества [2–8].

В настоящее время, практически во всех странах на уровне правительств используются динамические модели прогнозирования развития эпидемии, которые более адекватно описывают эпидемический процесс. Не вдаваясь в подробности разнообразных динамических моделей, можно отметить такое их общее свойство, как чувствительность их параметров, что приводит к новой проблеме адекватной и устойчивой их оценке,

без которой эффективное функционирование динамических моделей невозможно.

На начальном этапе эпидемии, исследование которого является предметом настоящей работы, произвести качественную оценку параметров динамических моделей не представлялось возможным из-за недостатка надежных реальных данных и понимания многих аспектов поведения вируса SARS-CoV-2 и болезни COVID-19. По этой причине, на рассматриваемом этапе развития эпидемии получили широкое распространение статические модели, в том числе логистического типа. Результаты применения этих моделей использовались многими исследователями для оценки параметров динамических моделей.

В работе рассматривается применение нового подхода к обучению моделей с использованием реальных данных и прогнозированию, называемого *Рандомизированным машинным обучением* (РМО) [9]. Основным достоинством этого метода является независимость от реальных характеристик используемых данных. Для корректного применения метода не требуется подтверждения или предположения о нормальности данных (или иных их вероятностных свойствах), а полученные в результате обучения распределения получены в условиях максимальной энтропии (максимальной неопределенности), таким образом отражая наиболее “плохой” сценарий развития исследуемого процесса. Эти свойства энтропийного подхода установлены в работах Больцмана [10], Джейнса [11, 12], Шеннона [13]. Еще одной важной особенностью метода является получение, вместе с оптимальными распределениями параметров, энтропийно-оптимальных распределений шумов (стохастических компонент данных), содержащихся в данных. Это свойство существенно отличает метод от классических подходов, в которых делаются различные предположения о характеристиках шумов.

2. Исходные данные и модели. Одним из важных и основных индикаторов эпидемии является общее (накопленное, куммулятивное) количество инфицированных. В случае эпидемии инфекционной болезни, важность этого показателя становится еще больше, так как количество инфицированных необходимо для оценивания и прогнозирования медицинских ресурсов, а также для оценивания других показателей эпидемического процесса, в том числе и его динамических характеристик.

Для моделирования рассматриваемого показателя в настоящее время (на начало 2021 года) в большинстве случаев применяются динамические модели, основанные на дифференциальных уравнениях с параметрами, которые оцениваются на основе реальных данных или с помощью других моделей. Статические модели, основанные на нели-

нейных уравнениях, параметры которых предполагается оценивать по реальным данным, стали популярными на начальной стадии эпидемии COVID-19 главным образом из-за недостатка реальных данных, и невозможности вследствие этого качественно оценить основные параметры динамических моделей. Кроме этого, опыт прошлых эпидемий, а также идея о снижении доли инфицированных в популяции из-за выздоровления или смерти, позволяет предположить, что кривая заражений будет иметь форму логистической кривой с участками плавного роста, быстрого роста и последующего снижения количества зараженных.

2.1. Модель. Динамика инфицированных членов популяции N в биологической системе может быть описана следующим уравнением [1]:

$$\frac{dN}{dt} = \lambda N \left(1 - \frac{N}{K} \right), \quad (1)$$

где λ — скорость роста инфицированных, N — количество инфицированных, K — объем популяции. Решением этого уравнения является кривая Ферхюльста [14, 15]

$$N(t) = \frac{K}{1 + B e^{-\lambda t}}, \quad B = \frac{K - N_0}{N_0}, \quad (2)$$

где N_0 — количество инфицированных в популяции в начальный момент времени [15].

Уравнение вида (2) активно использовалось в начале 2020 г. для построения предсказательных моделей заражения (предсказания общего количества заболевших) [2–8] и показали свою эффективность на начальном этапе развития эпидемии. Эти модели относятся к так называемым *логистическим моделям роста* (Logistic Growth Model, LGM), которые применяют для описания накопленного количества зараженных членов замкнутой популяции (кривую заражения).

Рассмотрим данную модель в общем виде. Для этого будем использовать уравнение трехпараметрической логистической кривой, которая определяет преобразование скалярного входа x в выход \hat{y} с использованием логистической нелинейной функции

$$\hat{y} = \Phi(x, \mathbf{a}) = \frac{a_3}{1 + a_1 e^{-a_2 x}}, \quad (3)$$

где $\mathbf{a} = (a_1, a_2, a_3)$ — вектор параметров модели. Входом модели является порядковый номер (или индекс) дня, а выходом — накопленное количество инфицированных.

Основной задачей для использования модели в прогнозировании, является ее обучение на реальных данных (оценивание ее параметров). После получения оценок параметров, возникает следующая задача, связанная с ее использованием для получения прогноза. Стандартные методы машинного обучения (статистического оценивания) состоят в вычислении точечных или интервальных оценок параметров модели, с последующим построением прогноза с использованием обученной модели. Этот подход является в большинстве случаев эффективным, однако при использовании нелинейных моделей он сопряжен с существенными трудностями в установлении свойств получаемых оценок, которые требуется учитывать при построении прогноза.

Подход к оцениванию параметров модели, предлагаемый в работе, основан на теории *рандомизированного машинного обучения* (РМО) [9, 16, 17], который предполагает получение не точечных оценок параметров, а их распределений вероятностей, причем вместе с распределениями вероятностей шумов, неизбежно содержащихся в данных. Важно отметить, что никаких априорных предположений о вероятностных свойствах реальных данных, которые бы учитывались при оценивании, не требуется. Для построения прогноза необходимо генерировать полученные оценки распределений параметров и шумов, что приводит к вычислению не одной прогнозной траектории, а ансамбля траекторий, который требуется анализировать статистическими методами.

Применение РМО приводит к построению *рандомизированной модели*, которая определяет специальную методику ее использования при прогнозировании — *рандомизированное прогнозирование* [18–21].

При применении методов математического моделирования эффективным является путь, который состоит в использовании непрерывных математических объектов и, соответственно, непрерывных моделей. Это обусловлено как достижениями в различных областях математики, так и точностью, которую могут обеспечить только непрерывные объекты. В то же время, работа в рамках непрерывной теории сопряжена с существенными вычислительными трудностями в случае многомерных объектов. Одним из путей решения этой проблемы является переход к дискретным объектам. Адаптация теории РМО на дискретный случай для нелинейных моделей была предложена в [22]. При применении этого подхода, можно построить *рандомизированную логистическую модель роста (РЛМР)*, которая является частным случаем *нелинейной рандомизированной дискретной статической модели (НРДСМ)*:

$$v = \hat{y} + \xi = \Phi(x, \mathbf{a}) + \xi, \quad (4)$$

где Φ определяется (3), x определяет вход модели, которым в рассматриваемой задаче является порядковый номер дня, ξ — шум при каждом измерении входа, действующий аддитивно, \hat{y} — выход модели (общее количество инфицированных), v — выход искаженный шумом.

Параметры и шумы реализуются интервальными дискретными случайными величинами с распределениями

$$a_{kl} \in A_k, \quad p_{kl} \in [0,1], \quad k = \overline{1,d}, \quad \ell = \overline{1,M}, \quad (5)$$

$$\xi_{jh} \in \Xi_j, \quad q_{jh} \in [0,1], \quad j = \overline{1,m}, \quad h = \overline{1,L}, \quad (6)$$

где a_{kl} , ξ_{jh} — значения случайных величин, p_{kl} , q_{jh} — вероятности их реализации, A_k , Ξ_j — интервалы значений случайных величин, m — количество точек данных, $d = 3$.

2.2. Данные. Для обучения и прогнозирования используются данные Университета Джонса Хопкинса [23, 24] о ежедневном количестве инфицированных. На рисунке 1 представлены данные вместе с 7-дневным скользящим средним для нескольких стран Европейского союза. Из данных видно, что зимой–весной 2020 года во всех странах наблюдался пик заражений, после которого произошел спад, продолжавшийся все лето, после которого возникла так называемая “вторая волна”. Учитывая, что исследования, проводимые в настоящей работе, направлены на моделирование и прогнозирование роста заражений вблизи пика, для обучения модели выбирались данные за 7 дней до пика “первой волны”, положение пика определялось по максимуму скользящего среднего за 7 дней.

Для обучения данные были масштабированы на отрезок $[0,1]$ с целью уменьшения сложностей, связанных с переполнением при вычислениях. При прогнозировании полученный выход модели преобразовывался к исходному масштабу.

2.3. Обучение. Рандомизированное машинное обучение предполагает обучение модели (3) с использованием реальных данных для получения оптимальных распределений параметров модели. Используемая здесь техника энтропийного оценивания позволяет определять такие распределения, которые доставляют максимум их энтропии. Полученные энтропийно-оптимальные распределения будут отражать наиболее неопределенный («плохой») сценарий, что в условиях полного отсутствия информации о реальных характеристиках является единственным доступным в этих условиях решением [9, 25, 26].

Вычисление оптимальных распределений обеспечивается решением задачи условной максимизации энтропии распределений параметров и шумов измерений при условиях нормировки соответствующих рас-

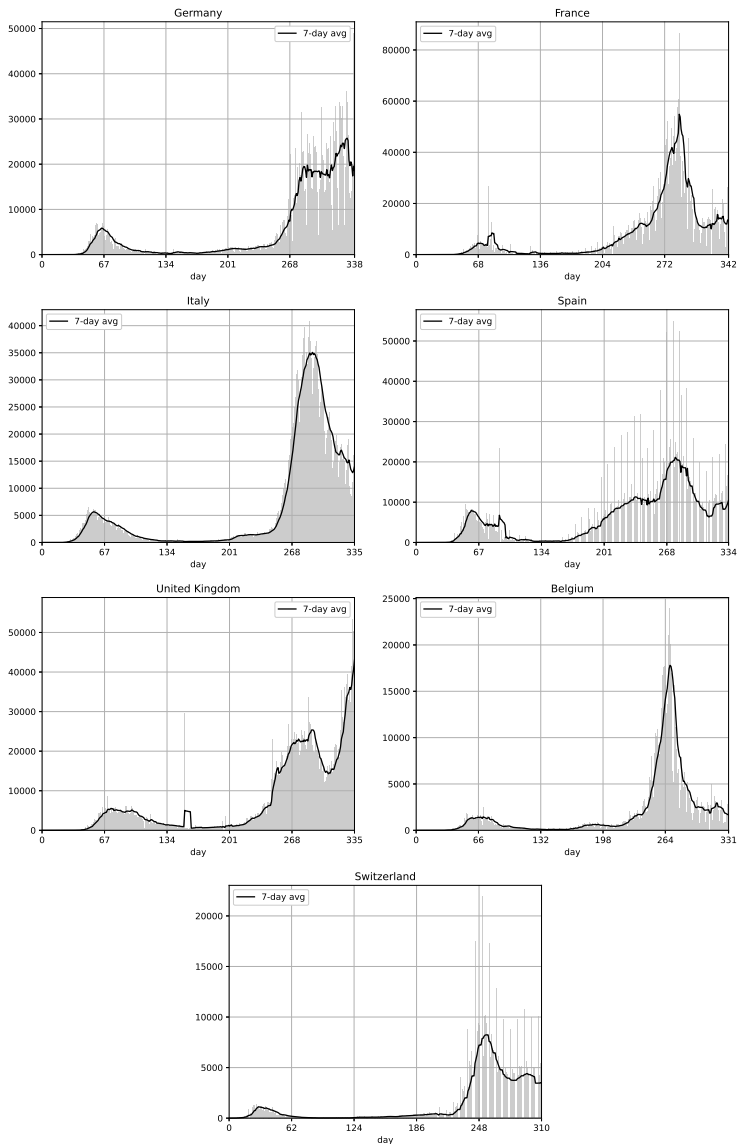


Рис. 1. Дневное количество инфицированных

пределений и выполнении условий на баланс среднего выхода модели с измерением выхода объекта, которая формулируется следующим образом [9]:

$$H(P, Q) = - \sum_{k=1}^d \sum_{\ell=1}^M p_{k\ell} \ln p_{k\ell} - \sum_{j=1}^m \sum_{h=1}^L q_{jh} \ln q_{jh} \rightarrow \max_{P, Q}, \quad (7)$$

где P и Q — распределения параметров и шумов (5) и (6), а $d = 3$ при условиях:

$$\sum_{\ell=1}^M p_{k\ell} = 1, \quad \sum_{h=1}^L q_{jh} = 1, \quad k = \overline{1, d}, \quad j = \overline{1, m}, \quad (8)$$

$$\mathbb{E}[v_j] = \mathbb{E}[\Phi(x_j, \mathbf{a}) + \xi_j] = y_j, \quad (9)$$

где y_j — реальные данные выхода, $\Phi(x, \mathbf{a})$ определяется (3), ξ_j — шум.

Условие (9) определяет баланс среднего выхода модели с реальными данными выхода:

$$\begin{aligned} \mathbb{E}[v_j] &= \mathbb{E}[\Phi(x_j, \mathbf{a}) + \xi_j] = \mathbb{E}[\Phi(x_j, \mathbf{a})] + \mathbb{E}[\xi_j] = \\ &= \sum_{\substack{\ell_k=1 \\ k=1, d}}^M \Phi(x_j, a_{1\ell_1}, \dots, a_{d\ell_d}) p_{1\ell_1} \cdots p_{d\ell_d} + \sum_{h=1}^L \xi_{jh} q_{jh} = \\ &= \bar{\Phi}(x_j) + \sum_{h=1}^L \xi_{jh} q_{jh} = y_j, \quad j = \overline{1, m}. \end{aligned} \quad (10)$$

Сумма в выражении для $\bar{\Phi}$ содержит M^d членов, суммирование осуществляется для всех комбинаций значений случайных величин $a_{k\ell}$.

Для решения задачи (7)-(9) используем метод множителей Лагранжа, применение которого позволяет получить выражения энтропийно-оптимальных распределений параметров и шумов измерений, зависящих

от параметра λ (множителей Лагранжа):

$$p_{k\ell}^*(\lambda) = \frac{\exp\left(-\sum_{j=1}^m \lambda_j \frac{\partial \Phi_j}{\partial p_{k\ell}}\right)}{\sum_{\ell=1}^M \exp\left(-\sum_{j=1}^m \lambda_j \frac{\partial \Phi_j}{\partial p_{k\ell}}\right)}, \quad k = \overline{1,d}, \ell = \overline{1,M}, \quad (11)$$

$$q_{jh}^*(\lambda) = \frac{\exp(-\lambda_j \xi_{jh})}{\sum_{h=1}^L \exp(-\lambda_j \xi_{jh})}, \quad j = \overline{1,m}, h = \overline{1,L}. \quad (12)$$

Множители λ определяются решением системы уравнений, получающихся подстановкой выражений (11)-(12) в балансовые соотношения (9):

$$\sum_{\substack{\ell_k=1 \\ k=1,d}}^M \Phi(\mathbf{x}_j, a_{1\ell_1}, \dots, a_{d\ell_d}) \prod_{\substack{\ell_s=1 \\ s=1,d}}^M p_{s\ell_s}^*(\lambda) + \sum_{h=1}^L \xi_{jh} q_{jh}^*(\lambda) = y_j, \quad j = \overline{1,m}. \quad (13)$$

Решение системы (13) позволяет определить требуемые энтропийно-оптимальные распределения параметров и шумов измерений. Для решения этой системы, очевидно, требуется привлекать численные методы, так как ее решение аналитически сопряжено с существенными трудностями. В работе для этой цели применяется усеченный метод Ньютона TNC [27, 28].

2.4. Прогнозирование. В результате обучения модель оказывается снабжена энтропийно-оптимальными оценками распределений параметров и измерительных шумов, формируя таким образом *рандомизированную предсказательную модель (РПМ)*. Применение РПМ для прогнозирования основано на специальной методике *рандомизированного прогнозирования*, которая определяет технологию использования энтропийно-оптимальных распределений параметров и шумов для построения прогноза. Указанные распределения можно использовать двумя способами. Первый состоит в использовании их средних значений в качестве точечных оценок параметров и шумов модели (4). Второй способ обладает большей эффективностью и состоит в генерации указанных распределений и построением ансамбля выхода модели для каждой реа-

лизации соответствующих случайных величин. Этот способ используется в вычислениях в настоящей работе, наряду с первым способом, который участвует в сравнительном исследовании. Для выполнения сравнительного исследования также используется стандартный способ прогнозирования, основанный на подгонке кривой нелинейным методом наименьших квадратов.

Вычисляя выход РПМ для каждой реализации параметров и шумов, мы получаем ансамбль траекторий, причем, для каждой реализации параметров, соответствующий шум в каждой точке данных генерируется независимо в количестве нескольких экземпляров.

Пусть имеется выборка параметров из распределения P^* объема S , определяемого (11). Для каждой реализации параметра \mathbf{a}_i , $i = \overline{1, S}$ генерируется U экземпляров шума с распределением Q^* определяемого (12) для каждого входа РПМ. Используя эти параметры и шумы, ансамбль выхода модели будет определяться блочным вектором или матрицей

$$\mathcal{V} = [\mathbf{v}_1, \dots, \mathbf{v}_W] = \begin{bmatrix} v_{11} & \dots & v_{1m} \\ \vdots & \ddots & \vdots \\ v_{W1} & \dots & v_{Wm} \end{bmatrix}, \quad (14)$$

где векторы \mathbf{v}_k (строки матрицы) соответствуют траектории выхода РПМ для одной реализации параметров и шумов, а общее количество таких векторов (строк) $W = SU$.

По полученному ансамблю можно вычислить различные траектории, характеризующие выборочные статистические величины. В настоящей работе используются средняя траектория (обозначаемая далее mean), медианная траектория (med), область стандартного отклонения (std). Кроме этого, для целей сравнения, может быть вычислена прогнозная траектория, соответствующая среднему значению параметров (avg).

3. Дизайн эксперимента. Данные об общем количестве инфицированных, используемые в экспериментальном исследовании для каждой страны, были организованы следующим образом.

Каждый набор данных был разделен на 3 непересекающиеся группы (интервала): интервал обучения T_{train} , интервал тестирования T_{test} и интервал прогнозирования T_{pred} . Следуя принятой в области анализа данных и машинного обучения методике, модель требуется обучить на интервале обучения, далее ее необходимо протестировать на интервале тестирования с целью улучшения предсказательной способности, и далее осуществить прогнозирование на интервале прогнозирования.

Таблица 1. Параметры набора данных

	d_0	T_{train}	T_{test}	T_{pred}
Германия	08-03-2020	[40, 57]	[58, 71]	[72, 101]
Франция	08-03-2020	[44, 69]	[70, 83]	[84, 113]
Италия	29-02-2020	[29, 47]	[48, 61]	[62, 91]
Испания	09-03-2020	[37, 51]	[52, 65]	[66, 95]
Великобритания	14-03-2020	[43, 66]	[67, 80]	[80, 109]
Швейцария	05-03-2020	[9, 21]	[22, 35]	[36, 65]
Бельгия	16-03-2020	[41, 63]	[64, 77]	[78, 107]

Для обучения предсказательной модели для каждой страны используются данные начиная со дня, когда впервые было зафиксировано общее количество инфицированных, превышающее 1000 человек (100 человек для Швейцарии), и заканчиваясь за 7 дней до пика весны 2020 года. Интервал тестирования составлял 14 дней, и интервал прогнозирования — 30 дней. Конфигурация всех наборов данных по каждой стране указаны в таблице 1, в колонке d_0 указана дата, соответствующая первому дню интервала обучения для соответствующей страны.

Рандомизированная модель (4)-(6) содержит несколько гиперпараметров, которые настраиваются на интервале тестирования. Этими параметрами являются интервалы A_k , интервалы шума устанавливаются для всех распределений равными $[-0,3; 0,3]$, что соответствует шуму в пределах 30%, параметры L и M принимаются постоянными и равными 5.

Основной метрикой для сравнения и настройки гиперпараметров является коэффициент детерминации R^2 , определяемый формулой:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}, \quad (15)$$

где y — реальные данные, \hat{y} — выход модели, \bar{y} — среднее значение, вычисленное по реальным данным.

Настройка гиперпараметров РПМ проводится следующим образом. Формируется набор конфигураций интервалов A_k путем вычисления сетки по каждому параметру с определенным шагом. Далее проводится обучение с каждой конфигурацией, и на интервале тестирования выбирается конфигурация интервалов с максимальным значением R^2 для средней по ансамблю траектории $mean$. Начальные интервалы устанавливаются в пределах 20% от точечной оценки, полученной методом наименьших квадратов на том же наборе данных.

После настройки гиперпараметров на интервале тестирования происходит переобучение РПМ на совмещенном наборе данных $\mathcal{T}_{train} \cup \mathcal{T}_{test}$ с полученными при настройке оптимальными интервалами A_k^* , и дальнейшее прогнозирование на интервале \mathcal{T}_{pred} , при этом при прогнозировании используется распределение шума, полученное для последней точки на интервале $\mathcal{T}_{train} \cup \mathcal{T}_{test}$.

Полученные рандомизированные прогнозы сравниваются с подгонкой кривой с помощью метода наименьших квадратов на том же наборе данных.

Кроме показателя R^2 также вычислялась средне-квадратичная ошибка (Mean-squared error, MSE) для истинных (реальных, true) значений y и предсказанных (модельных, predicted) значений \hat{y} по формуле

$$MSE = \frac{1}{n} \sum_{i=1}^{n-1} (y_i - \hat{y}_i)^2. \quad (16)$$

Все программное обеспечение для экспериментов реализовано на платформе Python 3.8 в среде Windows 10, метод наименьших квадратов реализован функцией `curve_fit`, метод оптимизации TNC, используемый для обучения РПМ, реализован функцией `minimize` библиотеки `scipy.optimize`.

4. Результаты и обсуждение. Полученные в экспериментах результаты изображены на рисунках 2-8 со следующими метками:

- `real` — реальные данные;
- `ols` — метод наименьших квадратов;
- `avg` — рандомизированное прогнозирование со средними по распределению значениями параметров модели;
- `mean` — рандомизированное прогнозирование со средним по ансамблю;
- `med` — рандомизированное прогнозирование с медианой по ансамблю;
- `7-avg` — 7-дневное скользящее среднее.

На рисунках указаны общие и ежедневные объемы инфицированных. На рисунках с общими объемами светло-серым цветом отмечены траектории, составляющие полученный ансамбль, темно-серым цветом — область стандартного отклонения по ансамблю. На рисунках с ежедневными объемами инфицированных светло-серым цветом указаны реальные данные и нанесены кривые, соответствующие 7-дневному скользящему среднему.

Таблица 2. Интервалы параметров РПМ

	A_1	A_2	A_3
Германия	[195231, 292847]	[0.1618, 0.2427]	[0.5902, 0.8853]
Франция	[8833, 13249]	[0.0940, 0.1411]	[1.0129, 1.5193]
Италия	[15476, 23215]	[0.1511, 0.2267]	[0.4993, 0.7489]
Испания	[207419, 311129]	[0.1753, 0.2630]	[0.5333, 0.8000]
Великобритания	[65218, 97828]	[0.1291, 0.1937]	[0.4687, 0.7030]
Швейцария	[1197, 1796]	[0.2108, 0.3162]	[0.5060, 0.7590]
Бельгия	[13241, 19862]	[0.1229, 0.1844]	[0.6245, 0.9368]

Все эксперименты производились для 1000 реализаций из распределения параметров модели, и 100 реализаций по распределениям шумов для каждого значения параметра. Генерация распределений шума проводилось для каждой точки соответствующего интервала. Таким образом, полученный ансамбль состоял из 10^5 траекторий. Вертикальная синяя пунктирная линия нанесена в точке начала интервала тестирования, красная пунктирная линия нанесена в точке начала интервала прогнозирования.

В таблице 2 указаны интервалы параметров A_k , полученные в результате тестирования модели и использованные при финальном обучении на итоговом интервале $\mathcal{T}_{train} \cup \mathcal{T}_{test}$.

В таблицах 3-4 приведены значения метрик, используемых в экспериментах для интервала обучения $\mathcal{T}_{train} \cup \mathcal{T}_{test}$ и прогнозирования \mathcal{T}_{pred} .

Анализируя полученные результаты, можно отметить следующее. Данные, используемые в экспериментах, соответствуют периоду «первой волны» эпидемии COVID-19 (весна 2020 года). В этот период все страны столкнулись с массовой нехваткой тестов, неготовностью медицинской и лабораторной инфраструктуры, а также с неготовностью протоколов лечения и диагностики, в следствие чего данные о заболевших в тот период имеют существенные искажения. Кроме этого, в тот период совершенствование тест-систем и протоколов тестирования происходило фактически в реальном времени, что приводило к получению более точных и полных данных. Эти соображения могут объяснить рост числа заболевших в определенные периоды времени, в то время как ожидалось их снижение вследствие введенных общенациональных локдаунов практически во всех европейских странах. Отмеченные эффекты наблюдаются в полученных в экспериментах результатах, в частности, такая ситуация видна в Великобритании. Стоит также отметить, что продолжающийся, хотя и более плавный, рост числа заболевших может быть связан с недостаточными

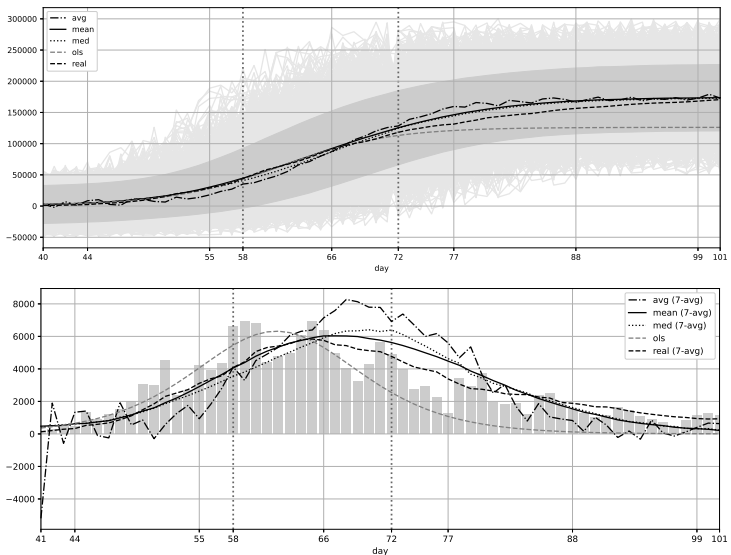


Рис. 2. Общее и ежедневное количество инфицированных в Германии

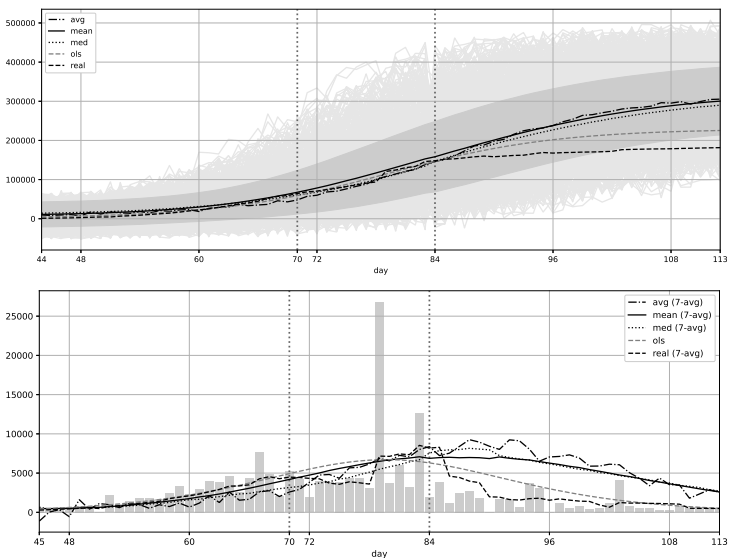


Рис. 3. Общее и ежедневное количество инфицированных во Франции

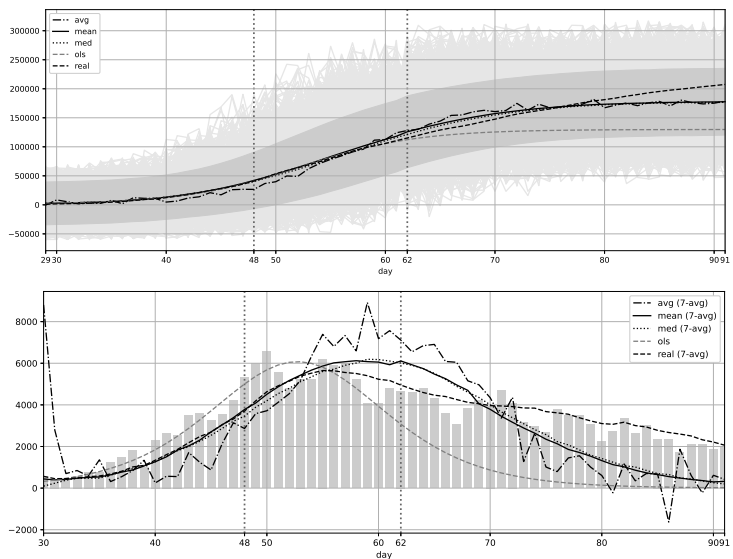


Рис. 4. Общее и ежедневное количество инфицированных в Италии

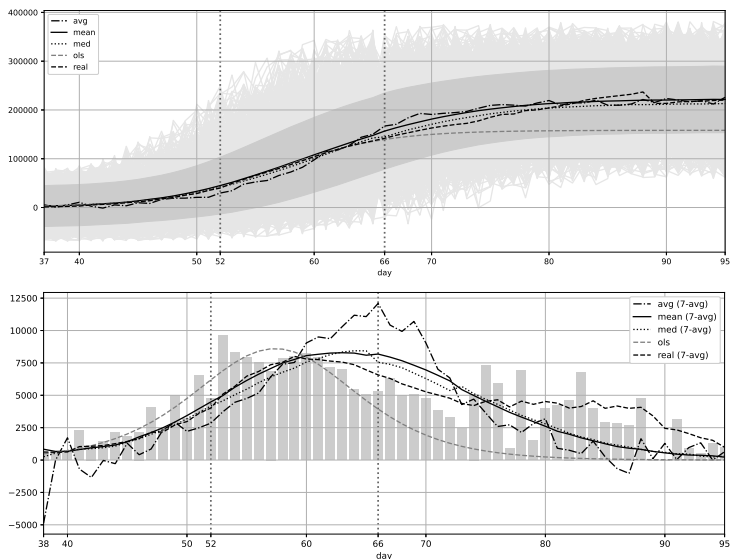


Рис. 5. Общее и ежедневное количество инфицированных в Испании

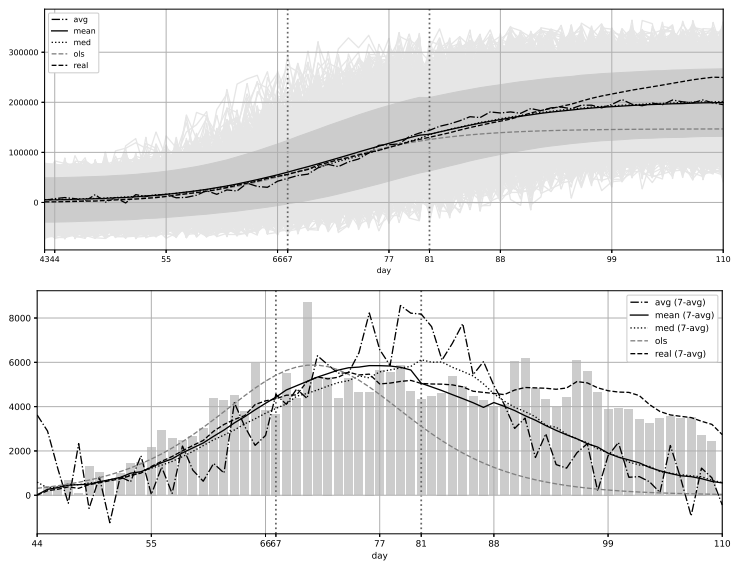


Рис. 6. Общее и ежедневное количество инфицированных в Великобритании

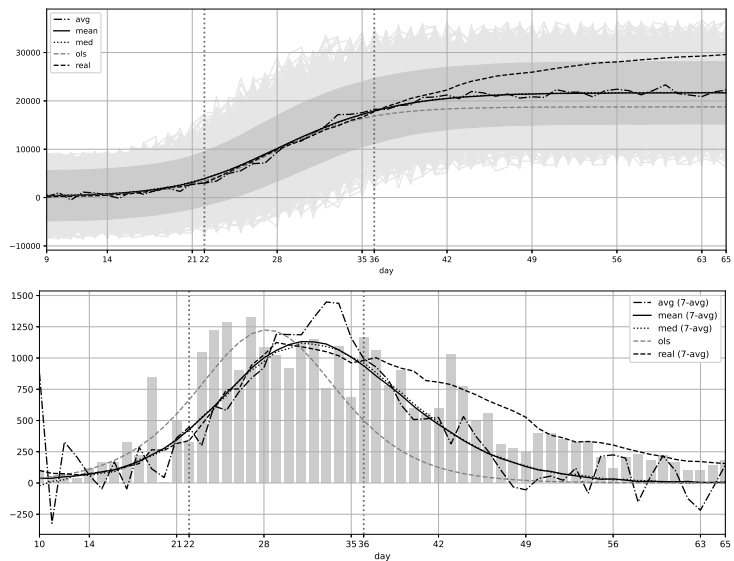


Рис. 7. Общее и ежедневное количество инфицированных в Швеции

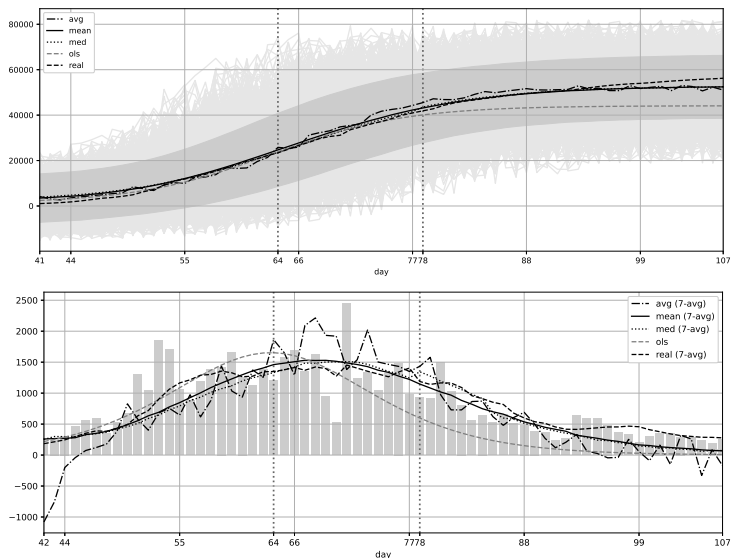


Рис. 8. Общее и ежедневное количество инфицированных в Бельгии

Таблица 3. Значение показателей качества на интервале обучения $T_{train} \cup T_{test}$

	ols	avg	mean	med
R^2				
Германия	0.9983	0.9483	0.9933	0.9875
Франция	0.9878	0.9639	0.9416	0.9372
Италия	0.9993	0.9583	0.9967	0.9933
Испания	0.9993	0.9533	0.9922	0.9956
Великобритания	0.9983	0.9754	0.9812	0.9691
Швейцария	0.9987	0.9846	0.9969	0.9958
Бельгия	0.9944	0.9651	0.9846	0.9758
MSE				
Германия	0.0001	0.0025	0.0003	0.0006
Франция	0.0007	0.0020	0.0033	0.0035
Италия	0.0000	0.0013	0.0001	0.0002
Испания	0.0000	0.0017	0.0003	0.0002
Великобритания	0.0000	0.0006	0.0005	0.0008
Швейцария	0.0000	0.0005	0.0001	0.0001
Бельгия	0.0003	0.0019	0.0009	0.0013

Таблица 4. Значение показателей качества на интервале прогнозирования \mathcal{T}_{pred}

	ols	avg	mean	med
	R^2			
Германия	-2.5876	-0.1733	0.4272	0.5883
Франция	-11.5808	-55.8622	-44.6427	-47.4572
Италия	-1.9247	0.6243	0.6975	0.7083
Испания	-2.1955	0.5739	0.8489	0.8536
Великобритания	-1.8281	0.2007	0.1640	0.2189
Швейцария	-4.0840	-1.2270	-1.5965	-1.4652
Бельгия	-3.1254	0.6807	0.7199	0.7511
	MSE			
Германия	0.0298	0.0097	0.0048	0.0034
Франция	0.0363	0.1640	0.1317	0.1398
Италия	0.0532	0.0068	0.0055	0.0043
Испания	0.0422	0.0056	0.0020	0.0019
Великобритания	0.0655	0.0185	0.0194	0.0181
Швейцария	0.0689	0.0302	0.0352	0.0334
Бельгия	0.0222	0.0017	0.0015	0.0013

противоэпидемическими мерами, а также с массовым несоблюдением этих мер населением.

Согласно показателям качества (см. табл. 3), полученным на интервале обучения, стандартный подход, основанный на методе наименьших квадратов, кажется более эффективным, качество приближения реальных данных с использованием трехпараметрической логистической модели превосходит соответствующие расчетные данные рандомизированной модели. Однако, на интервале прогнозирования ситуация существенно меняется в пользу последних (см. табл. 4).

Рандомизированный подход к прогнозированию, основная идея которого состоит в анализе ансамбля выхода модели, может быть реализован и для средних по распределению значений параметров модели. Такой способ реализации предсказательной модели соответствует стандартному подходу к прогнозированию, основанному на точечных оценках параметров. Из полученных результатов такого способа прогнозирования видно, что модельные значения выхода имеют большое отклонение от модельных значений, полученных при генерации соответствующих распределений, и иногда находятся в области отрицательных значений. Последнее наблюдение обусловлено свойством рандомизированной модели, в которой на выходе действует аддитивный шум, по причине которого модельные значения могут сильно отклоняться от средних по ансамблю траекторий. Это свойство такого прогноза может быть сглажено другим способом

учета шума при реализации модели, например, полной его фильтрацией. В целом необходимо отметить, что реализация указанного прогноза приводится в работе с целью сравнения как со стандартным методом, так и с подходом, основанным на генерации энтропийно-оптимальных распределений параметров и шумов.

Таким образом, можно констатировать преимущество энтропийно-рандомизированного подхода и его эффективность даже в условиях неполных и искаженных (зашумленных) данных. Необходимо также отметить, что при применении метода рандомизированного машинного обучения какие-либо априорные предположения о вероятностных свойствах данных не требуется.

5. Заключение. В работе предложен подход к прогнозированию развития эпидемии COVID-19, основанный на новом методе рандомизированного машинного обучения, состоящего в оценке распределений вероятностей параметров модели вместе с измерительными шумами. Применение данного метода продемонстрировано в эксперименте, одной из целей которого является сравнительное исследование со стандартными методами. Показаны преимущества предложенного метода в условиях малого количества зашумленных данных, каковыми являются эпидемические данные начального этапа. Важной особенностью используемого метода является его независимость от вероятностных свойств данных. Полученные результаты показывают, что даже в таких сложных условиях возможно построение адекватного прогноза, существенно превосходящего качества прогноза, полученного методом наименьших квадратов. Настоящая работа содержит результаты первого применения метода рандомизированного машинного обучения к моделированию эпидемического процесса. Успешное его применение к другим задачам интеллектуального анализа данных, таких как задачи классификации и задачи прогнозирования с помощью динамических регрессионных моделей, позволяет надеяться на его эффективность и в задачах, рассматриваемых в настоящей работе.

Литература

1. van den Driessche P. *Mathematical Epidemiology* / ed. by Brauer F., van den Driessche P., Wu J. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008. Vol. 1945. pp. 147–157.
2. Kumar J., Hembram K.P.S.S. Epidemiological study of novel coronavirus (COVID-19). ArXiv. 2020. URL: <http://arxiv.org/abs/2003.11376> (accessed 02.09.2021).
3. Yang W., Zhang D., Peng L., Zhuge C., and Hong L. Rational evaluation of various epidemic models based on the COVID-19 data of China. ArXiv. 2020. URL: <http://arxiv.org/abs/2003.05666> (accessed 02.09.2021).
4. Tátrai D., Várallyay Z. COVID-19 epidemic outcome predictions based on logistic fitting and estimation of its reliability. ArXiv. 2020. URL: <http://arxiv.org/abs/2003.14160> (accessed 02.09.2021).

5. Morais A.F. Logistic approximations used to describe new outbreaks in the 2020 COVID-19 pandemic. ArXiv. 2020. URL: <http://arxiv.org/abs/2003.11149> (accessed 02.09.2021).
6. Shen C.Y. Logistic growth modelling of COVID-19 proliferation in China and its international implications. *International Journal of Infectious Diseases*. 2020. vol. 96. pp. 582–589. URL: <https://doi.org/10.1016/j.ijid.2020.04.085> (accessed 02.09.2021).
7. Wang P., Zheng X., Li J., Zhu B. Prediction of epidemic trends in COVID-19 with logistic model and machine learning technics. *Chaos, Solitons & Fractals*. 2020. vol. 139. P. 110058. URL: <https://doi.org/10.1016/j.chaos.2020.110058> (accessed 02.09.2021).
8. Chen D.-G., Chen X., Chen J.K. Reconstructing and forecasting the COVID-19 epidemic in the United States using a 5-parameter logistic growth model. *Global Health Research and Policy*. 2020. vol. 5, no. 1. P. 25. URL: <https://doi.org/10.1186/s41256-020-00152-5> (accessed 02.09.2021).
9. Popkov Yu.S., Popkov A.Yu., Dubnov Yu.A. Randomizirovanoe mashinnoe obuchenie pri ogranichennyh naborah dannyh: ot jempiricheskoj verojatnosti k jentropijnoj randomizacii. [Randomized machine learning with limited datasets: from empirical probability to entropy randomization]. M.: LENAND, 2019.
10. Boltzman L. O svjazii mezhdv vtorym nachalom meh anicheskoj teorii teploty i teoriej verojatnostej v teoremah o teplovom ravnovesii. [On the connection between the second principle of the mechanical theory of heat and the theory of probabilities in theorems on thermal equilibrium]. Boltzman L.E. Izbrannye trudy. [Selected Works]. Ed. Slag L.S. Moscow: Nauka, 1984.
11. Jaynes E.T. Information theory and statistical mechanics. *Physical review*. 1957. vol. 106, no. 4. pp. 620–630.
12. Jaynes E.T. Probability theory: the logic of science. Cambridge university press. 2003.
13. Shannon C.E. Communication theory of secrecy systems. *Bell Labs Technical Journal*. 1949. vol. 28. no. 4. pp. 656–715.
14. Verhulst P.-F. Notice sur la loi que la population suit dans son accroissement. *Corresp. Math. Phys*. 1838. Vol. 10. pp. 113–126.
15. Singer H.M. The COVID-19 pandemic: growth patterns, power law scaling, and saturation. *Physical Biology*. 2020. vol. 17. no. 5. P. 055001.
16. Popkov Yu.S., Dubnov Yu.A., Popkov A.Yu. Randomized machine learning: Statement, solution, applications. *Intelligent Systems (IS), 2016 IEEE 8th International Conference on IEEE*. 2016. pp. 27–39.
17. Popkov Y.S., Dubnov Y.A., Popkov A.Y. Introduction to the Theory of Randomized Machine Learning. *Learning Systems: From Theory to Practice* / ed. by Sgurev V., Piuri V., Jotsov V. Cham: Springer International Publishing, 2018. pp. 199–220.
18. Popkov Y.S., Dubnov Y.A., Popkov A.Y. New method of randomized forecasting using entropy-robust estimation: Application to the World population prediction. *Mathematics*. 2016. vol. 4, no. 1. pp. 1–16.
19. Popkov Y.S., Volkovich Z., Dubnov Y.A., Avros R., Ravve E. Entropy 2-Soft Classification of Objects. *Entropy*. 2017. vol. 19, no. 4. P. 178.
20. Dubnov Y.A. Entropy-Based Estimation in Classification Problems. *Automation and Remote Control*. 2019. vol. 80. pp. 502–512.
21. Popkov Y.S., Popkov A.Y., Dubnov Y.A., Solomatine D. Entropy-Randomized Forecasting of Stochastic Dynamic Regression Models. *Mathematics*. 2020. No. 8. P. 1119.
22. Popkov A.Y. Randomized machine learning of nonlinear models with application to the prediction of the development of epidemic process. *Automation and Remote Control*. 2021.

23. Dong E., Du H., Gardner L. An interactive web-based dashboard to track COVID-19 in real time. *The Lancet infectious diseases*. 2020. vol. 20. no. 5. pp. 533–534.
24. COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University. URL: <https://github.com/CSSEGISandData/COVID-19> (accessed 02.09.2021).
25. Golan A., Judge G., Miller D. *Maximum Entropy Econometrics: Robust Estimation with Limited Data*. New York : John Wiley & Sons. 1996.
26. Golan A. Information and entropy econometrics. A review and synthesis. *Foundations and trends in Econometrics*. 2008. vol. 2, no. 1–2. pp. 1–145.
27. Nocedal J., Wright S. *Numerical optimization*. Springer Science & Business Media. 2006.
28. Nash S.G. Newton-type minimization via the Lanczos method. *SIAM Journal on Numerical Analysis*. 1984. vol. 21. no. 4. pp. 770–788.

Попков Юрий Соломонович — д-р техн. наук, профессор, академик РАН, главный научный сотрудник, Федеральный исследовательский центр "Информатика и управление" Российской академии наук; главный научный сотрудник, Институт проблем управления Российской академии наук. Область научных интересов: энтропийные методы, макросистемы, рандомизированное машинное обучение. Число научных публикаций — 220. popkov@isa.ru; ул. Вавилова, 44/2, 119133, Москва, Россия; р.т.: +79852111090.

Дубнов Юрий Андреевич — научный сотрудник, Федеральный исследовательский центр "Информатика и управление" Российской академии наук; Старший преподаватель, Национальный исследовательский университет "Высшая школа экономики". Область научных интересов: машинное обучение, байесовское оценивание. Число научных публикаций — 30. yury.dubnov@phystech.edu; ул. Вавилова, 44/2, 119133, Москва, Россия; р.т.: +79099634498.

Попков Алексей Юрьевич — канд. техн. наук, ведущий научный сотрудник, Федеральный исследовательский центр "Информатика и управление" Российской академии наук. Область научных интересов: энтропийные методы, рандомизированное машинное обучение, интеллектуальный анализ данных, разработка программного обеспечения. Число научных публикаций — 47. aropkov@isa.ru; ул. Вавилова, 44/2, 119133, Москва, Россия; р.т.: +7(499)135-62-60.

Поддержка исследований. Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (проект № 20-04-60119).

Y. POPKOV, Y. DUBNOV, A. POPKOV
**FORECASTING DEVELOPMENT OF COVID-19 EPIDEMIC IN
EUROPEAN UNION USING ENTROPY-RANDOMIZED APPROACH**

Popkov Y., Dunbov Y., Popkov A. Forecasting Development of COVID-19 Epidemic in European Union Using Entropy-Randomized Approach.

Abstract. The paper is devoted to the forecasting of the COVID-19 epidemic by the novel method of randomized machine learning. This method is based on the idea of estimation of probability distributions of model parameters and noises on real data. Entropy-optimal distributions correspond to the state of maximum uncertainty which allows the resulting forecasts to be used as forecasts of the most “negative” scenario of the process under study. The resulting estimates of parameters and noises, which are probability distributions, must be generated, thus obtaining an ensemble of trajectories that considered to be analyzed by statistical methods. In this work, for the purposes of such an analysis, the mean and median trajectories over the ensemble are calculated, as well as the trajectory corresponding to the mean over distribution values of the model parameters. The proposed approach is used to predict the total number of infected people using a three-parameter logistic growth model. The conducted experiment is based on real COVID-19 epidemic data in several countries of the European Union. The main goal of the experiment is to demonstrate an entropy-randomized approach for predicting the epidemic process based on real data near the peak. The significant uncertainty contained in the available real data is modeled by an additive noise within 30%, which is used both at the training and predicting stages. To tune the hyperparameters of the model, the scheme is used to configure them according to a testing dataset with subsequent retraining of the model. It is shown that with the same datasets, the proposed approach makes it possible to predict the development of the epidemic more efficiently in comparison with the standard approach based on the least-squares method.

Keywords: epidemic modelling, SARS-CoV-2, COVID-19, randomized machine learning, entropy, entropy estimation, forecasting, randomized forecasting.

Popkov Yuri — Ph.D., Dr.Sci., Professor, Academician of RAS, Chief research scientist, Federal Research Center “Computer Science and Control” of Russian Academy of Sciences; Chief research scientist, Institute of Control Problems of Russian Academy of Sciences. Research interests: entropy, macrosystems, randomized machine learning, optimization. The number of publications — 220. popkov@isa.ru; 44/2, Vavilov Str., 119133, Moscow, Russia; office phone: +79852111090.

Dubnov Yuri — Research scientist, Federal Research Center “Computer Science and Control” of Russian Academy of Sciences; Senior lecturer, National Research University Higher School of Economics. Research interests: machine learning, bayesian estimation. The number of publications — 30. yury.dubnov@phystech.edu; 44/2, Vavilov Str., 119133, Moscow, Russia; office phone: +79099634498.

Popkov Alexey — Ph.D., Leading research scientist, Federal Research Center “Computer Science and Control” of Russian Academy of Sciences. Research interests: entropy methods, machine learning, data mining, software development. The number of publications — 47. apopkov@isa.ru; 44/2, Vavilova Str., 119133, Moscow, Russia; office phone: +7(499)135-62-60.

Acknowledgements. This work was supported by Russian Foundation for Basic Research (project no. 20-04-60119).

References

1. van den Driessche P. *Mathematical Epidemiology* / ed. by Brauer F., van den Driessche P., Wu J. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008. Vol. 1945. pp. 147–157.
2. Kumar J., Hembram K.P.S.S. Epidemiological study of novel coronavirus (COVID-19). ArXiv. 2020. URL: <http://arxiv.org/abs/2003.11376> (accessed 02.09.2021).
3. Yang W., Zhang D., Peng L., Zhuge C., and Hong L. Rational evaluation of various epidemic models based on the COVID-19 data of China. ArXiv. 2020. URL: <http://arxiv.org/abs/2003.05666> (accessed 02.09.2021).
4. Tátrai D., Várallyay Z. COVID-19 epidemic outcome predictions based on logistic fitting and estimation of its reliability. ArXiv. 2020. URL: <http://arxiv.org/abs/2003.14160> (accessed 02.09.2021).
5. Morais A.F. Logistic approximations used to describe new outbreaks in the 2020 COVID-19 pandemic. ArXiv. 2020. URL: <http://arxiv.org/abs/2003.11149> (accessed 02.09.2021).
6. Shen C.Y. Logistic growth modelling of COVID-19 proliferation in China and its international implications. *International Journal of Infectious Diseases*. 2020. vol. 96. pp. 582–589. URL: <https://doi.org/10.1016/j.ijid.2020.04.085> (accessed 02.09.2021).
7. Wang P., Zheng X., Li J., Zhu B. Prediction of epidemic trends in COVID-19 with logistic model and machine learning technics. *Chaos, Solitons & Fractals*. 2020. vol. 139. P. 110058. URL: <https://doi.org/10.1016/j.chaos.2020.110058> (accessed 02.09.2021).
8. Chen D.-G., Chen X., Chen J.K. Reconstructing and forecasting the COVID-19 epidemic in the United States using a 5-parameter logistic growth model. *Global Health Research and Policy*. 2020. vol. 5, no. 1. P. 25. URL: <https://doi.org/10.1186/s41256-020-00152-5> (accessed 02.09.2021).
9. Popkov Yu.S., Popkov A.Yu., Dubnov Yu.A. Randomizirovannoe mashinnoe obuchenie pri ogranichennykh naborah dannyh: ot jempiricheskoy veroyatnosti k jentropijnoj randomizacii. [Randomized machine learning with limited datasets: from empirical probability to entropy randomization]. M.: LENAND, 2019.
10. Boltzman L. O svyazi mezhdv vtorym nachalom meh anicheskoj teorii teploty i teoriej veroyatnostej v teoremah o teplovom ravnesii. [On the connection between the second principle of the mechanical theory of heat and the theory of probabilities in theorems on thermal equilibrium]. Boltzman L.E. *Izbrannyye trudy*. [Selected Works]. Ed. Slag L.S. Moscow: Nauka, 1984.
11. Jaynes E.T. *Information theory and statistical mechanics*. Physical review. 1957. vol. 106, no. 4. pp. 620–630.
12. Jaynes E.T. *Probability theory: the logic of science*. Cambridge university press. 2003.
13. Shannon C.E. Communication theory of secrecy systems. *Bell Labs Technical Journal*. 1949. vol. 28, no. 4. pp. 656–715.
14. Verhulst P.-F. Notice sur la loi que la population suit dans son accroissement. *Corresp. Math. Phys.* 1838. Vol. 10. pp. 113–126.
15. Singer H.M. The COVID-19 pandemic: growth patterns, power law scaling, and saturation. *Physical Biology*. 2020. vol. 17, no. 5. P. 055001.
16. Popkov Yu.S., Dubnov Yu.A., Popkov A.Yu. Randomized machine learning: Statement, solution, applications. *Intelligent Systems (IS)*, 2016 IEEE 8th International Conference on IEEE. 2016. pp. 27–39.

17. Popkov Y.S., Dubnov Y.A., Popkov A.Y. Introduction to the Theory of Randomized Machine Learning. Learning Systems: From Theory to Practice / ed. by Sgurev V., Piuri V., Jotsov V. Cham: Springer International Publishing, 2018. pp. 199–220.
18. Popkov Y.S., Dubnov Y.A., Popkov A.Y. New method of randomized forecasting using entropy-robust estimation: Application to the World population prediction. Mathematics. 2016. vol. 4, no. 1. pp. 1–16.
19. Popkov Y.S., Volkovich Z., Dubnov Y.A., Avros R., Ravve E. Entropy 2-Soft Classification of Objects. Entropy. 2017. vol. 19, no. 4. P. 178.
20. Dubnov Y.A. Entropy-Based Estimation in Classification Problems. Automation and Remote Control. 2019. vol. 80. pp. 502–512.
21. Popkov Y.S., Popkov A.Y., Dubnov Y.A., Solomatine D. Entropy-Randomized Forecasting of Stochastic Dynamic Regression Models. Mathematics. 2020. No. 8. P. 1119.
22. Popkov A.Y. Randomized machine learning of nonlinear models with application to the prediction of the development of epidemic process. Automation and Remote Control. 2021.
23. Dong E., Du H., Gardner L. An interactive web-based dashboard to track COVID-19 in real time. The Lancet infectious diseases. 2020. vol. 20. no. 5. pp. 533–534.
24. COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University. URL: <https://github.com/CSSEGISandData/COVID-19> (accessed 02.09.2021).
25. Golan A., Judge G., Miller D. Maximum Entropy Econometrics: Robust Estimation with Limited Data. New York : John Wiley & Sons. 1996.
26. Golan A. Information and entropy econometrics. A review and synthesis. Foundations and trends in Econometrics. 2008. vol. 2, no. 1–2. pp. 1–145.
27. Nocedal J., Wright S. Numerical optimization. Springer Science & Business Media. 2006.
28. Nash S.G. Newton-type minimization via the Lanczos method. SIAM Journal on Numerical Analysis. 1984. vol. 21. no. 4. pp. 770–788.

В.В. ЗАХАРОВ, Ю.Е. БАЛЫКИНА
**БАЛАНСОВАЯ МОДЕЛЬ ЭПИДЕМИИ COVID-19 НА ОСНОВЕ
ПРОЦЕНТНОГО ПРИРОСТА**

Захаров В.В., Балыкина Ю.Е. **Балансовая модель эпидемии COVID-19 на основе процентного прироста.**

Аннотация. В статье изучается возможность использования альтернативного подхода к прогнозированию статистических показателей эпидемии вируса нового типа. Представлен систематический обзор моделей прогнозирования эпидемий новых инфекций в зарубежной и российской научной литературе. Анализируется точность модели SIR при прогнозировании весенней волны эпидемии COVID-19 в России. В качестве альтернативного подхода к моделированию эпидемии предлагается использование вместо традиционной модели SIR новой дискретной стохастической модели распространения эпидемии CIR, основанной на балансе показателей эпидемии в текущий и прошлые моменты времени. Новая модель описывает динамику общего количества заболевших (C), общего количества выздоровевших и умерших (R) и числа активных случаев (I). Параметрами системы являются процентный прирост величины $C(t)$ и характеристика динамического баланса эпидемиологического процесса, впервые введенная в этой статье. Сформулирован принцип динамического баланса эпидемиологического процесса, предполагающий наличие у любого процесса свойства близости значений общего количества заболевших в прошлые периоды и значений общего количества выздоровевших и умерших в текущий момент времени. Для вычисления значений характеристики динамического баланса используется задача целочисленного программирования. Продемонстрировано, что в общем случае динамическая характеристика эпидемиологического процесса не является постоянной величиной. Эпидемиологический процесс, динамическая характеристика которого не является постоянной величиной, называется нестационарным. Для построения среднесрочных прогнозов показателей эпидемиологического процесса на промежутках стационарности эпидемиологического процесса разработан специальный алгоритм. Исследован вопрос об использовании этого алгоритма на промежутках стационарности и нестационарности. Приведены примеры применения модели CIR для построения прогнозов рассматриваемых показателей эпидемии в России в мае-июне 2020 года.

Ключевые слова: COVID-19, модели распространения, моделирование эпидемии новых вирусов, SIR-модели, прогнозирование.

1. Введение. Вспышка коронавирусной инфекции COVID-19, вызванная новым вирусом SARS-CoV-2, быстро распространилась по миру в конце 2019 года, затронув уже более 200 стран. 31 января 2020 года Всемирная организация здравоохранения (ВОЗ) объявила эту вспышку чрезвычайной ситуацией, имеющей международное значение, а 11 марта 2020 года классифицировала ее как пандемию. По состоянию на середину мая 2021 года в мире зарегистрировано более 159 млн. подтвержденных случаев COVID-19. По признанию ВОЗ, наличие своевременных математических моделей играет ключевую роль в принятии решений, основанных на фактах, лицами, определяющими политику здравоохранения. Рассмотрение инфекционных заболеваний

с математической точки зрения может выявить не только важные закономерности пандемии, но и возможные варианты эпидемического контроля. Математические модели передачи болезней помогают получить представление о динамике распространения инфекционных заболеваний и потенциальной роли различных видов стратегий вмешательства со стороны общественного здравоохранения [1]. Важно в условиях максимальной неопределенности своевременно выявлять параметры распространения новых инфекционных заболеваний, чтобы оценить потенциальные опасности и масштабы эпидемии.

Особенность пандемии COVID-19 состоит в отсутствии достоверной статистической информации о распространении вируса SARS-CoV-2 за прошлые годы. Это влечёт за собой проблему корректного использования имеющихся данных о параметрах развивающейся эпидемии. Научно-исследовательские группы в разных странах мира ведут работы по созданию методов краткосрочного и долгосрочного прогнозирования динамики распространения нового вируса.

Во втором разделе проанализировано современное состояние исследований по прогнозированию распространения эпидемий и описаны основные подходы к моделированию. В третьем разделе проводится анализ точности прогнозирования весенней волны эпидемии в России при использовании дискретной модели SIR. В четвертом разделе представлена новая балансовая модель распространения эпидемии на основе процентного прироста (CIR), вводится понятие характеристики динамического баланса эпидемии. В пятом разделе обсуждаются результаты применения модели CIR и метода прецедентов (CBR – case based reasoning) для построения краткосрочных прогнозов динамики эпидемии COVID-19 в России в мае-июне 2020 года в реальном времени.

2. Обзор научной литературы. Мы провели систематический обзор моделей прогнозирования эпидемий новых инфекций. Поиск производился в наукометрических базах данных Web of Science, Scopus, RCI. Дополнительно были проанализированы результаты поиска в базе Elsevier – одного из крупнейших европейских издательств. В качестве ключевых слов использовались “forecasting”, “prediction”, “model”, “emerging infection”, “coronavirus”, “covid”. Большая часть исследований, удовлетворяющих критериям поиска, касалась распространения эпидемии нового коронавируса SARS-CoV-2 (COVID-19). Модели эпидемии различаются по типу модели, способу получения исходных данных, гипотезе и распределению ключевых входных параметров. Все исследования старались ответить на следующие вопросы: (1) заразность рассматриваемого вируса, и связанная с этим оценка

значения базового репродуктивного числа; (2) оценка ключевых временных периодов распространения эпидемии (достижение пика, плато, затухание); (3) прогнозирование распространения инфекции (включая краткосрочное и долгосрочное прогнозирование); (4) оценка воздействия различных мер, направленных на сдерживание эпидемии. Также востребованным вопросом является влияние неизвестной ранее эпидемии на макро- и микроэкономические показатели регионов и стран.

В целом можно выделить четыре основных подхода к моделированию распространения инфекционных заболеваний [2]: камерные модели, модели на основе временных рядов, агентно-ориентированные модели и модели, построенные с использованием методов машинного обучения и эвристических подходов (рис. 1).



Рис. 1. Основные подходы к моделированию распространения эпидемий инфекционных заболеваний

Модели на основе временных рядов. Говоря о моделях на основе временных зависимостей, следует отметить, что регрессионный анализ и анализ временных рядов являются одними из самых известных методов прогнозирования распространения заболеваемости. Одной из наиболее часто используемых моделей является модель $ARIMA(p,d,q)$ – интегрированная модель авторегрессии – скользящего среднего. Аббревиатура $ARIMA$ носит описательный характер и отражает ключевые аспекты самой модели. AR (AutoRegressive) отвечает за авторегрессионную часть, т.е. за порядок запаздывания (p); I (Integrated) отвечает за степень разности (d); MA (Moving Average) отвечает за размер окна скользящего среднего, также называемого порядком скользящего среднего (q). Каждый из этих компонентов явно указывается в модели в качестве параметра.

Некоторые исследователи использовали $ARIMA$ для прогнозирования распространения пандемии новой коронавирусной инфекции. Так, в [3] авторы оценивали количество ежедневных новых случаев

заболевания в Иране. Результаты прогнозирования показывали в районе 9 558 новых случаев в день на конец апреля 2020 года. Официально на данную дату было зафиксировано 1 073 новых случая заражения. В работе [4] использовался метод простого скользящего среднего для прогнозирования подтвержденных случаев COVID-19 в Пакистане. Авторы прогнозировали более 35 000 случаев заболевания к концу мая 2020 года. (Фактически на конец мая было зафиксировано в 2 раза больше, 72 460 случаев новой коронавирусной инфекции). Тандон с соавторами [5] сравнивает результаты прогнозирования распространения эпидемии COVID-19 в Индии с помощью 7 различных моделей временных рядов. Горизонт моделирования составлял 20 дней. Авторы заключают, что лучшие результаты среди сравниваемых показала модель ARIMA. По данным авторов, к 3 мая 2020 года число случаев заражения должно было быть около 30 990 случаев. Фактически, на данную дату было зафиксировано 42 505 случаев заражения.

Очевидно, что хотя модели временных рядов и являются популярным инструментом прогнозирования, применение данного подхода для оценки распространения новых инфекций имеет свои ограничения. В частности, отсутствие статистики за предыдущие периоды и, как следствие, неизвестные значения параметров не позволяют построить модели достаточной степени точности.

Камерные модели. В базовой камерной модели рассматривается популяция из N человек. В каждый момент времени каждый человек относится к одной из трех групп (камер): в группу восприимчивых S (Susceptible) входят люди, которые еще не сталкивались с инфекцией, затем, по мере того как вирус распространяется среди населения, они переходят в группу I инфицированных (Infected), а затем – в группу R (Removed) выбывших (выздоровевших или умерших), при этом возможность повторного заражения в данной модели не предусматривается. Также считается, что размер популяции остается неизменным: $N = S(t) + I(t) + R(t)$. Впервые так называемая модель эпидемии SIR, описываемая системой из трех обыкновенных дифференциальных уравнений для переменных S , I , R , была представлена в статье У. Кермака и А. Маккендрика в 1927 году [6, 7].

Модель SIR определяется всего двумя параметрами: эффективная частота контактов (β), которая влияет на переход из группы восприимчивых в группу инфицированных, и скорость выздоровления (γ), которая влияет на переход из группы инфицированных во множество выздоровевших или умерших (рис. 2). Соответ-

ствующее соотношение $R_0 = \frac{\beta}{\gamma}$ носит название базового репродуктивного числа. Оно определяется как среднее количество новых случаев, вызванных одним случаем, в течение инфекционного периода среди полностью восприимчивого населения. Эпидемия возникает тогда и только тогда, когда $R_0 > 1$.



Рис. 2. Общая схема камерной модели SIR

Некоторые исследователи предпочитают использовать SIR из-за небольшого количества требуемых на входе параметров. Однако, это преимущество возникает из-за порой чрезмерного упрощения модели за счет относительно нереалистичных предположений. Например, модель предполагает однородное смешивание популяции, а это означает, что все особи в популяции имеют равную вероятность вступить в контакт друг с другом. Это не отражает человеческие социальные структуры, в которых большинство контактов происходит в ограниченных сообществах. Модель SIR также предполагает замкнутую популяцию без миграции, рождений или смертей по причинам, отличным от эпидемии.

В качестве примера можно привести исследование Дил и Маке-на [8], в котором использовали модель SIR для прогнозирования подтвержденных случаев COVID-19 в регионе Восточного Средиземноморья (Иран, Ирак, Саудовская Аравия, ОАЭ, Ливан, Египет и Пакистан). По оценкам авторов, к 20 июня 2020 года предполагалось достичь 2,12 миллиона случаев в Иране, 0,58 миллиона в Саудовской Аравии и 0,51 миллиона в Пакистане. В действительности, число зафиксированных случаев заражения на 20.06.2020 официально составляло 202 584 случаев заражения в Иране, 176 617 случаев – в Пакистане и 154 233 случаев – в Саудовской Аравии (по данным Центра CSSE Университета Джонса Хопкинса [9]).

Кроме того, параметры в традиционной модели SIR не позволяют количественно оценить неопределенность параметров модели. Распространенной стратегией прогнозирования развития эпидемии является расчет модели SIR по нескольким возможным значениям для каждого параметра, результатом при этом является диапазон будущих траекторий. Более сложные модели используют распределения для каждого параметра вместо точечной оценки для характеристики веро-

ятности различных будущих траекторий. Также, в более сложных адаптациях камерной структуры SIR, можно формально включить наблюдаемые данные, чтобы значения параметров калибровались на основе поступающих данных [10].

Модель SEIR [11] похожа на модель SIR, за исключением того, что в неё добавляется камера E (Exposed) для людей, которые уже были инфицированы, но сами еще не являются заразными (за переход в данную группу отвечает параметр α). Далее, по истечении инкубационного периода, люди из группы E переходят уже в группу распространителей вируса I (рис. 3). При этом модель допускает возможность того, что вероятность передачи при контакте с зараженным человеком может отличаться от вероятности передачи при контакте с инфицированным человеком.

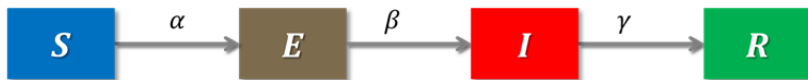


Рис. 3. Общая схема камерной модели SEIR

Помимо SEIR, часто используется модель SIRD, в которой предусмотрены отдельные камеры для выздоровевших людей (тех, кто пережил болезнь и теперь имеет иммунитет) и умерших людей. В качестве примера можно привести исследование [12].

Основная проблема всех камерных моделей состоит в том, что полученные с их помощью прогнозы достаточно чувствительны к значениям параметров на входе модели, таким как R_0 и другим. При этом, данные характеристики возможно получить, лишь набрав достаточный объем статистических сведений о динамике заболевания. В случае рассмотрения новых инфекций это представляет собой большую проблему. Например, в случае коронавирусной инфекции COVID-19 в первую очередь рассматривали данные Китая и Италии как стран, где эпидемия развивалась раньше всех. Но при этом использовать напрямую полученные значения параметров при моделировании в других странах нельзя, в силу демографических и географических особенностей рассматриваемых регионов (например, средний уровень иммунитета жителей, возрастной состав, вероятность мутаций вируса, климат и т.д.) [13].

В качестве примера использования модели SEIR, можно привести исследование Ченга и Бурку [14], в котором представлена платформа `icumonitoring.ch` для прогнозирования загруженности отделений интенсивной терапии (ОИТ) на уровне больниц на основе модели

SEIR. Данные обновляются каждые 3-4 дня, и предполагается, что такая платформа может помочь менеджерам ICU оценить потребность в дополнительных ресурсах. Простая модель SIRD была предложена в [12] для прогнозирования всплеск COVID-19 в Китае, Италии и Франции. В данном исследовании также была проведена оценка потребностей медицинских учреждений (например, оценивалось необходимое количество аппаратов ИВЛ). На сегодняшний день опубликовано несколько десятков моделей SIR и её разновидностей только по COVID-19. В частности, модели SIR [15-17], SIR-X [18], SEIR [19-25], SIRD [26], SEIQR [27], SEIHR [28].

В целом, камерные модели хорошо себя зарекомендовали и являются своего рода «золотым стандартом» при моделировании динамики распространения эпидемии. В то же время, в случае рассмотрения новых инфекций данный тип моделей показывает не лучшие результаты, особенно на ранних стадиях эпидемии. Этому есть несколько причин, основная из которых – недостаток достоверной информации о значениях параметров. Фактически, недостоверные исходные данные могут дать нереалистичные значения параметров и, следовательно, ненадежные прогнозы модели [29-30].

Сетевые модели. В современной математической эпидемиологии сетевые модели представляют собой один из новейших методов анализа и моделирования сложных эпидемиологических систем. По сравнению с камерными моделями они являются более детальными и позволяют рассматривать каждого участника в отдельности, при этом взаимодействие между людьми представляется в виде сложного графа социальных связей. В такие модели можно включать различную степень детализации: разбивать популяцию на возрастные группы, учитывать количество социальных связей, наличие мест с большей плотностью населения и эффективность ограничительных мер. В то время как камерные модели позволяют оценить общую динамику распространения эпидемии, сетевые модели дают возможность симитировать эффективность тех или иных мер по сдерживанию распространения инфекции [31].

Модели с использованием методов машинного обучения. Искусственный интеллект и машинное обучение давно используются в эпидемиологии. Это мощный инструмент для поиска взаимосвязи между входными и выходными данными в случаях, когда аналитическое исследование затруднительно. Применение таких эвристических подходов для раннего обнаружения эпидемиологических рисков в некоторых случаях позволяет улучшить качество прогнозирования. За последние годы в методах машинного обучения было достигнуто много

успехов. Важным представителем моделей на основе машинного обучения являются искусственные нейронные сети (ИНС). Они представляют собой модель искусственного интеллекта, «вдохновленную» структурой биологических нейронов человека. В случае прогнозирования распространения эпидемии обучение ИНС заключается в вычислении коэффициентов связей между входными сигналами, определяющих их силу, и выполняется на основе эмпирических данных, включающих статистику заболеваемости и значения различных факторов, способных оказывать влияние на распространение заболеваемости. В исследовании [32] рассматриваются возможности применения нейронных сетей для прогнозирования распространения COVID-19. Результаты работы сети разработанной архитектуры для некоторых регионов достигали 87%. При этом, авторы акцентируют внимание, что для корректного обучения ИНС необходим большой объем исторических данных. Также следует упомянуть и подход к моделированию на основе прецедентов (англ. case-based reasoning). Данный подход основывается на идее поиска возможных решений задачи на основе уже существующих решений для аналогичных ситуаций [33]. В работах [34, 35] описана новая модель CBRR (case-based rate reasoning) на базе данного подхода для прогнозирования будущих значений основных параметров эпидемии коронавируса в России, позволяющая строить краткосрочные прогнозы на основе аналогов динамики процентного прироста в других странах. Также описан новый эвристический метод оценивания продолжительности переходного процесса процентного прироста между заданными уровнями, учитывающий информацию о динамике эпидемиологических процессов в странах цепочки распространения. Подробный обзор возможностей использования методов машинного обучения для прогнозирования распространения COVID-19 можно найти в [36].

В целом, каждый подход имеет свои достоинства и недостатки. При этом, в случае попыток прогнозирования вновь возникающих эпидемий, основной проблемой является отсутствие исторических данных, на основании которых можно было бы оценить значения входных параметров моделей. Другой общей проблемой являются «грязные» исходные данные. Информация берется из разных источников, а разные регионы и страны используют отличные друг от друга методики учета новых случаев заражения, смерти, тестирования и т.д. Учитывая тот факт, что для вновь возникающих инфекционных заболеваний отсутствует накопленная за многие годы статистика, данное ограничение является критичным при определении ключевых параметров моделей. В качестве примера можно привести различные оцен-

ки базового репродуктивного числа для нового коронавирусной инфекции [37-39].

3. Модель SIR и её представление. Рассмотрим детерминированную систему дифференциальных уравнений, традиционно используемую для моделирования эпидемий, известную как модель SIR [6]:

$$\frac{dS(t)}{dt} = -\beta \frac{S(t)I(t)}{N}, \quad (1)$$

$$\frac{dI(t)}{dt} = \beta \frac{S(t)I(t)}{N} - \gamma I(t), \quad (2)$$

$$\frac{dR(t)}{dt} = \gamma I(t), \quad (3)$$

$$N = I(t) + S(t) + R(t). \quad (4)$$

В этой модели переменная S принимает значения, равные численности восприимчивого к вирусу населения популяции, I – количества инфицированных людей, R – общего количества выздоровевших и умерших людей. При этом численность популяции N считается неизменной в течение всего эпидемиологического процесса. Коэффициенты β и γ в модели, как правило, являются неотрицательными величинами, не изменяющимися во времени. Значения переменных в начальный момент времени t_0 заданы.

Выбор значений коэффициентов β и γ при использовании этой модели для прогнозирования динамики эпидемии конкретного типа вируса осуществляется на основании имеющейся статистики прошлых периодов. Построенные для выбранных значений коэффициентов траектории такой системы являются, как правило, детерминированными функциями времени, которые, как показывают многочисленные примеры из научной литературы, часто являются достаточно хорошими приближениями фактических траекторий, развивающихся в реальном времени. Вместе с тем, на точность оценки влияют многочисленные факторы, такие, например, как качество имеющейся статистики и эволюционная изменчивость самого вируса и его различных штаммов. В случае появления нового вируса подходящей статистики нет. Неопределенность параметров модели его распространения становится существенной преградой для построения прогнозов динамики эпидемиологической ситуации. Коэффициенты β и γ фактически являются случайными величинами. Функции распределения этих случайных величин не известны заранее.

Естественно предположить, что количество восприимчивых к новому вирусу людей в начале эпидемии близко к значению численности популяции N , то есть $S \approx N$. Тогда, с учетом близости значений S и N , получим систему дифференциальных уравнений типа SIR вида:

$$\frac{dI(t)}{dt} = (\beta - \gamma)I(t), \quad (2')$$

$$\frac{dR(t)}{dt} = \gamma I(t). \quad (3')$$

Данная система уравнений описывает динамику количества инфицированных $I(t)$ и общего количества $R(t)$ выздоровевших и умерших людей на начальном этапе эпидемии, когда количество невосприимчивых к вирусу людей относительно мало. Разность $\beta - \gamma$ является показателем экспоненциального роста функции $I(t)$, а отношение $R_0 = \frac{\beta}{\gamma}$ называется базовым репродуктивным числом (индексом) вируса. В начале эпидемии значение $R_0 > 1$.

Решение системы (2')-(3') имеет простой аналитический вид

$$I(t) = I(t_0)e^{(\beta - \gamma)(t - t_0)},$$

$$R(t) = \frac{\gamma I(t_0)}{\beta - \gamma} \left(e^{(\beta - \gamma)(t - t_0)} - 1 \right) + R(t_0),$$

в том числе и для дискретных значений времени $t = 0, 1, 2, \dots$.

Рассмотрим вопрос об оценке точности приближений фактической динамики эпидемии с помощью дискретных решений системы (2')-(3). Предположим, что информация мониторинговой системы здравоохранения обновляется ежедневно (то есть $t = 0, 1, 2, \dots$) как, например, это происходит на странице стопкоронавирус.рф/information/ оперативных данных сайта стопкоронавирус.рф оперативного штаба Правительства РФ. После начала эпидемии мониторинговыми системами здравоохранения ежедневно регистрируются общее количество выявленных случаев инфицирования, количество новых случаев заболевания Δ , количество инфицированных людей

(число активных случаев) I , количество умерших и выздоровевших людей R .

Рассмотрим данные о числе активных случаев и общем количестве выздоровевших и умерших в России в период с 20 апреля по 19 мая 2020 года и внесем их в таблицу 1 (составлена по данным ежедневных Отчетов Коммуникационного центра Правительства РФ по ситуации с коронавирусом, размещенным в открытом доступе на портале стопкоронавирус.рф.). Выберем такие постоянные значения коэффициентов β и γ , при которых дискретная траектория системы (2')-(3') наиболее точно соответствует фактическим данным, имеющимся в этой таблице.

Таблица 1. Данные оперативного штаба и отклонение от расчетных данных модели SIR (период с 20 апреля по 19 мая 2020 года)*

t	Дата	Данные $I(t)$	Данные $I(t)$ модели SIR	Отклонение	Данные $R(t)$	Данные $R(t)$ модели SIR	Отклонение
1	20 апр	43270	43270	0	3851		
2	21 апр	48434	46332	-2102	4329	2545	-1784
3	22 апр	53066	49612	-3454	4933	5271	338
4	23 апр	57327	53122	-4205	5446	8189	2743
5	24 апр	62421	56881	-5540	6201	11314	5113
6	25 апр	67657	60906	-6751	6931	14660	7729
7	26 апр	73435	65215	-8220	7514	18243	10729
8	27 апр	79007	69828	-9179	8140	22079	13939
9	28 апр	84235	74767	-9468	9323	26186	16863
10	29 апр	88138	80055	-8083	11261	30584	19323
11	30 апр	93806	85716	-8090	12692	35293	22601
12	01 май	100042	91777	-8265	14389	40336	25947
13	02 мая	107819	98265	-9554	16235	45734	29499
14	03 мая	116768	105211	-11557	17919	51515	33596
15	04 мая	125817	112647	-13170	19451	57703	38252
16	05 мая	134054	120607	-13447	21316	64330	43014
17	06 мая	143065	129128	-13937	22864	71424	48560
18	07 мая	151732	138249	-13483	25428	79020	53592
19	08 мая	159528	148012	-11516	28331	87152	58821
20	09 мая	164933	158462	-6471	33743	95859	62116
21	10 мая	173467	169647	-3820	36221	105180	68959
22	11 мая	179534	181619	2085	41810	115159	73349
23	12 мая	186615	194432	7817	45628	125843	80215
24	13 мая	192056	208144	16088	50215	137280	87065
25	14 мая	196410	222819	26409	55835	149524	93689

26	15 мая	202199	238524	36325	60644	162631	101987
27	16 мая	206304	255329	49025	65739	176662	110923
28	17 мая	211748	273311	61563	70004	191681	121677
29	18 мая	217747	292551	74804	72931	207758	134827
30	19 мая	220974	313136	92162	78967	224967	146000

При выборе значений коэффициентов β и γ будем исходить из следующих общих оценок:

- в качестве оценки разности будем использовать линейную аппроксимацию для $\ln I = (\beta - \gamma)t + \ln I_0$;
- в качестве оценки параметра, отвечающего за скорость выздоровления, можно взять $\frac{1}{T}$, где T – среднее время перехода в состояние R . Данная оценка получается, исходя из предположения о неизменности значений коэффициентов.

В итоге для построения модели SIR были взяты следующие параметры: $T = 17$, $\gamma = 0.06$, $\beta = 0.13$. Данные значения согласуются со значениями, используемыми другими исследователями при моделировании динамики распространения COVID-19 [40]. Результаты расчётов модели SIR представлены на рисунке 4 и в таблице 1. Можно видеть, что в течение 20 дней, с 20 апреля по 9 мая, модель достаточно хорошо отражает динамику распространения эпидемии. В последние 10 дней наблюдается нарастающее расхождение между расчетными данными и фактическими. Средняя абсолютная ошибка в процентах (mean absolute percentage error, MAPE) расчетной траектории числа активных случаев заболевания на интервале с 20 апреля по 19 мая 2020 года составила 11%.

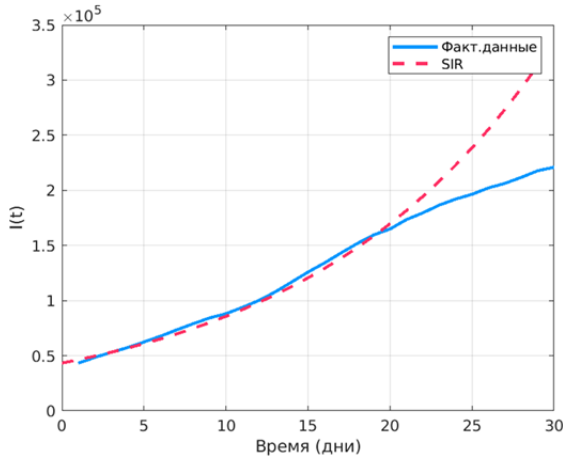


Рис. 4. Фактическая траектория числа активных случаев заболевания и расчетная траектория по модели SIR

4. Балансовая модель эпидемии на основе процентного прироста. Обозначим через $C(t)$ общее число подтвержденных случаев инфицирования (Cumulative Cases, Confirmed Cases) с момента начала эпидемии до дня t включительно, а $\Delta(t)$ – количество новых случаев, зарегистрированных в день t . С учетом введенных обозначений для $t = 1, 2, 3, \dots$ имеем

$$C(t) = C(t-1) + \Delta(t).$$

Запишем $C(t)$ в виде

$$C(t) = \alpha(t)C(t-1),$$

где

$$\begin{aligned} \alpha(t) &= 1 + \frac{r(t)}{100}, \\ r(t) &= 100 \frac{\Delta(t)}{C(t-1)}, \\ \Delta(t) &= C(t) - C(t-1). \end{aligned} \tag{5}$$

То есть окончательно имеем следующую формулу

$$C(t) = \left(1 + \frac{r(t)}{100}\right) \times C(t-1). \quad (6)$$

Здесь параметр $r(t)$ интерпретируется как отношение в процентах абсолютного прироста общего количества выявленных случаев инфицирования до дня t включительно к общему количеству выявленных случаев инфицирования до предыдущего дня включительно.

Будем далее называть параметр $r(t)$ *процентным приростом* общего количества выявленных случаев инфицирования в день t . Учитывая, что $\Delta(t)$ является случайной величиной, принимающей неотрицательные значения, процентный прирост также является неотрицательной случайной величиной.

Значение величины $C(t)$ в любой день T вычисляется по следующей формуле

$$C(T) = \prod_{t=1}^T \left(1 + \frac{r(t)}{100}\right). \quad (7)$$

Зафиксируем некоторые значения $t_0 \geq 0$ и $T > t_0$, такие, что $R(T) > C(t_0) > 1$. Учитывая не убывание функций $C(t)$ и $R(t)$, а также то, что $C(t) \geq R(t)$ при любом $t \geq t_0$, такое значение T существует. Фактически существование таких значений t_0 и T означает, что пациенты, заболевшие к моменту времени t_0 , выздоравливают или умирают за конечное время.

Рассмотрим следующую задачу:

$$\min_{t_0 \leq t \leq T} t, \quad (8)$$

$$C(t) \geq R(T). \quad (9)$$

Учитывая свойства функций $C(t)$ и $R(t)$ множество допустимых решений такой задачи не пусто. Обозначим решение задачи (8)-(9) через $\tau(T)$.

Обратим внимание на то, что для $t = \tau(T) - 1$ условие (9) не выполняется. Тогда с учетом не убывания функции $C(t)$, очевидно, выполняется неравенство $R(T) \geq C(\tau(T) - 1)$. Таким образом, справедлива следующая теорема.

Теорема 1 (Принцип динамического баланса). Пусть заданы значения $t_0 \geq 0$ и $T > t_0$, такие, что $R(T) > C(t_0) > 1$. Тогда для решения $\tau(T)$ задачи (8)-(9) значение $R(T)$ в день T удовлетворяет неравенствам

$$C(\tau(T)) \geq R(T) \geq C(\tau(T) - 1). \quad (10)$$

Следствие. В условиях теоремы 1 функция $R(T)$ представима в виде

$$R(T) = \lambda_r C(\tau(T) - 1) + (1 - \lambda_r) C(\tau(T)), \quad \lambda_r \in [0, 1]. \quad (11)$$

Его справедливость следует из известной в аналитической геометрии формулы представления точек, лежащих на отрезке.

Условие (10) означает, что количество выздоровевших и умерших людей в определенный день t зависит от общего количества зафиксированных заболевших в прошлом, а именно $t - \tau(t)$ дней назад. Таким образом, с помощью условия (10) можно установить динамический баланс между значениями функций $R(t)$, $C(\tau(t) - 1)$ и $C(\tau(t))$.

Заметим, что для любого значения t справедливо равенство

$$C(t) = I(t) + R(t). \quad (12)$$

Это балансовое соотношение означает, что группу выявленных заболевших в любой день можно разделить на тех, кто еще болеет, и тех, кто к этому дню выздоровел или, к сожалению, умер. Воспользуемся формулами (6), (11) и (12) и запишем систему дискретных уравнений

$$C(t) = \left(1 + \frac{r(t)}{100}\right) \times C(t-1), \quad (13)$$

$$I(t) = \left(1 + \frac{r(t)}{100}\right) \times C(t-1) - R(t), \quad (14)$$

$$R(t) = \lambda_i C(\tau(t)-1) + (1 - \lambda_i) C(\tau(t)). \quad (15)$$

Определение 1. Модель, динамика которой описывается системой дискретных уравнений (13)-(15), будем в дальнейшем сокращенно называть *CIR-модель*.

CIR-модель использует понятие процентного прироста общего количества заболевших $C(t)$ и учитывает балансовое соотношение между общим количеством заболевших в прошлые моменты времени $\tau(t)$ и $\tau(t)-1$ и общим количеством выздоровевших и умерших $R(t)$ в момент t . В связи с этим мы можем также называть эту модель *балансовой моделью эпидемии на основе процентного прироста*.

Определение 2. Функцию $\vartheta(t) = t - \tau(t)$ будем называть *характеристикой динамического баланса эпидемии*.

Исследуя статистику пандемии коронавирусной инфекции COVID-19 в России и других странах можно отметить, что на отдельных промежутках времени характеристика динамического баланса $\vartheta(t)$ является константой, однако, в общем случае она меняется во времени, хотя ее волатильность ограничена.

Определение 3. Будем называть эпидемиологический процесс *стационарным* на промежутке времени $[t_1, t_2]$, если $\vartheta(t) = const$ для всех $t \in [t_1, t_2]$.

Рассмотрим данные эпидемиологического процесса из табл. 1. Оценим, каковы на промежутке $[23, 24, \dots, 30]$ значения функций $\tau(t)$, $\vartheta(t)$, интервалы возможных значений функции $R(t)$, а также отклонения $R(t)$ от этих интервалов. На промежутке $[1, 2, \dots, 22]$ по данным таблицы 1 определить эти значения не представляется возможным.

По имеющимся в столбцах 3 и 4 таблицы 2 данным можно определить моменты времени, соответствующие допустимым решениям задачи (8)-(9) в период с 12 по 19 мая, то есть для $T = 23, 24, \dots, 30$ соответственно (для значений, меньших 23, множество допустимых решений является пустым). Эти решения представлены в столбце 5 таблицы 2, в столбце 6 – вычисленные значения характеристики динамического баланса $\vartheta(t)$, в столбце 7 – оценка интервалов будущих возможных значений $R(T)$ при постоянной характеристике динамического баланса, в столбце 8 – величина отклонения фактического значения $R(T)$ от интервала оценки. Заметим, что в период с 12 по 19

мая функция $\mathcal{G}(t)$ не равна константе, то есть рассматриваемый эпидемиологический процесс в этот период времени не является стационарным.

Заметим, что по имеющимся в таблице данным интервальная оценка будущих значений общего количества выздоровевших и умерших в дни с 12 по 19 мая может быть получена при известных заранее значениях характеристики динамического баланса $\mathcal{G}(t)$ за 22 или 23 дня до этих дат. Увеличение значения функции $\mathcal{G}(t) = t - \tau(t)$ 18 и 19 мая на один день по отношению к периоду с 12 по 17 мая, когда она была постоянна (22 дня), привело к отклонению фактических значений $R(18)$ и $R(19)$ на 1697 и 1982 соответственно, или примерно на 30% от прироста новых случаев заболевания 20 апреля.

Таким образом, условие динамического баланса можно использовать для прогнозирования количества активных случаев на промежуток, продолжительность которого зависит от значения характеристики динамического баланса.

Таблица 2. Значения характеристики динамического баланса и оценки (период с 20 апреля по 19 мая 2020 года)

t	Дата	Общее число заболевших $C(t)$	Всего выздоровели или умерли $R(t)$	$\tau(t)$	$\mathcal{G}(t)$	Интервалы возможных значений функции $R(t)$	Отклонение $R(t)$ от интервала
1	20 апр	47121	3851				
2	21 апр	52763	4329				
3	22 апр	57999	4933				
4	23 апр	62773	5446				
5	24 апр	68622	6201				
6	25 апр	74588	6931				
7	26 апр	80949	7514				
8	27 апр	87147	8140				
9	28 апр	93558	9323				
10	29 апр	99399	11261				
11	30 апр	106498	12692				
12	01 мая	114431	14389				
13	02 мая	124054	16235				
14	03 мая	134687	17919				
15	04 мая	145268	19451				
16	05 мая	155370	21316				
17	06 мая	165929	22864				

18	07 мая	177160	25428				
19	08 мая	187859	28331				
20	09 мая	198676	33743				
21	10 мая	209688	36221				
22	11 мая	221344	41810				
23	12 мая	232243	45628	1	22	≤ 47121	0
24	13 мая	242271	50215	2	22	[47121;52763]	0
25	14 мая	252245	55835	3	22	[52763;57999]	0
26	15 мая	262843	60644	4	22	[57999;62773]	0
27	16 мая	272043	65739	5	22	[62773;68622]	0
28	17 мая	281752	70004	6	22	[68622; 74588]	0
29	18 мая	290678	72931	6	23	[74588;80949]	1657
30	19 мая	299941	78967	7	23	[80949;87147]	1982

В нашем примере, при имеющихся данных, 20 апреля можно сделать относительно точный прогноз на период с 12 мая по 19 мая. Для получения прогноза для периода с 21 апреля до 19 мая требуется дополнительно учитывать более ранние данные. Можно также отметить, что если бы эпидемиологический процесс на промежутке с 12 по 19 мая был стационарным, точность прогноза была бы выше.

5. CIR-модель и метод прецедентов для прогнозирования процентного прироста. В модели CBRR (Case-Based Rate Reasoning), подробно рассмотренной в статьях [34, 35] мы предложили новый подход для прогнозирования динамики показателя $C(t)$, основанный на методе прецедентов. Он предусматривает использование процедуры эвристического выбора будущих значений процентного прироста общего количества выявленных случаев инфицирования в России на основе информации о динамических процессах развития пандемии в других странах. В соответствии с этим алгоритмом была сформирована цепочка распространения эпидемии, в которую вошли Россия (страна-последователь), и страны-предшественники Италия, Испания, Великобритания и Франция. На рисунке 5 представлены графики процентного прироста заболевших в этих странах после выхода на уровень общего количества заболевших, равный пять тысяч.

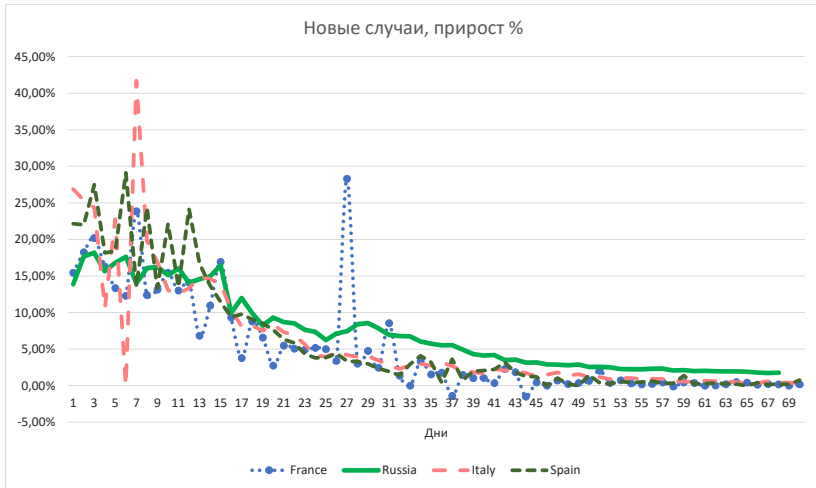


Рис. 5. Динамика процентного прироста новых случаев заболевания COVID-19 во Франции, Италии, Испании и России

Общая процедура формирования прогноза процентного прироста при использовании СВRR-модели заключается в следующем. Рассматриваемый временной горизонт развития эпидемии разбивается на M промежутков $(T_{m-1}, T_m]$, $0 \leq T_{m-1} < T_m \leq T$, $m = 1, 2, \dots, M$. Границы промежутков соответствуют прогнозируемым моментам времени прохождения процентным приростом заданных значений (например, 10, 5, 2, 1%). Количество дней в промежутке $(T_{m-1}, T_m]$ равно $T_m - T_{m-1}$. Динамику изменения прогнозируемых значений общего количества инфицированных внутри каждого промежутка можно описать, например, следующим образом:

$$\tilde{C}(T_{m-1} + k) = \left(1 + \frac{r(T_{m-1} + k)}{100}\right) \tilde{C}(T_{m-1} + k - 1), \quad k = 1, 2, \dots, (T_m - T_{m-1}),$$

а для вычисления общего количества инфицированных в конце промежутка $(T_{m-1}, T_m]$ воспользуемся формулой

$$\tilde{C}(T_m) = \left(1 + \frac{r(T_{m-1} + 1)}{100}\right) \left(1 + \frac{r(T_{m-1} + 2)}{100}\right) \dots \left(1 + \frac{r(T_m)}{100}\right) \tilde{C}(T_{m-1}).$$

На рисунках 6 и 7 представлены построенные на промежутке от 23.04.2020 до 30.06.2020 спрогнозированные траектории, которые сравниваются с фактическими траекториями. Среднее значение отклонения траектории прогноза общего количества заболевших от факти-

ческой составило 0,37%. 30 июня отклонение оказалось равным минус 1,17%.

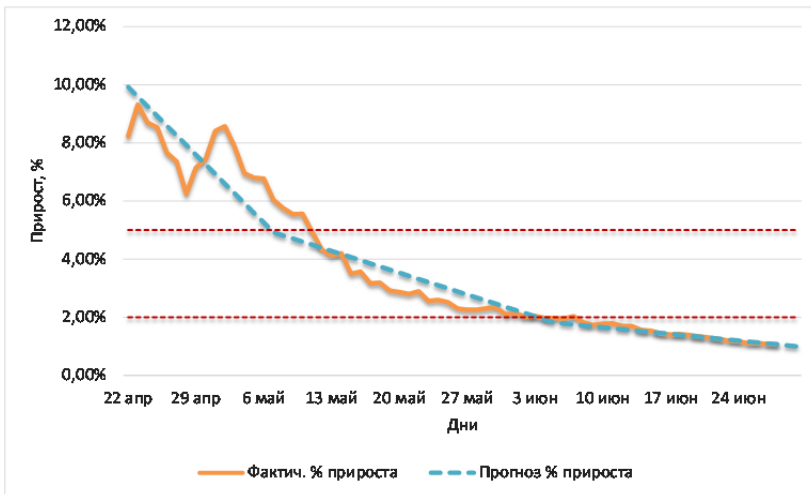


Рис. 6. Процентный прирост общего числа случаев заболевания

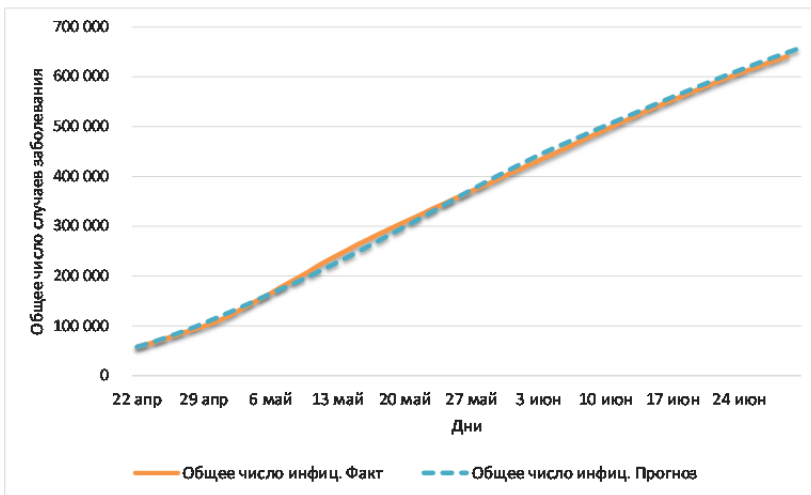


Рис. 7. Динамика общего числа зарегистрированных случаев заболевания – фактическая и расчётная траектории

Рассмотрим результаты применения разработанного подхода к построению прогноза показателей эпидемии в России на промежутке с

6 июня по 30 июня 2020 года на основании имеющейся 5 июня информации. При предположении о стационарности эпидемиологического процесса с 6 по 30 июня мы взяли для расчета постоянное значение характеристики динамического баланса, равное 25 дней. Спрогнозированные значения общего количества выздоровевших и умерших $R(t)$ и активных случаев $I(t)$ приведены, соответственно, во втором и в четвертом столбцах таблицы 3, а их фактически реализовавшиеся значения – в столбцах 3 и 5. Отклонение расчётной траектории от фактической траектории числа активных больных составило, в среднем, 3,63%. На пике эпидемии по количеству активных случаев (15 июня) отклонение достигало 6,38%. Значение MAPE для данного прогноза составило 4,14%.

Для динамики общего количества выздоровевших и умерших ошибка прогноза MAPE на рассматриваемом интервале составила 4,69%. При этом максимальное отклонение составляло 7,3%. Графики расчётных и фактических значений этих показателей приведены на рисунках 8 и 9.

Предположим, что значения характеристики динамического баланса $\vartheta(t)$ на промежутке прогнозирования нам известны заранее. Эти значения внесены в столбец 6 таблицы 3. График функции $\vartheta(t)$ представлен на рисунке 10.

Таблица 3. Прогноз от 5 июня 2020 года и фактические значения $R(t)$ и $I(t)$. Значения характеристики динамического баланса и оценки отклонения (период с 06 июня по 30 июня 2020 года)

Дата	Прогноз $R(t)$	Факт $R(t)$	Прогноз $I(t)$	Факт $I(t)$	$\vartheta(t)$	Отклонение $R(t)$	Отклонение $I(t)$
06 июн	226794	227113	241750	231576	25	-320	10174
07 июн	237257	232590	239645	235083	25	4667	4562
08 июн	247258	236659	237995	239999	26	10599	-2004
09 июн	257544	248539	236049	236714	26	9005	-665
10 июн	267443	259141	234471	234516	26	8302	-45
11 июн	276898	267682	233315	234754	26	9216	-1439
12 июн	286215	276085	232266	235338	26	10130	-3072
13 июн	295310	281470	231405	238659	27	13840	-7254
14 июн	304323	286998	230584	241966	27	17325	-11382
15 июн	313130	291630	229922	245580	27	21500	-15658
16 июн	322001	301590	229143	243868	27	20411	-14725
17 июн	331165	311820	228011	241481	27	19345	-13470

18 ИЮН	340182	321623	226962	239468	27	18559	-12506
19 ИЮН	348954	332333	226085	236730	27	16621	-10645
20 ИЮН	357885	342594	224972	234358	27	15291	-9386
21 ИЮН	366511	347822	224080	236858	27	18689	-12778
22 ИЮН	374866	352622	223369	239658	28	22244	-16289
23 ИЮН	383337	364788	222445	234917	27	18549	-12472
24 ИЮН	392099	377335	221127	229546	27	14764	-8419
25 ИЮН	401209	383769	219353	230225	27	17440	-10872
26 ИЮН	410361	392933	217423	227861	27	17428	-10438
27 ИЮН	419310	402321	215575	225325	27	16989	-9750
28 ИЮН	428009	408160	213849	226277	27	19849	-12428
29 ИЮН	436693	412596	212006	228560	28	24097	-16554
30 ИЮН	445471	421979	209929	225870	28	23492	-15941

Если использовать фактические значения $\vartheta(t)$ из таблицы 3, то, как видно на рисунке 11, графики фактических и спрогнозированных значений $R(t)$ практически совпадают. Аналогично на рисунке 12 практически совпадают графики общего количества активных больных. На интервале с 6 по 30 июня отклонение от фактической траектории числа активных больных составило в среднем 1,13%. На пике эпидемии (15 июня) отклонение составляет 1,01%.

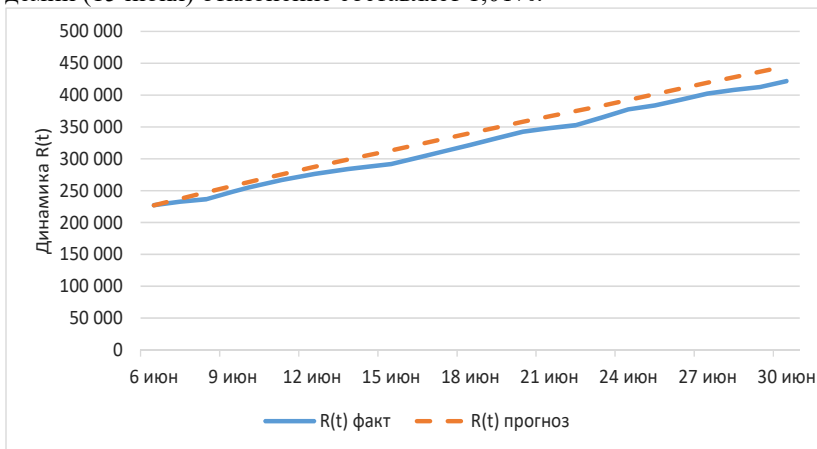


Рис. 8. Динамика $R(t)$ – фактическая и расчётная траектории на интервале с 6 по 30 июня 2020 года

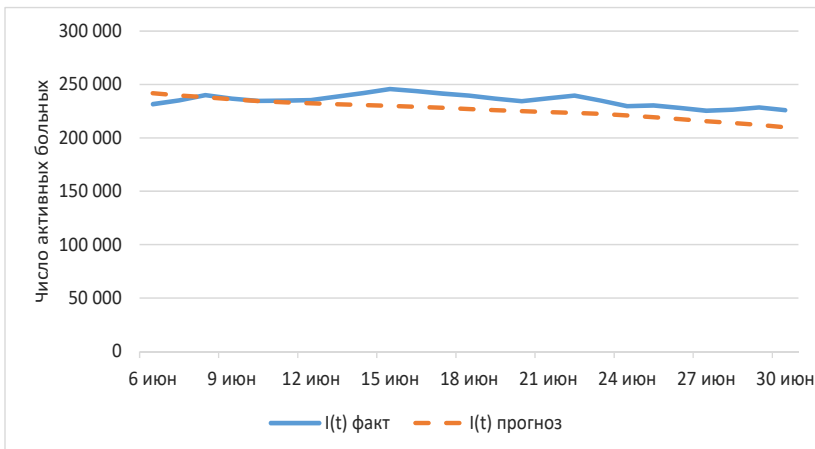


Рис. 9. Динамика $I(t)$ – фактическая и расчётная траектории на интервале с 6 по 30 июня 2020 года

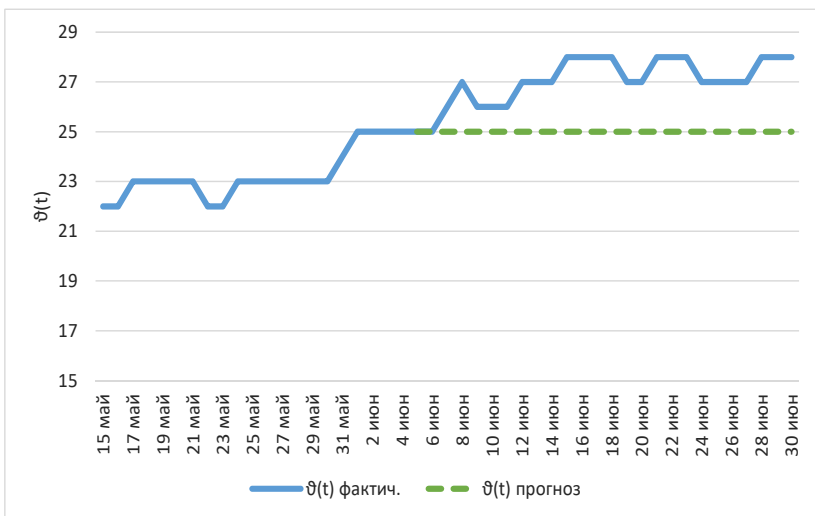


Рис. 10. Фактические и модельные значения характеристики динамического баланса $\vartheta(t)$

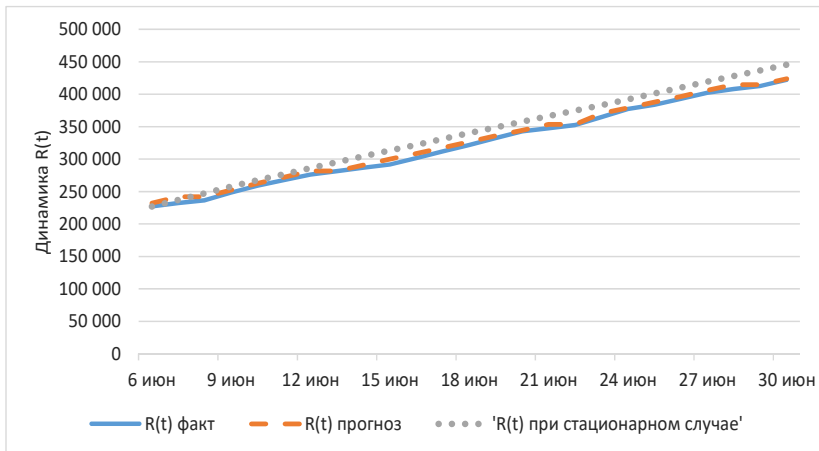


Рис. 11. Динамика $R(t)$ – фактическая и расчётная траектории на интервале с 6 по 30 июня 2020 года при известных заранее значениях $g(t)$

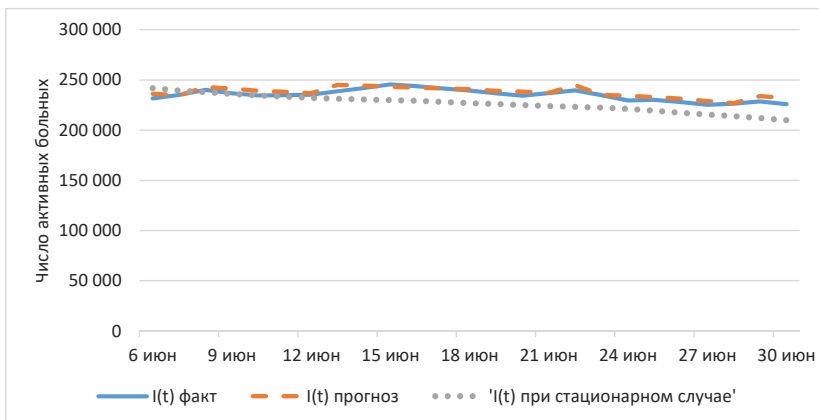


Рис. 12. Динамика $I(t)$ – фактическая и расчётная траектории на интервале с 6 по 30 июня 2020 года при известных заранее значениях $g(t)$

6. Заключение. В статье представлена новая дискретная SIR-модель распространения эпидемии, описывающая с использованием случайных значений процентного прироста и характеристики динамического баланса динамику общего количества заболевших, общего количества выздоровевших и умерших и количества активных случаев болезни. Для прогнозирования значений общего количества заболе-

ших применяется CBRR-модель, ранее предложенная и апробированная авторами. Она сначала прогнозирует на среднесрочном промежутке времени значения процентного прироста общего количества заболевших с использованием метода прецедентов и интеллектуальных алгоритмов извлечения необходимых данных с использованием специального класса модели PDRM (predictive dynamic regression model) [34], а затем использует полученные значения для построения прогноза общего количества случаев заражения. Горизонт прогнозирования CIR-модели, как правило, ограничен величиной, равной значению характеристики динамического баланса эпидемии, вычисляемой на основании принципа динамического баланса эпидемиологического процесса, сформулированного в теореме 1. Поскольку характеристика динамического баланса является случайной величиной, зависящей от времени и принимающей целочисленные значения, возможности использования предложенной модели для прогнозирования динамики нестационарных эпидемиологических процессов пока представляются ограниченными. Однако, как показывают проведенные расчеты, точность прогнозирования при применении предложенной модели даже в случае нестационарности эпидемиологических процессов на промежутке прогнозирования достаточно высока. Так, отклонение расчётной траектории от фактической траектории числа активных больных составило, в среднем, 3,63%. На пике эпидемии по количеству активных случаев (15 июня) отклонение достигало 6,38%. Значение MAPE для данного прогноза составило 4,14%. В случае использования фактических значений $\mathcal{I}(t)$ отклонение траектории числа активных больных сокращалось до 1,13%. Для динамики общего количества выздоровевших и умерших ошибка прогноза MAPE на рассматриваемом интервале с 6 по 30 июня 2020 г. составила 4,69%.

Литература

1. Foppa I.M. A historical introduction to mathematical modeling of infectious diseases. Academic Press. 2017.
2. Shinde G.R., Kalamkar, A.B., Mahalle P.N., et al. Forecasting Models for Coronavirus (COVID-19): A Survey of the State-of-the-Art. SN Computer Science. 2020. vol. 1.
3. Moftakhar L., Seif M., Safe M.S. Exponentially increasing trend of infected patients with COVID-19 in Iran: a comparison of neural network and ARIMA forecasting models. Iran Journal of Public Health. vol. 49. pp. 92–100.
4. Chaudhry R.M. et al. Coronavirus disease 2019 (COVID-19): forecast of an emerging urgency in Pakistan. Cureus. 2020. vol. 12. no 5.
5. Tandon H., Ranjan P., Chakraborty T., Suhag V. Coronavirus (covid-19): Arima based time-series analysis to forecast near future. arXiv:2004.07859.

6. Kermack W.O., McKendrick A.G. A contribution to the mathematical theory of epidemics. Proceedings of the Royal Society of London. Series A. 1927. vol. 115. no. 772. pp. 700–721.
7. Anderson R.M., May R.M. Infectious diseases of humans: Dynamics and control. Oxford: Oxford University Press. 1991. P. 757
8. Dil S., Dil N., Maken Z.H. COVID-19 trends and forecast in the eastern mediterranean region with a particular focus on Pakistan. Cureus. 2020. vol. 12. no 6.
9. Johns Hopkins Coronavirus Resource Center. Available at: <https://coronavirus.jhu.edu/data> (accessed: 20.04.2021)
10. Liao Z., Lan P., Liao Z. et al. TW-SIR: time-window based SIR for COVID-19 forecasts. Sci Rep. 2020. vol. 10.
11. Rădulescu A., Williams C., Cavanagh K. Management strategies in a SEIR-type model of COVID 19 community spread. Sci Rep. 2020. vol. 10.
12. Fanelli D., Piazza F. Analysis and forecast of COVID-19 spreading in China, Italy and France. Chaos Solitons Fractals. 2020. vol. 134.
13. Miller A., et al. Correlation between universal BCG vaccination policy and reduced morbidity and mortality for COVID-19: an epidemiological study. medRxiv 2020.03.24.20042937.
14. Cheng Z., et al. Icumonitoring.ch: a platform for short-term forecasting of intensive care unit occupancy during the COVID-19 epidemic in Switzerland. Swiss Medical Weekly. 2020. vol. 150.
15. Rodrigues H.S. Application of SIR epidemiological model: new trends. International Journal of Applied Mathematics and Informatics. 2016. vol. 10. pp. 92-97.
16. Iwami S., Takeuchi Y., Liu X. Avian–human influenza epidemic model. Mathematical biosciences. 2007. vol. 207. no. 1. pp. 1–25.
17. Teles P. Predicting the evolution of SARS-COVID-2 in Portugal using an adapted SIR model previously used in South Korea for the MERS outbreak. medRxiv 2020.03.18.20038612.
18. Maier B.F., Brockmann D. Effective containment explains subexponential growth in recent confirmed COVID-19 cases in China. Science. 2020. vol. 368. no. 6492. pp. 742-746.
19. Chinazzi M., Davis J.T., et al. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. Science. 2020. vol. 368. no. 6489. pp. 395-400.
20. Tang B., Wang X., Li Q., Bragazzi N.L., Tang S., Xiao Y., et al. Estimation of the transmission risk of 2019-nCov and its implication for public health interventions. Journal of Clinical Medicine. 2020. vol. 9. no. 2.
21. Tian H., Liu Y., Li Y. An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. Science. 2020. vol. 368. no. 6491. pp. 638-642.
22. López L., Rodó X. A modified SEIR model to predict the COVID-19 outbreak in Spain and Italy: Simulating control scenarios and multi-scale epidemics. Results in Physics. 2021. vol. 21.
23. Feng S., Feng Z., Ling C., Chang C., Feng Z. Prediction of the COVID-19 epidemic trends based on SEIR and AI models. PLoS ONE. 2021. vol. 16. no. 1.
24. Криворотько О.И., Кабанихин С.И., Зятков Н.Ю., Приходько А.Ю., Прохошин Н.М., Шишленин М.А. Математическое моделирование и прогнозирование COVID-19 в Москве и Новосибирской области. URL: <https://arxiv.org/abs/2006.12619v1> (дата обращения: 30.04.2021)
25. Матвеев А.В. Математическое моделирование оценки эффективности мер против распространения эпидемии COVID-19. Национальная безопасность и стратегическое планирование. 2020. №. 1. С. 23-39.

26. Anastasopoulou C., Russo L., Tsakris A., Siettos C. Data-based analysis, modelling and forecasting of the COVID-19 outbreak. *PLoS One*. 2020. vol. 15. no. 3.
27. Mandal S., Bhatnagar T., Arinaminpathy N. Prudent public health intervention strategies to control the coronavirus disease 2019 transmission in India A mathematical model-based approach. *Indian Council of Medical Research*. 2020. vol. 151. pp. 190-199.
28. Choi S., Ki M. Estimating the reproductive number and the outbreak size of COVID-19 in Korea. *Epidemiology and Health*. 2020. vol. 42.
29. Tolles J., Luong T. Modeling Epidemics with Compartmental Models. *JAMA*. 2020. vol. 323. no. 24. pp. 2515–2516.
30. Fudolig M., Howard R. The local stability of a modified multi-strain SIR model for emerging viral strains. *PLoS ONE*. 2020. vol. 15. no. 12.
31. Adam D. Special report: The simulations driving the world’s response to COVID-19. *Nature*. 2020. vol. 580. pp. 316-318.
32. Wieczorek M., Siłka J., Woźniak M. Neural network powered COVID-19 spread forecasting model. *Chaos, Solitons & Fractals*. 2020. vol. 140.
33. Кондратьев М.А. Методы прогнозирования и модели распространения заболеваний. Компьютерные исследования и моделирование. 2013. т. 5. № 5 с. 863-882.
34. Zakharov V., Balykina Y., Petrosian O., Gao H. CBRR Model for Predicting the Dynamics of the COVID-19 Epidemic in Real Time. *Mathematics*. 2020. vol. 8. no. 10.
35. Захаров В.В., Балькина Ю.Е. Прогнозирование динамики эпидемии коронавируса (COVID-19) на основе метода прецедентов. Вестник Санкт-Петербургского университета. Прикладная математика. Информатика. Процессы управления. 2020. т. 16. Вып. 3. С. 249-259.
36. Dairi A., Harrou F., Zeroual A., Hittawe M.M., Sun Y. Comparative study of machine learning methods for COVID-19 transmission forecasting. *Journal of Biomedical Informatics*. 2021. vol. 18.
37. Mizumoto K., Chowell G. Transmission potential of the novel coronavirus (COVID-19) onboard the diamond Princess Cruises Ship, 2020. *Infectious Disease Modelling*. 2020. vol. 5. pp. 264-270.
38. Zhang S., et al. Estimation of the reproductive number of novel coronavirus (COVID-19) and the probable outbreak size on the Diamond Princess cruise ship: A data-driven analysis. *Int J Infect Dis*. 2020. vol. 93. pp. 201-204.
39. Jung S., et al. Real-Time Estimation of the Risk of Death from Novel Coronavirus (COVID-19) Infection: Inference Using Exported Cases. *J. Clin. Med*. 2020. vol. 9. no. 2.
40. Cooper I., Mondal A., Antonopoulos C. G. A SIR model assumption for the spread of COVID-19 in different communities. *Chaos, Solitons and Fractals*. 2020. vol. 139.

Захаров Виктор Васильевич — д-р физ.-мат. наук, профессор, заведующий кафедрой, кафедра математического моделирования энергетических систем, факультет прикладной математики – процессов управления, Санкт-Петербургский государственный университет. Область научных интересов: представление и обработка данных и знаний с неопределенностью, применение методов математики и информатики в социокультурных и эпидемиологических исследованиях. Число научных публикаций — 230. v.zaharov@spbu.ru; Университетский пр., 35, 198504, г. Санкт-Петербург, Россия; р.т.: +7 (812) 328-20-00.

Балькина Юлия Ефимовна — канд. физ.-мат. наук, доцент, факультет прикладной математики – процессов управления, Санкт-Петербургский государственный

университет. Область научных интересов: математическое моделирование, анализ данных, применение методов математического моделирования в медицине. Число научных публикаций — 80. j.balykina@spbu.ru; Университетский пр., 35, 198504, г. Санкт-Петербург, Россия; р.т.: +7 (812) 328-20-00.

V. ZAKHAROV, Y. BALYKINA
**BALANCE MODEL OF COVID-19 EPIDEMIC BASED ON
PERCENTAGE GROWTH RATE**

Zakharov V., Balykina Y. Balance Model of COVID-19 Epidemic Based on Percentage Growth Rate.

Abstract. The paper examines the possibility of using an alternative approach to predicting statistical indicators of a new COVID-19 virus type epidemic. A systematic review of models for predicting epidemics of new infections in foreign and Russian literature is presented. The accuracy of the SIR model for the spring 2020 wave of COVID-19 epidemic forecast in Russia is analyzed. As an alternative to modeling the epidemic spread using the SIR model, a new CIR discrete stochastic model is proposed based on the balance of the epidemic indicators at the current and past time points. The new model describes the dynamics of the total number of cases (C), the total number of recoveries and deaths (R), and the number of active cases (I). The system parameters are the percentage increase in the C(t) value and the characteristic of the dynamic balance of the epidemiological process, first introduced in this paper. The principle of the dynamic balance of epidemiological process assumes that any process has the property of similarity between the value of the total number of cases in the past and the value of the total number of recoveries and deaths at present. To calculate the values of the dynamic balance characteristic, an integer linear programming problem is used. In general, the dynamic characteristic of the epidemiological process is not constant. An epidemiological process the dynamic characteristic of which is not constant is called non-stationary. To construct mid-term forecasts of indicators of the epidemiological process at intervals of stationarity of the epidemiological process, a special algorithm has been developed. The question of using this algorithm on the intervals of stationarity and non-stationarity is being examined. Examples of the CIR model application for making forecasts of the considered indicators for the epidemic in Russia in May-June 2020 are given.

Keywords: COVID-19, epidemic modeling of new viruses, SIR models, forecast.

Zakharov Victor — Ph.D., Dr.Sci., Professor, Head of department, Mathematical modeling of energy systems department, Faculty of applied mathematics and control processes, Saint Petersburg State University. Research interests: uncertain knowledge and data representation and processing, application of mathematics and computer science in sociocultural and epidemiological studies. The number of publications — 230. v.zakharov@spbu.ru; 35, University pr., 198504, St. Petersburg, Russia; office phone: +7 (812) 328-20-00.

Balykina Yulia — Ph.D., Associate Professor, Faculty of applied mathematics and control processes, Saint Petersburg State University. Research interests: mathematical modeling, data mining, application of mathematical modeling in medicine. The number of publications — 80. j.balykina@spbu.ru; 35, University pr., 198504, St. Petersburg, Russia; office phone: +7 (812) 328-20-00.

References

1. Foppa I.M. A historical introduction to mathematical modeling of infectious diseases. Academic Press. 2017.
2. Shinde G.R., Kalamkar, A.B., Mahalle P.N., et al. Forecasting Models for Coronavirus (COVID-19): A Survey of the State-of-the-Art. SN Computer Science. 2020. vol. 1.

3. Mofhtakhar L., Seif M., Safe M.S. Exponentially increasing trend of infected patients with COVID-19 in Iran: a comparison of neural network and ARIMA forecasting models. *Iran Journal of Public Health*. vol. 49. pp. 92–100.
4. Chaudhry R.M. et al. Coronavirus disease 2019 (COVID-19): forecast of an emerging urgency in Pakistan. *Cureus*. 2020. vol. 12. no 5.
5. Tandon H., Ranjan P., Chakraborty T., Suhag V. Coronavirus (covid-19): Arima based time-series analysis to forecast near future. arXiv:2004.07859.
6. Kermack W.O., McKendrick A.G. A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London. Series A*. 1927. vol. 115. no. 772. pp. 700–721.
7. Anderson R.M., May R.M. *Infectious diseases of humans: Dynamics and control*. Oxford: Oxford University Press. 1991. P. 757
8. Dil S., Dil N., Maken Z.H. COVID-19 trends and forecast in the eastern mediterranean region with a particular focus on Pakistan. *Cureus*. 2020. vol. 12. no 6.
9. Johns Hopkins Coronavirus Resource Center. Available at: <https://coronavirus.jhu.edu/data> (accessed: 20.04.2021)
10. Liao Z., Lan P., Liao Z. et al. TW-SIR: time-window based SIR for COVID-19 forecasts. *Sci Rep*. 2020. vol. 10.
11. Rădulescu A., Williams C., Cavanagh K. Management strategies in a SEIR-type model of COVID 19 community spread. *Sci Rep*. 2020. vol. 10.
12. Fanelli D., Piazza F. Analysis and forecast of COVID-19 spreading in China, Italy and France. *Chaos Solitons Fractals*. 2020. vol. 134.
13. Miller A., et al. Correlation between universal BCG vaccination policy and reduced morbidity and mortality for COVID-19: an epidemiological study. medRxiv 2020.03.24.20042937.
14. Cheng Z., et al. Icumonitoring.ch: a platform for short-term forecasting of intensive care unit occupancy during the COVID-19 epidemic in Switzerland. *Swiss Medical Weekly*. 2020. vol. 150.
15. Rodrigues H.S. Application of SIR epidemiological model: new trends. *International Journal of Applied Mathematics and Informatics*. 2016. vol. 10. pp. 92-97.
16. Iwami S., Takeuchi Y., Liu X. Avian–human influenza epidemic model. *Mathematical biosciences*. 2007. vol. 207. no. 1. pp. 1–25.
17. Teles P. Predicting the evolution of SARS-COVID-2 in Portugal using an adapted SIR model previously used in South Korea for the MERS outbreak. medRxiv 2020.03.18.20038612.
18. Maier B.F., Brockmann D. Effective containment explains subexponential growth in recent confirmed COVID-19 cases in China. *Science*. 2020. vol. 368. no. 6492. pp. 742-746.
19. Chinazzi M., Davis J.T., et al. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science*. 2020. vol. 368. no. 6489. pp. 395-400.
20. Tang B., Wang X., Li Q., Bragazzi N.L., Tang S., Xiao Y., et al. Estimation of the transmission risk of 2019-nCov and its implication for public health interventions. *Journal of Clinical Medicine*. 2020. vol. 9. no. 2.
21. Tian H., Liu Y., Li Y. An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. *Science*. 2020. vol. 368. no. 6491. pp. 638-642.
22. López L., Rodó X. A modified SEIR model to predict the COVID-19 outbreak in Spain and Italy: Simulating control scenarios and multi-scale epidemics. *Results in Physics*. 2021. vol. 21.
23. Feng S., Feng Z., Ling C., Chang C., Feng Z. Prediction of the COVID-19 epidemic trends based on SEIR and AI models. *PLoS ONE*. 2021. vol. 16. no. 1.

24. Krivorotko O.I., Kabanikhin S.I., Zyatkov N.Y., Prikhodko A.Y., Prokshoshin N.M., Shishlenin M.A. Matematicheskoe modelirovanie i prognozirovanie COVID-19 v Moskve i Novosibirskoj oblasti. [Mathematical modeling and prediction of COVID-19 in Moscow city and Novosibirsk region]. Available at: <https://arxiv.org/abs/2006.12619v1> (accessed: 30.04.2021) (In Russ.).
25. Matveev A. V. Matematicheskoe modelirovanie ocenki jeffektivnosti mer protiv rasprostraneniya jepidemii COVID-19. Nacional'naja bezopasnost' i strategicheskoe planirovanie. [The mathematical modeling of the effective measures against the COVID-19 spread]. National Security and Strategic Planning. 2020. no. 1. pp. 23-39.
26. Anastasopoulou C., Russo L., Tsakris A., Siettos C. Data-based analysis, modelling and forecasting of the COVID-19 outbreak. PloS One. 2020. vol. 15. no. 3.
27. Mandal S., Bhatnagar T., Arinaminpathy N. Prudent public health intervention strategies to control the coronavirus disease 2019 transmission in India A mathematical model-based approach. Indian Council of Medical Research. 2020. vol. 151. pp. 190-199.
28. Choi S., Ki M. Estimating the reproductive number and the outbreak size of COVID-19 in Korea. Epidemiology and Health. 2020. vol. 42.
29. Tolles J., Luong T. Modeling Epidemics with Compartmental Models. JAMA. 2020. vol. 323. no. 24. pp. 2515-2516.
30. Fudolig M., Howard R. The local stability of a modified multi-strain SIR model for emerging viral strains. PLoS One. 2020. vol. 15. no. 12.
31. Adam D. Special report: The simulations driving the world's response to COVID-19. Nature. 2020. vol. 580. pp. 316-318.
32. Wieczorek M., Silka J., Woźniak M. Neural network powered COVID-19 spread forecasting model. Chaos, Solitons and Fractals. 2020. vol. 140.
33. Kondratyev M.A. Metody prognozirovaniya i modeli rasprostraneniya zabojevanij. Komp'yuternye issledovanija i modelirovanie. [Forecasting methods and models of disease spread]. Computer Research and Modeling. 2013. vol. 5. no. 5. pp. 863-882.
34. Zakharov V., Balykina Y., Petrosian O., Gao H. CBRR Model for Predicting the Dynamics of the COVID-19 Epidemic in Real Time. Mathematics. 2020. vol. 8. no. 10.
35. Zakharov V.V., Balykina Yu.E. Prognozirovanie dinamiki jepidemii koronavirusa (COVID-19) na osnove metoda precedentov. [Predicting the dynamics of the coronavirus (COVID-19) epidemic based on the case-based reasoning approach]. Vestnik of Saint Petersburg University. Applied Mathematics. Computer Science. Control Processes, 2020. vol. 16., no. 3. pp. 249–259. (In Russ.).
36. Dairi A., Harrou F., Zeroual A., Hittawe M.M., Sun Y. Comparative study of machine learning methods for COVID-19 transmission forecasting. Journal of Biomedical Informatics. 2021. vol. 18. Article ID 103791.
37. Mizumoto K., Chowell G. Transmission potential of the novel coronavirus (COVID-19) onboard the diamond Princess Cruises Ship, 2020. Infectious Disease Modelling. 2020. vol. 5. pp. 264-270.
38. Zhang S., et al. Estimation of the reproductive number of novel coronavirus (COVID-19) and the probable outbreak size on the Diamond Princess cruise ship: A data-driven analysis. Int J Infect Dis. 2020. vol. 93. pp. 201-204.
39. Jung S., et al. Real-Time Estimation of the Risk of Death from Novel Coronavirus (COVID-19) Infection: Inference Using Exported Cases. J. Clin. Med. 2020. vol. 9. no. 2.
40. Cooper I., Mondal A., Antonopoulos C. G. A SIR model assumption for the spread of COVID-19 in different communities. Chaos, Solitons and Fractals. 2020. vol. 139.

В.Ю. ОСИПОВ, С.В. КУЛЕШОВ, А.А. ЗАЙЦЕВА, А.Ю. АКСЕНОВ
**ПОДХОД К ЛОКАЛИЗАЦИИ ИСТОЧНИКА ЭПИДЕМИИ
COVID-19 В РОССИИ НА ОСНОВЕ МАТЕМАТИЧЕСКОГО
МОДЕЛИРОВАНИЯ**

Осипов В.Ю., Кулешов С.В., Зайцева А.А., Аксенов А.Ю. Подход к локализации источника эпидемии COVID-19 в России на основе математического моделирования.

Аннотация. В статье описаны результаты обработки статистических данных из открытых источников по развитию эпидемии COVID-19 и выполненного исследования по определению места и времени начала ее в России. В интересах предлагаемого исследования дан обзор существующих моделей процессов развития эпидемии и методов решения прямых и обратных задач его анализа. Предложена модель развития эпидемии COVID-19 в сети из девяти городов России: Москва, Санкт-Петербург, Нижний Новгород, Ростов-на-Дону, Краснодар, Екатеринбург, Новосибирск, Хабаровск, Владивосток. Города выбраны как по географическому положению, так и по количеству населения в них. Модель состоит из двадцати семи дифференциальных уравнений. Разработан алгоритм обратного анализа модели эпидемии. В качестве исходных данных для решения задачи выступали сведения по численности населения городов, интенсивности переходов процесса из одних состояний в другие, а также данные по инфицированности населения на заданные моменты времени. В статье также приводятся результаты детального анализа подходов решения к моделированию развития эпидемий по видам моделей (базовая модель SEIR, модель SIRD, адаптивная поведенческая модель, модифицированные SEIR-модели), и по странам (в Польше, во Франции, Испании, Греции и других), а также обзор прикладных задач, которые можно решить, используя моделирование распространения эпидемий. Рассматриваются дополнительные параметры среды, которые влияют на моделирование распространения эпидемий и могут учитываться для повышения точности результатов. По результатам моделирования установлены наиболее вероятные города-источники начала эпидемии в России, а также момент ее начала. Достоверность полученных оценок во многом определяется достоверностью использованных статистических данных по развитию COVID-19, находящихся в открытом доступе.

Ключевые слова: математическое моделирование, COVID-19, решение обратной задачи анализа, прогнозирование, модели SEIR, моделирование распространения эпидемий.

1. Введение. Одной из серьезных проблем 21 века, с которой столкнулось человечество, стала внезапная пандемия, вызванная вирусом COVID-19. Откуда появилась эта инфекция до сих пор неизвестно. Впервые о ней официально заговорили в декабре 2019 года после событий в китайском городе Ухань [1]. Некоторые считают, что именно там она и возникла [1, 2]. Существуют исследования, результаты которых утверждают о появлении коронавируса в США в 2019 году [3]. Другие, опираясь на факты проявления аналогичных признаков заболевания в США и Европе еще до декабря 2019 года, опровергают эти утверждения [2]. Установить первоисточник и место где появился

данный вид коронавируса, способный быстро распространяться среди людей, крайне затруднительно. С появлением этого заболевания в различных странах реакция была неоднозначной. Ряд стран долго не признавали угрозу и не предпринимали адекватных мер [4-6]. Действенные меры противодействия инфекции в мире, в том числе ограничительного характера по передвижению людей между странами, начали предприниматься, когда она уже широко распространилась. Несмотря на сложную запутанную ситуацию с распространением заболевания, по набранным статистическим данным удалось сформировать некоторую картину его развития, как за рубежом, так и в России. Однако, для более точного прогнозирования дальнейшего развития эпидемии и появления новых штаммов вируса [7], которые уже стали большой проблемой для формирования коллективного иммунитета и работоспособности предложенных ранее математических моделей, необходимо более-менее точно оценить время и место появления первых случаев заболевания.

К сожалению, из-за существенно отличающихся темпов развития эпидемии в разных странах и городах, по существующему состоянию трудно судить о месте и времени начала эпидемии. Показатели ее развития существенно зависят от численности и плотности населения, особенностей окружающей среды, реализуемых мер профилактики и лечения, появления и особенностей новых штаммов вируса, социальной позиции большинства и других факторов.

Для строгого обоснования места и времени начала эпидемии COVID-19 необходимо наличие и использование достоверных исходных данных и адекватных математических методов, учитывающих специфику анализируемого процесса. Несмотря на предпринимаемые многими учеными попытки математического обоснования и идентификации неизвестных параметров этого процесса, вопрос пока остается открытым [8-13].

В данной статье, на основе знания о специфике анализируемого процесса, обработки находящихся в открытом доступе статистических данных и известных решений, предложен метод, позволяющий в некоторой мере прояснить поставленный вопрос для Российской Федерации. С использованием этого метода установлены примерное время начала эпидемии и города России, являющиеся наиболее вероятными кандидатами на звание источника ее начала.

2. Анализ известных методов. Формализация процесса развития эпидемии может осуществляться в дискретном и непрерывном времени. Известны методы анализа эпидемии, основанные на линейных и нелинейных регрессионных моделях, нейросетевых решениях,

клеточных автоматах, системах дифференциальных уравнений и другие [8-13]. У каждой из этих моделей есть свои плюсы и минусы при использовании в определенных условиях. Для определения исходных состояний этих моделей и параметров используются различные варианты мониторинга ситуации и обработки статистических данных. Эти методы позволяют анализировать процессы развития эпидемии COVID-19 на уровне взаимодействия как индивидуумов, так и многочисленных групп населения.

Наиболее широкое распространение при анализе темпов развития вирусных эпидемий, в том числе и эпидемии COVID-19, получили модификации математических моделей типа SEIR (где: S - восприимчивые, E - в инкубации, I - инфекционные больные, R - переболевшие) [14], являющихся развитием классических моделей SIR, предложенных Kermack и McKendrick в 1927 году [15]. В России методы описания динамики развития эпидемических процессов были разработаны в 1960–70 годах В. Барояном и Л.А. Рвачевым [13, 16]. Предложенные ими уравнения описывают процесс развития эпидемии без учета статистической природы эпидемических процессов, давая оценку среднего ожидаемого количества заболевших. Развитие этих подходов учитывает также и влияние множества случайных известных и неизвестных факторов [17]. В [12] предложена адаптивная поведенческая модель (англ. Adaptive Compartmental Model: ACM) ACM-SEIR, которая формулируется в виде системы 7 сильно нелинейных обыкновенных дифференциальных уравнений. Эта модель предназначена для определения не только основных переменных S, E, I, R, но и переменных, отвечающих за действия правительства, реакцию населения, силу воздействия правительственных органов и других параметров. Базовая модель ACM-SEIR допускает обобщение за счет добавления в систему новых уравнений для других процессов типа S, E, I, R и т.д. Модель ACM-SEIR, будучи максимально приближенной к реальному процессу, не имеет ни аналитического, ни сколько-нибудь легко определяемого численного решения. Поэтому для качественного исследования модели требуется использовать подход, основанный на методах глубокого обучения.

В [8-13, 18-20] и других источниках в основном решаются прямые задачи анализа развития эпидемий. Например, в работе [21] на основе анализа многоуровневых пространственных данных показана сильная синхронизация сезонных эпидемий гриппа США, Дании и Норвегии.

Основываясь на опыте моделирования более ранних вирусных эпидемий, особенно вирусов гриппа и Эболы, проведены исследования

по прогнозированию развития эпидемии с COVID-19 в различных странах. Наибольшее количество исследований посвящено решению задачи прогнозирования развития эпидемии COVID-19 в странах, сильно пострадавших на начальном этапе: США, Индия, Бразилия, Франция, Россия, Великобритания, Турция, Италия, Китай. Развитие ситуации в других странах, таких как Греция, Израиль, Швеция, Норвегия, также интересно с точки зрения получения исходных данных для более точного прогнозирования ситуации.

Необходимо отметить, что для точного установления источников эпидемии и момента ее начала посредством моделирования имеет смысл использовать наборы статистических данных, собранные на начальных этапах ее развития, до введения различных противоэпидемических мер, так как они отражают достаточно объективную картину распространения коронавирусной инфекции. При этом решение задачи установления источников эпидемии и момента ее начала осуществимо с применением различных ее моделей и методов прогнозирования.

Рассмотрим применимость этих методов в ряде стран. В [9] для прогнозирования развития эпидемии в Польше, во Франции и Испании применен модифицированный подход на основе клеточных автоматов. Он использован для проверки (по доступным эпидемиологическим и социальным данным) потенциальных причин наблюдаемых особенностей эпидемии. Также этот подход применен для разработки руководящих принципов, которые будут наиболее эффективными с точки зрения государственных целей противодействия эпидемии. Он основан на модифицированной модели распространения гриппа с использованием двумерных клеточных автоматов, описанной в статье [22]. Предлагаемая в [9] модель устраняет недостатки традиционных моделей [23] за счет включения информации об инфекциях, привносимых на конкретную территорию прибывающими индивидуумами, а также учета реальной возрастной структуры населения и возрастной уязвимости людей. Предлагаемые в [9] результаты моделирования показывают, что один контакт за два дня с инфицированными людьми приводит к заражению более трех человек. Мобильность населения также влияет на скорость распространения эпидемии. Чем выше мобильность населения, тем выше рост числа инфицированных, особенно в долгосрочной перспективе. Предлагаемая модель также позволяет анализировать время реагирования в раннем периоде развития эпидемии. Это может быть основой для принятия соответствующих мер в зависимости от ожидаемых эффектов, характеризуемых параметрами продолжительности эпидемии и уровнем заболеваемости граждан.

Для анализа ситуации в Греции сделана попытка применить метод обработки статистических данных для мониторинга вспышек с использованием моделей временных рядов, распределений вероятностей и эпидемиологической модели [10]. Авторы [10] генерируют прогнозы и дают оценку прекращения или хотя бы начала нисходящей тенденции для вспышки COVID-19 в Греции на основе комбинации Newbold-Granger [24] классических авторегрессионных моделей. Строятся альтернативные сценарии развития эпидемии с помощью логнормального распределения. При этом оценки ожидаемого зарегистрированного уровня и тенденции к снижению вспышки делаются с помощью эпидемиологической модели tSIR, а также метода экспоненциального роста [10]. Кроме того, рассматриваются различные распределения вероятностей для моделирования данных зарегистрированных случаев. Используя выбранные модели прогнозирования, а также аппроксимации распределения вероятностей, авторы [10] рассматривают и анализируют различные сценарии в соответствии с серьезностью будущей ситуации. Эпидемиологическая модель tSIR используется как для оценки определенных мер, которые представляют собой ожидаемый зарегистрированный уровень, так и для оценки прекращения пандемии в каждой стране.

Много исследований посвящено моделированию распространения эпидемии COVID-19 в Китае, что вполне объяснимо с учетом уханьского штамма вируса, распространившегося по всему миру в первые месяцы развития эпидемии. В [25-26] прогнозировалось распространение эпидемий во Франции, Италии и Китае с использованием модели эпидемии SIRD. Также для моделирования динамики заболевания в крупных городах Китая применена эпидемиологическая модель SIR [27]. В работе [28] рассмотрена модель передачи инфекции для изучения ранней динамики развития эпидемии в Ухане. Показательно, что для анализа ситуации в Китае множество исследователей занимается развитием математических компартментных моделей [18] с использованием методов искусственного интеллекта. Например, в работе [29] применены подходы на основе машинного обучения, создание автоэнкодера для прогнозирования распространения коронавируса.

Достаточно большую сложность для моделирования и прогнозирования ситуации с эпидемией коронавируса представляют такие страны как США и Россия, как из-за распределенного географического положения, так и из-за неравномерной плотности населения внутри этих стран. Тем не менее, в таких исследованиях как [8, 30] делаются достаточно успешные попытки обобщения и прогнозирования с ис-

пользованием методов математического моделирования, основанных на системах дифференциальных уравнений и SIR-моделях.

В [31] выполнено объемное моделирование ситуации с коронавирусной инфекцией в Индии на основе статистических данных, полученных на начало 2020 года. Используются модифицированные уравнения классической SEIR модели. Интересным решением представляется применение модели количества помещенных в самоизоляцию или на карантин восприимчивых к вирусу людей, а также инфицированных лиц с явными и неявными симптомами. Последний параметр представляется достаточно спорным в связи с его малой достоверностью.

Еще одной значимой для исследователей страной является Бразилия, стоящая в настоящее время на третьем месте по абсолютному числу случаев заболевания новой коронавирусной инфекцией. Исследователи для прогнозирования ситуации в этой стране опираются не только на применение классических эпидемиологических моделей, но и современных методов, в том числе методов искусственного интеллекта [32]. В данном исследовании сделана попытка использования нейронных сетей, байесовской регрессии, методов k-ближайших соседей и алгоритмов случайного леса для прогнозирования количества новых случаев заболевания на недельном горизонте планирования.

Другие исследования представляют интерес с точки зрения применимости одних и тех же методов к эпидемиям различной природы. Так, в исследовании [19] предлагаемая авторами модель описывает распространение инфекции различной природы, будь это компьютерный вирус или эпидемия COVID-19.

В [20] приводится решение задачи построения сетевой модели аналитического вывода эпидемического порога, с учетом структуры сети и динамики болезни.

Решение прямой задачи распространения COVID-19 с использованием моделей, основанных на дифференциальных уравнениях, также известно из работ [33-35], что говорит об адекватности таких моделей.

Ряд работ используют такие исходные данные как зарегистрированные случаи COVID-19, данные о переписи и миграции и ежемесячные данные авиакомпаний о пассажирах. Их предпосылки основаны на предположении о том, что существующие модели, основанные исключительно на аспектах темпов роста или обобщенных сетевых концепциях, могут не обеспечивать требуемой точности прогнозов. Для более точного прогнозирования необходимо учитывать особенности внутреннего перемещения населения внутри страны. В [36] авторы предлагают модель, основанную на сетевой схеме перемещений, и учитывающую трафик населения между частями страны.

В некоторых работах производится учет более тонких закономерностей миграции, влияющей на распространение пандемии. При этом учитывается уровень дохода населения, плотность населения городов, образ жизни, влияние налоговой и политической базы, недвижимости [37].

Есть и такие работы, где исследуется влияние развития пандемии на жизнь трудовых мигрантов [38], позволяющие выявить новые закономерности, влияющие на перенос коронавируса между регионами одной страны.

Решаются в явном виде также обратные задачи анализа процесса развития эпидемий. Так в [8] сделана попытка решения обратной задачи восстановления коэффициентов для математической модели распространения коронавируса COVID-19 в регионах Российской Федерации. В [39] рассматривается проблема обратного отслеживания контактов для предотвращения распространения эпидемий. Предложена модель отслеживания источника заражения. Учитываются ошибки выборки, возникающие из-за неоднородности количества контактов, которые имеет человек. Однако исследование направлено не на поиск территории, на которой появился источник заражения, а на поиск ближайших источников заражения. Известны также работы по изучению и отслеживанию контактов при распространении не только COVID-19, но и других вирусов [40-46].

При моделировании эпидемий можно использовать большой класс дополнительных входных данных, которые напрямую не относятся к медицинским аспектам. Среди них данные относящиеся к социальным процессам, данные относящиеся к управленческим решениям в отношении введения ограничительных мер, и противоэпидемических мероприятий, а также погодных условий. Все эти процессы могут влиять на распространение заболеваний как непосредственно, так и опосредовано. В частности, изменение погодных условий может снижать количество социальных контактов на природе или вне помещений (в случае холодной погоды или наличия осадков) и увеличивать количество социальных контактов в помещениях (при отсутствии запрета на посещение общественных мест и культурных мероприятий). Одновременно с этим, осадки и холодная погода способствуют возникновению простудных заболеваний, которые, в свою очередь, не только снижают иммунитет и делают людей более восприимчивыми к инфекциям, но и могут ошибочно диагностироваться как симптомы основного заболевания (COVID-19). Социальные процессы также имеют большое влияние на распространение вирусов, причем зависимость между этими процессами имеет сильно нелинейный характер.

Среди измеримых показателей социальных процессов можно выделить: уровень осознания населением важности проводимых противоэпидемических мероприятий (количество вакцинированного населения, количество невакцинированного населения, активность использования средств индивидуальной защиты (СИЗ) в общественных местах, количество социальных контактов индивидов), реакцию населения на управленческие решения государственных органов (проведение акций протеста, порождающих массовое скопление людей, не использующих СИЗ, доля людей, прекративших использовать СИЗ в общественных местах), распространение сведений о бесполезности СИЗ и процесса вакцинации. Оценка эффективности управленческих решений может выражаться численно в оценке уменьшения количество социальных взаимодействий в общественных пространствах, исключении перемещения населения между регионами внутри страны и между странами, и частично коррелирует со значениями транспортных потоков, которые рассмотрены выше.

При всех потенциальных достоинствах использования этих дополнительных параметров, использующие их модели применимы скорее для прогнозирования, а не для ретроспективного анализа.

В случае решения обратной задачи удастся лишь проверить их значимость для модели и адекватность их использования за счет сравнения математически определенных значений с известными опубликованными статистическими значениями, и выявления некоторых закономерностей, которые напрямую не следуют из теоретических построений. В любом случае, использование таких параметров – это цель отдельного исследования.

Однако в каждом конкретном случае приходится учитывать свои особенности анализируемого процесса, требующие построения адекватных моделей и алгоритмов решения задач анализа.

С учетом вышесказанного рассмотрим в следующих разделах модель и метод решения обратной задачи анализа развития COVID-19 в России, учитывающий особенности ее городов и связи между ними.

3. Модель развития эпидемии. Для решения поставленной задачи поиска города - источника начала эпидемии COVID-19 в России будем исходить из следующих положений и допущений:

– изначально возникновение очага заболевания в любом городе может произойти только за счет появления носителей инфекции. Будем считать, что носители инфекции могут прибыть в конкретный город из уже имеющегося очага заболевания различным видом транспорта;

– распространение инфекции между городами возможно параллельно – последовательным путем. Носители инфекции могут одновременно отправиться в разные города и регионы, а также один и тот же носитель может посетить последовательно несколько городов друг за другом;

– интенсивность переноса заболеваний из одних городов в другие зависит от уровня перемещения людей между ними. Чем ближе географически или удобнее с точки зрения транспортного сообщения расположены города по отношению друг к другу, тем этот уровень выше;

– в близко расположенных городах при одних и тех же условиях среды существования законы развития инфекционных заболеваний будут очень мало отличаться, в том числе и по смещению начала эпидемии во времени. Это позволяет при моделировании процесса развития эпидемии COVID-19 в России ограничиться лишь анализом событий в ряде крупных удаленных городов, относящихся к различным регионам.

Для решения рассматриваемой задачи предлагается проанализировать развитие этой эпидемии в Санкт-Петербурге, Москве, Нижнем Новгороде, Ростове-на-Дону, Краснодаре, Екатеринбурге, Новосибирске, Хабаровске и Владивостоке. Данные города являются центрами соответствующих регионов России, находятся на достаточном удалении друг от друга и от столиц. Москва и Санкт-Петербург должны быть включены в список анализируемых городов, так как помимо наибольшей численности населения в них, они являются крупнейшими в РФ транспортными узлами. Для снижения сложности задачи допустим, что в каждом из этих городов люди могут находиться только в трех состояниях здоровья: здоровые, инфицированные и переболевшие. Полагаем что, умершие люди компенсируются новорожденными, а интенсивности инфицирования людей зависят от числа инфицированных на текущий момент времени.

Ориентируясь на эти города и выделенные состояния здоровья людей, процесс развития эпидемии COVID-19 можно представить в виде графа состояний, упрощенный вид которого показан на рисунке 1. В основу построения этого графа положена модель SEIR [13]. На этом графе 1, 4, ..., 22, 25 – состояния, соответствующие относительному числу здоровых людей в рассматриваемых городах; 2, 5, ..., 23, 26 – состояния, отражающие относительное число инфицированных людей; 3, 6, ..., 24, 27 – состояния, характеризующие относительное число переболевших людей. Нормировка осуществляется относительно суммарной численности людей во всех девяти городах. Дуги на

графе (рис. 1) соответствуют процессам перехода людей из одних состояний в другие. Заметим, что переходы между состояниями 2, 5, ..., 23, 26 обусловлены перемещением инфицированных людей между городами с использованием различных видов транспорта.

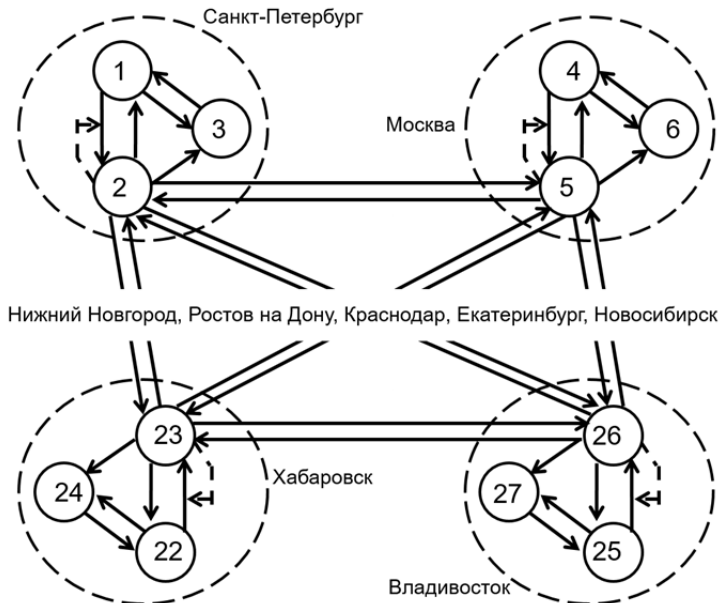


Рис. 1. Граф состояний процесса развития эпидемии в девяти городах

Опираясь на предельную теорему для суммарного потока, графу на рисунке 1 можно поставить в соответствие систему из 27 дифференциальных уравнений, вида:

$$\frac{dP_1(t)}{dt} = \lambda_{31}P_3(t) + \lambda_{21}P_2(t) - \lambda_{13}P_1(t) - \lambda_{12}P_2(t) \cdot P_1(t),$$

$$\begin{aligned} \frac{dP_2(t)}{dt} = & \lambda_{12}P_2(t) \cdot P_1(t) - \lambda_{21}P_2(t) - \lambda_{23}P_2(t) + \lambda_{52}P_5(t) + \lambda_{82}P_8(t) + \lambda_{11,2}P_{11}(t) + \lambda_{4,2}P_{14}(t) + \\ & + \lambda_{17,2}P_{17}(t) + \lambda_{20,2}P_{20}(t) + \lambda_{23,2}P_{23}(t) + \lambda_{26,2}P_{26}(t) - \\ & - (\lambda_{25} + \lambda_{28} + \lambda_{2,11} + \lambda_{2,14} + \lambda_{2,17} + \lambda_{2,20} + \lambda_{2,23} + \lambda_{2,26}) \cdot P_2(t), \end{aligned}$$

$$\frac{dP_3(t)}{dt} = \lambda_{13}P_1(t) + \lambda_{23}P_2(t) - \lambda_{31}P_3(t),$$

$$\frac{dP_{25}(t)}{dt} = \lambda_{27,25}P_{27}(t) + \lambda_{26,25}P_{26}(t) - \lambda_{25,27}P_{25}(t) - \lambda_{25,26}P_{26}(t) \cdot P_{25}(t),$$

$$\begin{aligned} \frac{dP_{26}(t)}{dt} = & \lambda_{25,26}P_{26}(t) \cdot P_{25}(t) - \lambda_{26,25}P_{26}(t) - \lambda_{26,27}P_{26}(t) + \lambda_{2,26}P_2(t) + \lambda_{5,26}P_5(t) + \lambda_{8,26}P_8(t) + \\ & + \lambda_{11,26}P_{11}(t) + \lambda_{4,26}P_{14}(t) + \lambda_{17,26}P_{17}(t) + \lambda_{20,26}P_{20}(t) + \lambda_{23,26}P_{23}(t) - \\ & - (\lambda_{26,2} + \lambda_{26,5} + \lambda_{26,8} + \lambda_{26,11} + \lambda_{26,14} + \lambda_{26,17} + \lambda_{26,20} + \lambda_{26,23}) \cdot P_{26}(t), \end{aligned}$$

$$\frac{dP_{27}(t)}{dt} = \lambda_{25,27}P_{25}(t) + \lambda_{26,27}P_{26}(t) - \lambda_{27,25}P_{27}(t).$$

В этой системе уравнений: $P_i(t)$ – вероятности нахождения процесса в i - состояниях на момент времени t ; λ_{ij} – интенсивности переходов процесса из i -х состояний j -е состояния. Заметим, что в этой системе уравнений интенсивности переходов процесса из состояний 1, 4, ..., 22, 25, соответственно в состояния 2, 5, ..., 23, 26 зависят от вероятностей нахождения его в этих состояниях. В явном виде в приведенной системе отражены уравнения только для Санкт-Петербурга и Владивостока. Для других городов уравнения выглядят аналогично. С использованием этой системы уравнений можно решать как прямые, так и обратные задачи анализа исследуемого процесса.

4. Алгоритм решения задачи. При решении прямой задачи анализа, когда заданы вероятности $P_i(t=0)$ нахождения процесса в i - состояниях на момент времени $t=0$ и определены интенсивности λ_{ij} , можно получить значения $P_i(t)$ на интересующий момент времени t .

В нашем случае требуется решить обратную задачу. На основе набранной статистики могут быть установлены значения P_{iz}^* нахождения людей в i -состояниях на Z -е моменты измерения. Однако неизвестно в каком городе и когда возник очаг заболевания.

Чтобы ответить на эти вопросы поступим следующим образом. Примем во внимание, что при всем здоровом населении городов веро-

ятности $P_2(t), P_3(t), \dots, P_{23}(t), P_{26}(t)$ равны нулю. Численность населения каждого города известна. Зададимся некоторым малым значением вероятности $P_{инф}$ инфицированности людей, которые изначально могут прибыть в любой из рассматриваемых городов. Затем можно разрешить приведенную выше систему дифференциальных уравнений относительно каждого потенциального города – источника развития эпидемии с этой вероятностью. Разрешение этой системы осуществимо с использованием, например, пакета прикладных программ MatLab, реализующего метод Рунге – Кутты четвертого порядка и другие. При разрешении можно сопоставить получаемые результаты с фактическими оценками. Город, для которого будет свойственна минимальная ошибка в расхождении модельных результатов с фактическими оценками, можно считать источником развития эпидемии.

С формальной точки зрения задача поиска такого города S_0 и момента t_0 , связанного с продолжительностью развития эпидемии, сводится к определению минимума целевой функции $W_0(S_0, t_0)$, суммарных относительных ошибок,

$$W_0(S_0, t_0) = \min_{S_k \in \Omega} \sum_{z=1}^Z \min_{t \in T} \sum_{i=1}^N \frac{|P_i(S_k, t) - P_{iz}^*|}{P_{iz}^*},$$

где Z – число временных моментов, на которые оценивается фактическое состояние развития эпидемии; N – число рассматриваемых городов, в нашем случае их девять; S_k – k -й город, рассматриваемый источником эпидемии. Когда поиск источника эпидемии осуществляется относительно результатов измерения на один момент времени, то

$$W_0(S_0, t_0) = \min_{\substack{t \in T \\ S_k \in \Omega}} \sum_{i=1}^N \frac{|P_i(S_k, t) - P_{i1}^*|}{P_{i1}^*}.$$

Особенность решения рассматриваемой задачи сводится к заданию начальных условий разрешения приведенной системы уравнений, при которых только для k -го города на момент времени $t = 0$ значение $P_k(0) = P_{инф}$, а для других городов $P_{i \neq k}(0) = 0$. При разрешении системы уравнений на каждый момент времени t рассчитывается суммарная ошибка

$$W(S_k, t) = \sum_{i=1}^N \frac{|P_i(S_k, t) - P_{i1}^*|}{P_{i1}^*}$$

и определяется момент времени t_0 , на который эта ошибка минимальна,

$$W(S_k, t_0) = \min_{t \in T} W(S_k, t).$$

Аналогичные расчеты выполняются для всех возможных городов – источников S_k инфекции. Из всех полученных $W(S_k, t_0)$ выбирается минимальное значение $W_0(S_0, t_0) = \min_{S_k \in \Omega} W(S_k, t_0)$. В итоге опре-

деляется город-источник S_0 инфекции и момент t_0 , характеризующий продолжительность развития инфекции (от нуля до t_0) на момент получения фактических оценок.

5. Исходные данные. Для решения рассматриваемой задачи в качестве исходных данных выступали: численность населения в анализируемых городах России, значения показателей инфицированности населения COVID-19 на 30 мая и 30 июня 2020 года, интенсивности переходов процесса развития эпидемии COVID-19 из одних состояний в другие. Для моделирования процесса развития эпидемии в России использовались данные по численности населения в городах и по статистической инфицированности населения в формате (численность населения, человек / статистическая вероятность инфицированности на 30 мая 2020 г. / статистическая вероятность инфицированности на 30 июня 2020 г.), представленные в таблице 1.

Таблица 1. Исходные данные по 9 анализируемым городам

Город	Численность населения, человек	Статистическая вероятность инфицированности на 30 мая 2020 г., ед.	Статистическая вероятность инфицированности на 30 июня 2020 г., ед.
Москва	12655050	0,006882783	0,00855920
Санкт-Петербург	5388759	0,000601774	0,00093500
Екатеринбург	1493749	0,000200231	0,00049410
Владивосток	606561	0,000075048	0,00016630
Хабаровск	616372	0,000077945	0,00016300
Ростов-на-Дону	1137904	0,000176902	0,00037120
Нижний Новгород	1252236	0,000368210	0,00067520
Новосибирск	1625631	0,000104480	0,00023120
Краснодар	932629	0,000144611	0,00023110

Интенсивности переходов процесса развития эпидемии в пределах каждого города определялись на основе анализа известной статистики [47]. Учитывались значения показателей инфицированности населения COVID-19 в мае и июне 2020 года. Для получения значений интенсивностей применялось правило, согласно которому каждая интенсивность – это отношение вероятности соответствующего перехода процесса за заданное время к этому времени, в нашем случае, в неделях.

Для определения интенсивностей перемещения людей из одних городов в другие анализировались регулярное движение транспорта (расписания полетов самолетов и движения железнодорожного транспорта), а также его вместимость.

С учетом этого для каждого города интенсивность перехода людей из здорового состояния в состояние инфицированности на текущий момент времени в неделях определялась как 0,6000, умноженное на вероятность нахождения людей уже в состоянии инфицированности. Интенсивность обратного перехода, минуя состояние «переболевшие», равнялась 0,0005. Переход из состояния «инфицированные» в состояние «переболевшие» осуществлялся с интенсивностью 0,0700. Интенсивность перехода из состояния «переболевшие» в состояние «здоровые» считалась равной 0,0100, а обратного перехода – 0,0020.

6. Результаты моделирования. Используя исходные данные, применяя рассмотренный метод и пакет прикладных программ MatLab, были получены результаты, характеризующие возможные города – источники коронавирусной инфекции (Таблица 2).

Таблица 2. Минимальные ошибки и моменты времени их проявления

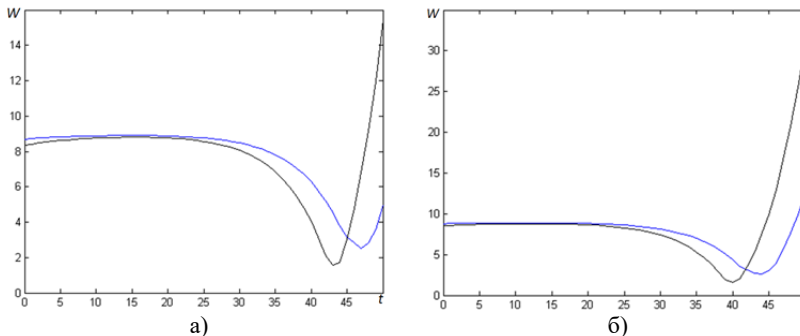
Город	Относительно данных на 30 мая 2020 г. минимум ошибки / момент времени (в неделях)	Относительно данных на 30 июня 2020 г. минимум ошибки / момент времени (в неделях)	Суммарные ошибки	Порядок инфицирования
Санкт-Петербург	1,5804 / 37	2,5819 / 41	4,1623	3
Москва	1,6190 / 28	2,6276 / 32	4,2466	7
Нижний Новгород	1,5679 / 37	2,6063 / 41	4,1742	4
Ростов-на-Дону	1,5951 / 37	2,6542 / 41	4,2493	9
Краснодар	1,6100 / 37	2,5836 / 41	4,1936	5

Екатеринбург	1,5996 / 39	2,6049 / 43	4,2045	6
Новосибирск	1,5630 / 40	2,5969 / 44	4,1599	2
Хабаровск	1,5741 / 43	2,5226 / 47	4,0967	1
Владивосток	1,5847 / 42	2,6625 / 46	4,2472	8

Данная таблица построена при сопоставлении модельных результатов и фактических данных по инфицированию людей в городах на 30 мая и 30 июня 2020 года. Анализ таблицы показывает, что относительно данных на 30 мая 2020 г. на первое место по порядку начала эпидемии выходит Новосибирск, затем Нижний Новгород и Хабаровск. Судя по данным на 30 июня 2020 года, Хабаровск выходит на первое место, за ним Санкт-Петербург и Новосибирск. По общим суммарным ошибкам лидирует Хабаровск с моментом начала развития эпидемии 43 недели до 30 мая 2020 г. Иначе говоря, судя по результатам моделирования, эпидемия в России начала развиваться в начале августа 2019 года в Хабаровске. При этом начало развития эпидемии за рубежом должно приходиться как минимум на два месяца раньше.

Следует также заметить, что, видимо, эпидемия COVID-19 в России начала развиваться не только с востока, но и с запада, в частности с Санкт-Петербурга. Значения минимальных ошибок в таблице 1 для Санкт-Петербурга близки к значениям ошибок для Хабаровска и Новосибирска.

На рисунке 2 показаны зависимости от времени суммарных относительных ошибок совпадения модельных результатов с фактическими данными относительно 30 мая и 30 июня 2020 года для четырех городов – прецедентов на первичные источники инфекции. Обратим внимание, что на каждом графике рисунка 2 разница по времени проявления минимумов ошибок составляет четыре недели, однако значения смещений пар минимумов по шкале времени отличаются.



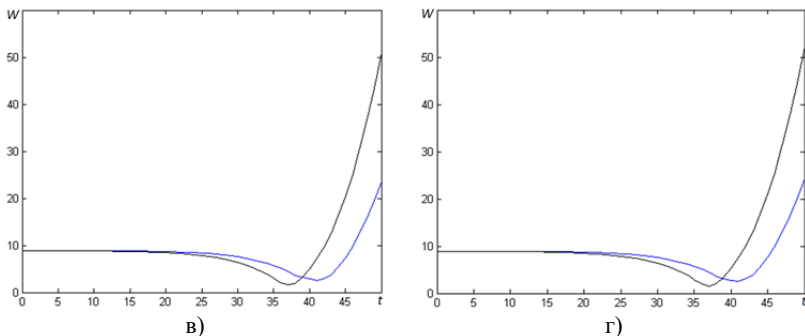
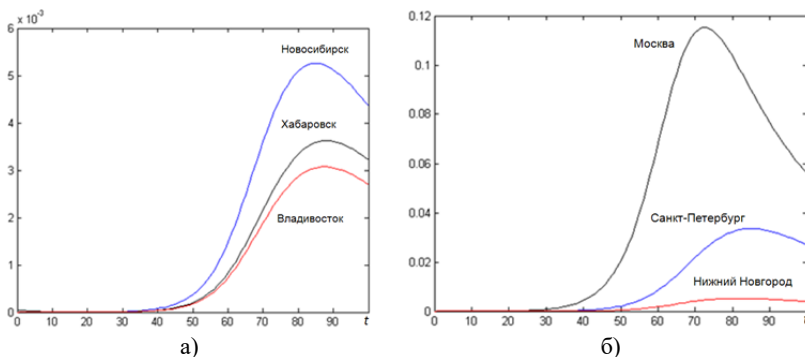


Рис. 2. Зависимости от времени (в неделях) суммарных относительных ошибок совпадения модельных результатов с фактическими данными относительно 30 мая и 30 июня 2020 года: а) при начале эпидемии в Хабаровске; б) в Новосибирске; в) Санкт-Петербурге; г) в Нижнем Новгороде

Рассмотрим, как изменяется эпидемиологическая ситуация во времени при начале развития заболевания в Хабаровске. На рисунке 3 отражены характерные этому случаю зависимости от времени (в неделях) вероятностей инфицированности людей городов (с учетом масштаба оси ординат). Отметим, что данная ситуация характерна только фиксированным мерам противодействия эпидемии.



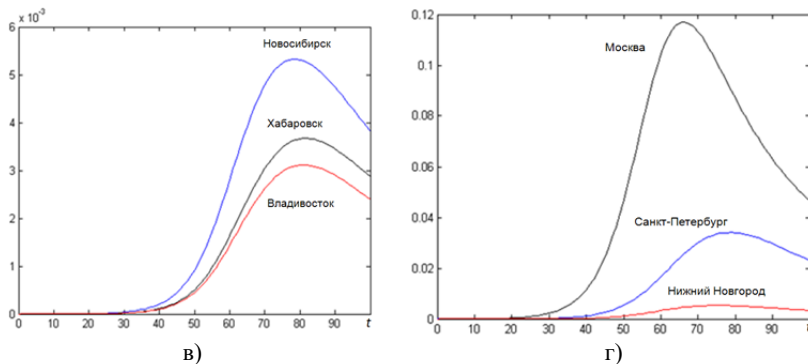


Рис. 3. Зависимости от времени (в неделях) вероятностей инфицированности людей в городах: а), б) - при начале эпидемии в Хабаровске; в), г) – в Санкт-Петербурге

При условии, что эпидемия началась с Хабаровска в начале августа 2019 года, пик заболеваемости для Москвы приходится на 72 неделю, т. е. на середину декабря 2020 г. Этот модельный результат хорошо согласуется со статистическими данными [33]. При начале эпидемии с Санкт-Петербурга пик заболеваемости для Москвы смещается влево на 66 неделю.

7. Заключение. В ходе выполненного исследования проведен анализ состояния проблемы определения места и времени появления вируса COVID-19 в России. Разработана модель развития эпидемии с COVID-19 в сети девяти городов России в виде 27 дифференциальных уравнений. Отличие этой модели от известных решений состоит в ранее не рассматриваемом пространстве состояний процесса и связей между ними. Предложен один из возможных алгоритмов решения обратной задачи анализа с использованием этой модели. Для фактического решения этой задачи были определены интенсивности переходов процесса развития эпидемии из одних состояний в другие.

По результатам моделирования установлено, что еще задолго до декабря 2019 года вирус COVID-19 мог «гулять» по многим странам. Период скрытого (мало заметного) распространения этого вируса мог достигать от 10 до 20 недель. Что касается развития эпидемии от вируса COVID-19 в России, то наиболее вероятно она началась почти одновременно с востока и запада. Первое место среди девяти исследованных городов России, возможных потенциальных источников начала эпидемии, занял Хабаровск, а второе место Новосибирск. На третьем месте находится Санкт-Петербург. Относительно Хабаровска начало

развития эпидемии приходится на начало августа 2019 года. Эпидемия же за рубежом началась как минимум на два месяца раньше, в июне 2019 года. Заметим, что достоверность полученных оценок во многом определяется достоверностью использованных статистических данных по развитию COVID-19, находящихся в открытом доступе.

Литература

1. Covid-19. URL: <https://ru.wikipedia.org/wiki/Covid-19> (дата обращения: 07.06.2021).
2. Так где же возник Covid-19? URL: <https://news-front.info/2021/05/03/tak-gde-zhe-voznik-covid-19> (дата обращения: 07.06.2021).
3. NIH study offers new evidence of early SARS-CoV-2 infections in U.S. URL: <https://www.nih.gov/news-events/news-releases/nih-study-offers-new-evidence-early-sars-cov-2-infections-us> (дата обращения: 07.06.2021).
4. Свой путь: весь год мир наблюдал за стратегией Швеции по борьбе с коронавирусом. URL: <https://tjournal.ru/analysis/249657-svoy-put-ves-god-mir-nablyudal-za-strategiyey-shvecii-po-borbe-s-koronavirusom-chto-iz-etogo-vyshlo> (дата обращения: 07.06.2021).
5. Sweden tells citizens to wear masks on public transport as it struggles with COVID-19 resurgence URL: <https://www.euronews.com/2020/12/18/sweden-tells-citizens-to-wear-masks-on-public-transport-as-it-struggles-with-covid-19-resu> (дата обращения: 07.06.2021).
6. Sweden set to begin easing coronavirus restrictions. URL: <https://www.ft.com/content/c64b7e37-f00b-4e74-acf8-49c039c57698> (дата обращения: 07.06.2021).
7. World Health Organization. URL: <https://www.who.int/> (дата обращения: 16.08.2021).
8. Криворотко О.И., Кабанихин С.И., Зятков Н.Ю. Математическая модель распространения коронавируса COVID-19 в Российской Федерации. URL: [https://www.nsu.ru/n/mathematics-mechanics-department/documents/Криворотко - Математическая модель распространения COVID-19 в РФ.pdf](https://www.nsu.ru/n/mathematics-mechanics-department/documents/Криворотко-Математическая%20модель%20распространения%20COVID-19%20в%20РФ.pdf) (дата обращения: 07.06.2021).
9. Medrek M., Pastuszak Z. Numerical simulation of the novel coronavirus spreading. Expert Systems with Applications. 2021. 166. 114109.
10. Katris C. A time series-based statistical approach for outbreak spread forecasting: Application of COVID-19 in Greece. A time series-based statistical approach for outbreak spread forecasting: Application of COVID-19 in Greece. Expert Systems with Applications. 2021. Vol. 166. 114077.
11. Nadella P., Swaminathan A., Subramanian S.V. Forecasting efforts from prior epidemics and COVID-19 predictions. Eur J Epidemiol. 2020. Vol. 35. pp. 727–729.
12. Levashkin S.P., Zakharova O.I., Kuleshov S.V., Zaytseva A.A. Adaptive-compartmental model of coronavirus epidemic and its optimization by the methods of artificial intelligence. Journal of Physics: Conference Series. 2021. vol. 1864.
13. SEIR and SEIRS models. URL: <https://docs.idmod.org/projects/emod-environmental/en/latest/model-seir.html> (дата обращения: 07.06.2021).
14. Bjørnstad O.N. Epidemics, Use R. Springer Nature Switzerland AG. 2018. 312 p.
15. Kermack W.O., McKendrick A.G. A contribution to the mathematical theory of epidemics. Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences. 1927. Vol. 115(772). pp. 700–721.
16. Бароян О.В., Рвачев Л.А. Математика и эпидемиология // М.: Знание. 1977. 64 с.

17. Десятков Б.М., Бородулин А.И., Котлярова С.С., Лаптева Н.А., Марченко М.Ю., Шабанов А.Н. Математическое моделирование эпидемических процессов и оценка их статистических характеристик // Химическая и биологическая безопасность. 2009. № 1–3. С. 43–45 URL: http://www.cbssafety.ru/rus/saf43_2009_st_2.pdf (дата обращения: 16.08.2021).
18. Xu H., Gonzalez M. J., Guo L., et al. Knowledge, Awareness, and Attitudes Towards COVID-19 Pandemic Among Different Populations in Central China: A Cross-Sectional Survey. *Journal of Medical Internet Research*. 2020.
19. Wuqiong L., Wee P.T., Mei L. Identifying Infection Sources and Regions in Large Networks. 2021. URL: <https://arxiv.org/pdf/1204.0354.pdf> (дата обращения: 07.06.2021).
20. Valdano E. et al. Analytical computation of the epidemic threshold on temporal networks. *Physical Review X*. 2015. Vol. 5. No. 2. 021005.
21. Morris S.E., Freiesleben de Blasio B., Viboud C., Wesolowski A., Bjørnstad O.N., Grenfell B.T. Analysis of multi-level spatial data reveals strong synchrony in seasonal influenza epidemics across Norway, Sweden, and Denmark. *PLoS ONE*. 2018. № 13(5).
22. Holko A., Medrek M., Pastuszak Z., Phusavat K. Epidemiological modeling with a population density map-based cellular automata simulation system. *Expert Systems With Applications*. 2016. Vol. 48. pp. 1–8.
23. Ahmed E., Agiza H.N. On modeling epidemics including latency, incubation and variable susceptibility. *Physica A: Statistical Mechanics and its Applications*. 1998. Vol. 253. No. 1-4. pp. 347–352.
24. Newbold P., Granger C.W.J. Experience with forecasting time series and combination of forecasts. *Journal of the Royal Statistical Society*. 1974. № 137(2). pp. 131–165.
25. Fanelli D., Piazza F. Analysis and forecast of COVID-19 spreading in China, Italy and France. *Chaos, Solitons & Fractals*. 2020. 134. Article 109761.
26. Anastassopoulou C., Russo L., Tsakris A., Siettos C. Data-based analysis, modelling and forecasting of the COVID-19 outbreak. *PLoS ONE*. 2020. № 15(3). e0230405.
27. Wu J.T., Leung K., Leung G.M. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: A modelling study. *The Lancet*. 2020. Vol. 395 (10225). pp. 689–697.
28. Li Q., Guan X., Wu P., Wang X., Zhou L., et al. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *New England Journal of Medicine*. 2020. URL: [https://valiasr.zums.ac.ir/files/i_management/files/nejmoa2001316_\(1\).pdf](https://valiasr.zums.ac.ir/files/i_management/files/nejmoa2001316_(1).pdf) (дата обращения: 16.08.2021).
29. Hu Z., Qiyang G., Shudi L., Li J., & Momiao X. Artificial Intelligence Forecasting of Covid-19 in China. *arXiv preprint*. 2020. arXiv:2002.07112. URL: <https://arxiv.org/abs/2002.07112> (дата обращения: 16.08.2021)
30. Macdonald J.C., Browne C., Gulbudak H. Modelling COVID-19 outbreaks in USA with distinct testing, lockdown speed and fatigue rates. *Royal Society Open Science*. 2021. Aug 4; 8(8):210227.
31. Sarkar K, Khajanchi S, Nieto J. Modeling and forecasting the COVID-19 pandemic in India. *Chaos Solitons Fractals*. 2020. Oct. 139. 110049.
32. da Silva R.G, Ribeiro M.H.D.M., Mariani V.C., Coelho L.D.S. Forecasting Brazilian and American COVID-19 cases based on artificial intelligence coupled with climatic exogenous variables. *Chaos Solitons Fractals*. 2020. Oct. 139. 110027.
33. Sharma N., Kumar A., Kumar A. Spatial Network based model forecasting transmission and control of COVID-19.
34. Кольцова Э.М., Куркина Е.С., Васецкий А.М. Математическое моделирование распространения эпидемии коронавируса Covid-19 в ряде европейских,

- азиатских стран, Израиле и России // Проблемы экономики и юридической практики. 2020. Т. XVI. No2. С. 154-165.
35. Обеснюк В.Ф. Динамика локальной эпидемической вспышки COVID-19 через призму компартмент-моделирования // Анализ риска здоровью. 2020. №2. С. 83-91. 36. Kumar A. Modeling geographical spread of COVID-19 in India using network-based approach. URL: <https://www.medrxiv.org/content/10.1101/2020.04.23.20076489v1> (дата обращения: 16.08.2021).
36. Haslag P.H., Weagley D. From L.A. to Boise: How Migration Has Changed During the COVID-19 Pandemic (March 26, 2021). URL: <https://ssrn.com/abstract=3808326> or <http://dx.doi.org/10.2139/ssrn.3808326> (дата обращения: 07.06.2021).
37. Das D. Regional disparities of growth and internal migrant workers in informal sectors in the age of COVID-19.
38. Kojaku S., Hébert-Dufresne L., Mones E. et al. The effectiveness of backward contact tracing in networks. *Nat. Phys.* 2021. 17, pp. 652–658.
39. Klinkenberg D., Fraser C., Heesterbeek H. The effectiveness of contact tracing in emerging epidemics. *PLoS one.* 2006. Vol. 1. No. 1. pp. e12.
40. Müller J., Kretzschmar M., Dietz K. Contact tracing in stochastic and deterministic epidemic models. *Mathematical biosciences.* 2000. Vol. 164. No. 1. pp. 39-64.
41. Ball F. G., Knock E. S., O'Neill P.D. Threshold behaviour of emerging epidemics featuring contact tracing *Advances in Applied Probability.* 2011. Vol. 43. No. 4. pp. 1048-1065.
42. Browne C., Gulbudak H., Webb G. Modeling contact tracing in outbreaks with application to Ebola. *Journal of theoretical biology.* 2015. Vol. 384. pp. 33-49.
43. Okolie A., Müller J. Exact and approximate formulas for contact tracing on random trees. *Mathematical biosciences.* 2020. Vol. 321. pp. 108320.
44. Bianconi G. et al. Message-passing approach to epidemic tracing and mitigation with apps. *Physical Review Research.* 2021. Vol. 3. No. 1. pp. L012014.
45. Endo A. Centre for the Mathematical Modelling of Infectious Diseases COVID-19 Working Group, Leclerc QJ, et al. Implication of backward contact tracing in the presence of over dispersed transmission in COVID-19 outbreaks. *Wellcome Open Res.* 2021. Vol. 5:239. URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7610176/> (дата обращения: 07.06.2021).
46. Coronavirus. URL: <https://coronavirus-monitor.info> (дата обращения: 07.06.2021).

Осипов Василий Юрьевич — д-р техн. наук, профессор, директор СПИИРАН, СПИИРАН, СПб ФИЦ РАН. Область научных интересов: математическое моделирование, интеллектуальные системы, нейронные сети, информационная безопасность. Число научных публикаций — 170. osipov_vasilyi@mail.ru; 14 линия, 39, 199178, Санкт-Петербург, Россия; р.т.: +7(812) 328-08-87.

Кулешов Сергей Викторович — д-р техн. наук, главный научный сотрудник, лаборатория автоматизации научных исследований, СПИИРАН, СПб ФИЦ РАН. Область научных интересов: ассоциативно-онтологический подход к анализу интернет-контента, цифровые программно-определяемые инфокоммуникационные системы, обработка изображений и видеоданных, сжатие данных, обработка текстов, поисковые системы. Число научных публикаций — 130. kuleshov@iias.spb.su; 14 линия, 39, 199178, Санкт-Петербург, Россия; р.т.: 3235139.

Зайцева Александра Алексеевна — канд. техн. наук, старший научный сотрудник, лаборатория автоматизации научных исследований, СПИИРАН, СПб ФИЦ РАН. Область научных интересов: методы и технологии обработки больших данных,

обработка текстов, поисковые системы. Число научных публикаций — 89. cher@iiias.spb.su; 14 линия, 39, , Санкт-Петербург, Россия; р.т.: +78123283311.

Аксенов Алексей Юрьевич — канд. техн. наук, старший научный сотрудник, лаборатория автоматизации научных исследований, СПИИРАН, СПб ФИЦ РАН. Область научных интересов: цифровая обработка сигналов, методы обработки и компрессии 3D-данных, в том числе полученных с помощью 3D-сканеров, поисковые системы. Число научных публикаций — 43. a_aksenov@iiias.spb.su; 14 линия, 39, 199178, Санкт-Петербург, Россия; р.т.: +7 (812) 323 51 39.

Поддержка исследований. Работа выполнена при финансовой поддержке РФФИ (проект № 20-04-60455) и бюджетной темы №0073-2019-0005.

V. OSIPOV, S. KULESHOV, A. ZAYTSEVA, A. AKSENOV
**APPROACH FOR THE COVID-19 EPIDEMIC SOURCE
LOCALIZATION IN RUSSIA BASED ON MATHEMATICAL
MODELING**

Osipov V., Kuleshov S., Zaytseva A., Aksenov A. Approach for the COVID-19 Epidemic Source Localization in Russia Based on Mathematical Modeling.

Abstract. The paper presents the results of statistical data from open sources on the development of the COVID-19 epidemic processing and a study carried out to determine the place and time of its beginning in Russia. An overview of the existing models of the processes of the epidemic development and methods for solving direct and inverse problems of its analysis is given. A model for the development of the COVID-19 epidemic via a transport network of nine Russian cities is proposed: Moscow, St. Petersburg, Nizhny Novgorod, Rostov-on-Don, Krasnodar, Yekaterinburg, Novosibirsk, Khabarovsk and Vladivostok. The cities are selected both by geographic location and by the number of population. The model consists of twenty seven differential equations. An algorithm for reverse analysis of the epidemic model has been developed. The initial data for solving the problem were the data on the population, the intensity of process transitions from one state to another, as well as data on the infection rate of the population at given time moments. The paper also provides the results of a detailed analysis of the solution approaches to modeling the development of epidemics by type of model (basic SEIR model, SIRD model, adaptive behavioral model, modified SEIR models), and by country (in Poland, France, Spain, Greece and others) and an overview of the applications that can be solved using epidemic spread modeling. Additional environmental parameters that affect the modeling of the spread of epidemics and can be taken into account to improve the accuracy of the results are considered. Based on the results of the modeling, the most likely source cities of the epidemic beginning in Russia, as well as the moment of its beginning, have been identified. The reliability of the estimates obtained is largely determined by the reliability of the statistics used on the development of COVID-19 and the available data on transportation network, which are in the public domain.

Keywords: mathematical modeling, COVID-19, inverse analysis problem solution, forecasting, SEIR models, epidemic spread modeling.

Osipov Vasily — Professor, Director of SPIIRAS, SPIIRAS, SPC RAS. Research interests: mathematical modeling, intelligent systems, neural networks, information security. The number of publications — 170. osipov_vasily@mail.ru; 39, 14 line, 199178, St. Petersburg, Russia; office phone: +7(812) 328-08-87.

Kuleshov Sergey — Ph.D., Chief researcher, Laboratory of research automation, SPIIRAS, SPC RAS. Research interests: associative-ontological approach to the analysis of Internet content, digital software-defined infocommunication systems, image and video data processing, data compression, text processing, search engines. The number of publications — 130. kuleshov@iias.spb.su; 39, 14th line, 199178, St. Petersburg, Russia; office phone: 3235139.

Zaytseva Alexandra — Ph.D., Senior researcher, Laboratory of research automation, SPIIRAS, SPC RAS. Research interests: methods and technologies for processing big data, word processing, search engines. The number of publications — 89. cher@iias.spb.su; 39, 14 Line, , St. Petersburg, Russia; office phone: +78123283311.

Aksenov Alexey — Ph.D., Senior researcher, Laboratory of research automation, SPIIRAS, SPC RAS. Research interests: digital signal processing, methods of processing and compression of 3D data, including, search engines. The number of publications — 43. a_aksenov@iias.spb.su; 39, 14 line, 199178, St. Petersburg, Russia; office phone: +7 (812) 323 51 39.

Acknowledgements. This research is supported by RFBR (grant 20-04-60455) and by the budget (project no.0073-2019-0005).

References

1. Covid-19. URL: <https://ru.wikipedia.org/wiki/Covid-19> (дата обращения: 07.06.2021).
2. Tak gde zhe voznik Covid-19? [So where did Covid-19 originate?] Available at: <https://news-front.info/2021/05/03/tak-gde-zhe-voznik-covid-19> (accessed: 07.06.2021). (In Russ.).
3. NIH study offers new evidence of early SARS-CoV-2 infections in U.S. URL: <https://www.nih.gov/news-events/news-releases/nih-study-offers-new-evidence-early-sars-cov-2-infections-us> (дата обращения: 07.06.2021).
4. Svoy put': ves' god mir nablyudal za strategiyey Shvetsii po bor'be s koro-navirusom. [Their path: The world has watched Sweden's strategy to combat Coronavirus all year]. Available at: <https://tjournal.ru/analysis/249657-svoy-put-ves-god-mir-nablyudal-za-strategiyey-shvecii-po-borbe-s-koronavirusom-chto-iz-etogo-vyshlo> (accessed: 07.06.2021). (In Russ.).
5. Sweden tells citizens to wear masks on public transport as it struggles with COVID-19 resurgence URL: <https://www.euronews.com/2020/12/18/sweden-tells-citizens-to-wear-masks-on-public-transport-as-it-struggles-with-covid-19-resu> (дата обращения: 07.06.2021).
6. Sweden set to begin easing coronavirus restrictions. URL: <https://www.ft.com/content/c64b7e37-f00b-4e74-acf8-49c039c57698> (дата обращения: 07.06.2021).
7. World Health Organization. URL: <https://www.who.int/> (дата обращения: 16.08.2021).
8. Krivorotko O.I., Kabanikhin S.I., Zyatkov N. Yu. Matematicheskaya model' rasprostraneniya koronavirusa COVID-19 v Rossiyskoy Federatsii. [Mathematical model of the spread of coronavirus COVID-19 in the Russian Federation]. Available at: <https://www.nsu.ru/n/mathematics-mechanics-department/documents/> (accessed: 07.06.2021). (In Russ.).
9. Medrek M., Pastuszak Z. Numerical simulation of the novel coronavirus spreading. Expert Systems with Applications. 2021. 166. 114109.
10. Katris C. A time series-based statistical approach for outbreak spread forecasting: Application of COVID-19 in Greece. A time series-based statistical approach for outbreak spread forecasting: Application of COVID-19 in Greece. Expert Systems with Applications. 2021. Vol. 166. 114077.
11. Nadella P., Swaminathan A., Subramanian S.V. Forecasting efforts from prior epidemics and COVID-19 predictions. Eur J Epidemiol. 2020. Vol. 35. pp. 727–729.
12. Levashkin S.P., Zakharova O.I., Kuleshov S.V., Zaytseva A.A. Adaptive-compartmental model of coronavirus epidemic and its optimization by the methods of artificial intelligence. Journal of Physics: Conference Series. 2021. vol. 1864.
13. SEIR and SEIRS models. Available at: <https://docs.idmod.org/projects/emod-environmental/en/latest/model-seir.html> (accessed: 07.06.2021).
14. Bjornstad O.N. Epidemics, Use R. Springer Nature Switzerland AG. 2018. 312 p.

15. Kermack W.O., McKendrick A.G. A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*. 1927. Vol. 115(772). pp. 700–721.
16. Baroyan O.V., Rvachev L.A. *Matematika i epidemiologiya* [Mathematics and Epidemiology]. M.: Znaniye. 1977. P. 64.
17. Desyatkov B.M., Borodulin A.I., Kotlyarova S.S., Lapteva N.A., Marchenko M.Yu., Shabanov A.N. *Matematicheskoye modelirovaniye epidemicheskikh protsessov i otsenka ikh statisticheskikh kharakteristik* [Mathematical modeling of epidemic processes and assessment of their statistical characteristics]. *Khimicheskaya i biologicheskaya bezopasnost'* [Chemical and biological safety]. 2009. No 1–3. pp. 43–45. Available at: http://www.cbsafety.ru/rus/saf43_2009_st_2.pdf (accessed: 16.08.2021). (In Russ).
18. Xu H., Gonzalez M. J., Guo L., et al. Knowledge, Awareness, and Attitudes Towards COVID-19 Pandemic Among Different Populations in Central China: A Cross-Sectional Survey. *Journal of Medical Internet Research*. 2020.
19. Wuqiong L., Wee P.T., Mei L. Identifying Infection Sources and Regions in Large Networks. 2021. Available at: <https://arxiv.org/pdf/1204.0354.pdf> (accessed: 07.06.2021).
20. Valdano E. et al. Analytical computation of the epidemic threshold on temporal networks. *Physical Review X*. 2015. Vol. 5. No. 2. 021005.
21. Morris S.E., Freiesleben de Blasio B., Viboud C., Wesolowski A., Bjørnstad O.N., Grenfell B.T. Analysis of multi-level spatial data reveals strong synchrony in seasonal influenza epidemics across Norway, Sweden, and Denmark. *PLoS ONE*. 2018. № 13(5).
22. Holko A., Medrek M., Pastuszak Z., Phusavat K. Epidemiological modeling with a population density map-based cellular automata simulation system. *Expert Systems With Applications*. 2016. Vol. 48. pp. 1–8.
23. Ahmed E., Agiza H.N. On modeling epidemics including latency, incubation and variable susceptibility. *Physica A: Statistical Mechanics and its Applications*. 1998. Vol. 253. No. 1–4. pp. 347–352.
24. Newbold P., Granger C.W.J. Experience with forecasting time series and combination of forecasts. *Journal of the Royal Statistical Society*. 1974. № 137(2). pp. 131–165.
25. Fanelli D., Piazza F. Analysis and forecast of COVID-19 spreading in China, Italy and France. *Chaos, Solitons & Fractals*. 2020. 134. Article 109761.
26. Anastassopoulou C., Russo L., Tsakris A., Siettos C. Data-based analysis, modelling and forecasting of the COVID-19 outbreak. *PLoS ONE*. 2020. № 15(3). e0230405.
27. Wu J.T., Leung K., Leung G.M. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: A modelling study. *The Lancet*. 2020. Vol. 395 (10225). pp. 689–697.
28. Li Q., Guan X., Wu P., Wang X., Zhou L., et al. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *New England Journal of Medicine*. 2020. URL: [https://valiasr.zums.ac.ir/files/i_management/files/nejmoa2001316_\(1\).pdf](https://valiasr.zums.ac.ir/files/i_management/files/nejmoa2001316_(1).pdf) (дата обращения: 16.08.2021).
29. Hu Z., Qiyang G., Shudi L., Li J., & Momiao X. Artificial Intelligence Forecasting of Covid-19 in China. arXiv preprint. 2020. arXiv:2002.07112. URL: <https://arxiv.org/abs/2002.07112> (дата обращения: 16.08.2021)
30. Macdonald J.C., Browne C., Gulbudak H. Modelling COVID-19 outbreaks in USA with distinct testing, lockdown speed and fatigue rates. *Royal Society Open Science*. 2021. Aug 4; 8(8):210227.
31. Sarkar K., Khajanchi S, Nieto J. Modeling and forecasting the COVID-19 pandemic in India. *Chaos Solitons Fractals*. 2020. Oct. 139. 110049.

32. da Silva R.G, Ribeiro M.H.D.M., Mariani V.C., Coelho L.D.S. Forecasting Brazilian and American COVID-19 cases based on artificial intelligence coupled with climatic exogenous variables. *Chaos Solitons Fractals*. 2020. Oct. 139. 110027.
33. Sharma N., Kumar A., Kumar A. Spatial Network based model forecasting transmission and control of COVID-19.
34. Kol'tsova E.M., Kurkina Ye.S., Vasetskiy A.M. Matematicheskoye modelirovaniye rasprostraneniya epidemii koronavirusa Covid-19 v ryade yevropeyskikh, aziatskikh stran, Izraile i Rossii [Mathematical modeling of the spread of the Covid-19 coronavirus epidemic in a number of European, Asian countries, Israel and Russia]. *Problemy ekonomiki i yuridicheskoy praktiki* [Problems of Economics and Legal Practice]. 2020. Vol. XVI. No.2. pp. 154-165. (In Russ).
35. Obesnyuk V.F. Dinamika lokal'noy epidemicheskoy vspyshki COVID-19 cherez prizmu kompartment-modelirovaniya [Dynamics of the local epidemic outbreak of COVID-19 through the prism of compartment modeling]. *Analiz riska zdorov'yu* [Health risk analysis]. 2020. Vol. 2. pp. 83-91. 36. Kumar A. Modeling geographical spread of COVID-19 in India using network-based approach. URL: <https://www.medrxiv.org/content/10.1101/2020.04.23.20076489v1> (дата обращения: 16.08.2021).
36. Haslag P.H., Weagley D. From L.A. to Boise: How Migration Has Changed During the COVID-19 Pandemic (March 26, 2021). URL: <https://ssrn.com/abstract=3808326> or <http://dx.doi.org/10.2139/ssrn.3808326> (дата обращения: 07.06.2021).
37. Das D. Regional disparities of growth and internal migrant workers in informal sectors in the age of COVID-19.
38. Kojaku S., Hébert-Dufresne L., Mones E. et al. The effectiveness of backward contact tracing in networks. *Nat. Phys*. 2021. 17, pp. 652–658.
39. Klinkenberg D., Fraser C., Heesterbeek H. The effectiveness of contact tracing in emerging epidemics. *PLoS one*. 2006. Vol. 1. No. 1. pp. e12.
40. Müller J., Kretzschmar M., Dietz K. Contact tracing in stochastic and deterministic epidemic models. *Mathematical biosciences*. 2000. Vol. 164. No. 1. pp. 39-64.
41. Ball F. G., Knock E. S., O'Neill P.D. Threshold behaviour of emerging epidemics featuring contact tracing *Advances in Applied Probability*. 2011. Vol. 43. No. 4. pp. 1048-1065.
42. Browne C., Gulbudak H., Webb G. Modeling contact tracing in outbreaks with application to Ebola. *Journal of theoretical biology*. 2015. Vol. 384. pp. 33-49.
43. Okolie A., Müller J. Exact and approximate formulas for contact tracing on random trees. *Mathematical biosciences*. 2020. Vol. 321. pp. 108320.
44. Bianconi G. et al. Message-passing approach to epidemic tracing and mitigation with apps. *Physical Review Research*. 2021. Vol. 3. No. 1. pp. L012014.
45. Endo A. Centre for the Mathematical Modelling of Infectious Diseases COVID-19 Working Group, Leclerc QJ, et al. Implication of backward contact tracing in the presence of over dispersed transmission in COVID-19 outbreaks. *Welcome Open Res*. 2021. Vol. 5:239. URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7610176/> (дата обращения: 07.06.2021).
46. Coronavirus. URL: <https://coronavirus-monitor.info> (дата обращения: 07.06.2021).

А.В. Смирнов, Е.Г. Молл, Н.Н. Тесля
**ИСПОЛЬЗОВАНИЕ НЕЧЕТКИХ КОАЛИЦИОННЫХ ИГР ПРИ
ПРИНЯТИИ СОЦИАЛЬНО ОРИЕНТИРОВАННЫХ РЕШЕНИЙ
ПРИ ГОСПИТАЛИЗАЦИИ В УСЛОВИЯХ ПАНДЕМИИ**

Смирнов А.В., Молл Е.Г., Тесля Н.Н. Использование нечетких коалиционных игр при принятии социально ориентированных решений при госпитализации в условиях пандемии.

Аннотация. Проблемы организации медицинской помощи в условиях пандемии COVID-19, связанные с неопределенностью и ограниченностью различных ресурсов, привели к необходимости совершенствования систем принятия решений при госпитализации пациентов. С помощью ситуационного управления можно улучшить процесс принятия решений, чтобы он лучше соответствовал текущей ситуации. При этом важным становится учет влияния психологических факторов на решения, принимаемые при госпитализации. В статье предлагается использование коалиционных игр для ситуационного управления при госпитализации больных. Игроками и участниками коалиции являются госпитали, бригады скорой помощи, пациенты и центры компьютерной томографии. Цель игры - сформировать коалицию участников, обеспечивающую максимальную выгоду по времени и стоимости госпитализации в момент принятия решения. Рассмотрены общая схема госпитализации, основные источники информации о ситуации, постановка и формализация проблемы. Проведен эксперимент, в котором проверялось формирование коалиции во время госпитализации на основе данных, полученных при анализе динамики пандемии COVID-19. В связи с малым объемом данных и отсутствием апробированных моделей развития ситуации при проведении расчета часть параметров была оценена с использованием эвристических моделей развития ситуации, основанных на анализе информации из открытых источников информации. Результат эксперимента содержит набор коалиций, обеспечивающих максимальную выгоду при указанных ограничениях. При этом время расчета коалиционной игры позволяет использовать предложенную модель поддержки принятия решений при госпитализации в диспетчерской службе станций скорой помощи.

Ключевые слова: госпитализация, принятие решений, пандемия, нечеткие множества, коалиционная игра.

1. Введение. Любая эпидемия требует от сектора здравоохранения максимально эффективной организации и консолидации ресурсов для борьбы с ней. В случае заболеваний с длительной историей, таких, как грипп, туберкулез, MERS, SARS, к настоящему времени уже разработаны и испытаны решения, снижающие негативное воздействие как с точки зрения смертности, так и с точки зрения общего воздействия на экономику и функционирование государства. Эти решения включают диагностику заболевания, госпитализацию, протоколы лечения с учетом большого количества возможных вариаций и осложнений. Однако эпидемия коронавируса COVID-19 показала, что при появлении ранее неизвестного заболевания имеющихся стандартных

решений может быть недостаточно, и во многих случаях требуется быстрое принятие принципиально новых решений на основе не только имеющегося опыта, но и анализа текущей ситуации.

В этой работе основное внимание уделяется госпитализации пациента в условиях неопределенности, вызванной пандемией. Для формирования решений, согласующихся с оперативной ситуацией предлагается учитывать влияние неопределенности на психологическое состояние человека, поскольку, согласно исследованиям, в условиях неопределенности и постоянного стресса могут допускаться ошибки при принятии решения о госпитализации [1]. В качестве подтверждения данного тезиса можно рассмотреть пример двух европейских стран, которые одними из первых столкнулись со вспышкой коронавируса. В Испании (рис. 1) после получения первых данных о COVID-19 массовая госпитализация была предпринята еще до роста количества подтвержденных случаев [2]. Как известно из данных, доступных в открытом доступе, это привело к серьезной перегрузке системы здравоохранения и увеличению общей смертности, в том числе от COVID-19 [3]. График также показывает, что во время второй волны был учтен опыт первой и количество госпитализаций было значительно меньше, что также оказало положительное влияние на динамику смертности. Кардинально противоположная ситуация наблюдалась в тот же период в Германии. Имеющиеся данные показывают, что несмотря на развитую систему здравоохранения, количество госпитализаций в Германии было значительно меньше, чем в Испании, что позволило системе здравоохранения адекватнее реагировать на пандемию, что выражается в низкой смертности от COVID-19 (рис. 2) [2]. Сравнивая общий стресс в первой и второй волнах, можно увидеть, что при правильном ведении и принятии обоснованных решений о госпитализации общее психологическое состояние как врачей, так и пациентов позволяет им более адекватно реагировать на решение о госпитализации.

При отсутствии готовых решений становится актуальным поддержка принятия решений с учетом контекста текущей ситуации. В статье предлагается рассмотреть одну из задач, связанную с поддержкой принятия решения о госпитализации пациентов во время пандемии. Подход, предлагаемый для решения данной задачи, основан на использовании теории кооперативных игр [4], где игроками являются пациенты, машины скорой помощи, госпитали и центры компьютерной томографии. Решением игры является коалиция игроков, обеспечивающая максимальную выгоду по времени и стоимости госпитализации.

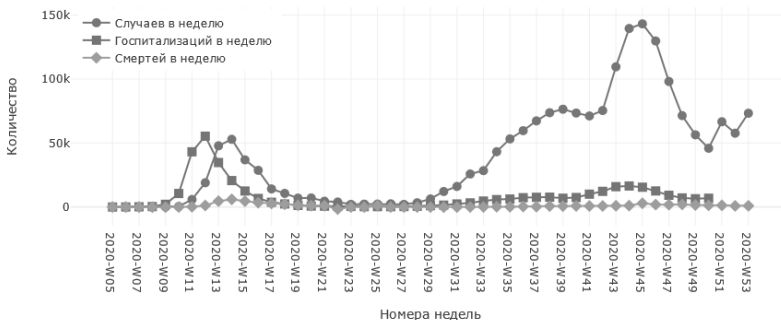


Рис. 1. Динамика распространения коронавирусной инфекции в Испании

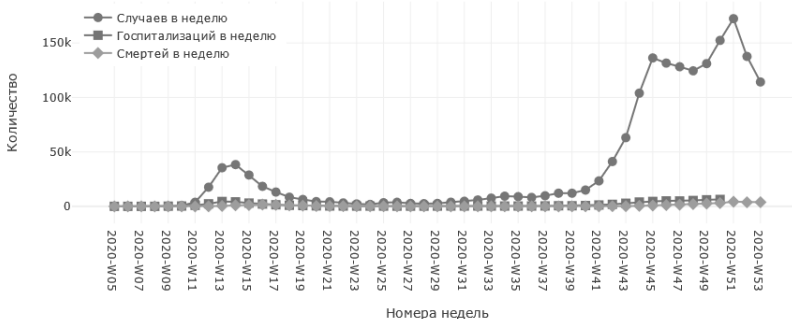


Рис. 2. Динамика распространения коронавирусной инфекции в Германии

Расчет каждой игры происходит как на основе количественных параметров, таких как наличие свободных мест (коек) в госпитале, среднее время госпитализации, количество машин скорой помощи, стоимость лечения, так и качественных параметров, связанных с психологическими факторами поведения пациентов, диспетчеров скорой помощи и медицинского персонала. Для расчета игры используется нечеткая функция принадлежности игрока к коалиции, позволяющая учесть нечеткую формализацию параметров игроков, связанных с динамикой психологического состояния при выборе стратегий их действия.

Статья построена следующим образом. Раздел 2 описывает текущее состояние исследований по использованию информационных технологий для поддержки управления ситуацией и принятия решений в здравоохранении. В разделе 3 описана постановка проблемы госпитализации в условиях пандемии с учетом психологических факторов для пациентов, бригад машин скорой помощи, медицинского персонала.

ла и персонала центра компьютерной томографии и проведен анализ данных по госпитализации, предоставленных СПб МИАЦ. Раздел 4 содержит формализацию коалиционной игры для процесса госпитализации и параметры, которые следует учитывать при принятии соответствующего решения. В разделе 5 приводится пример формирования коалиции с коалиционной игрой, использующий эвристики для расчета физического и психологического состояния пациента. В заключительном разделе дается общий вывод и направления дальнейшей работы.

2. Текущее состояние исследований. Использование информационных технологий в здравоохранении в настоящее время является активным и перспективным направлением исследований. Спектр их применения очень широк: от хранения медицинских карт, в которых содержатся все параметры пациента и его история здоровья [5], до использования искусственного интеллекта при выявлении диагнозов и поиска лекарств от конкретного заболевания у конкретного пациента [6-7].

В последние годы значительные прорывы были также сделаны в области телемедицины. Разработаны компактные носимые устройства, оснащенные датчиками для измерения основных жизненно важных параметров (частота сердечных сокращений, артериальное давление, сатурация крови кислородом, компактные устройства ЭКГ и др.) [8], параметров физической активности (фитнес-трекеры) [9], а также информационные системы для сферы здравоохранения, собирающие и анализирующие данные с носимых устройств. Развитие систем видеосвязи обеспечило возможность дистанционного осмотра и консультации лечащего врача (телемедицина). Получив доступ к накопленной информации, а также при наличии качественного канала связи с пациентом, врач может без вреда для своего здоровья оказать помощь, не требующую физического контакта, или удаленно проконсультировать родственников или опекунов больного по необходимым физическим манипуляциям, и сформировать рекомендации по лечению или поддержанию здоровья [10]. Ожидается, что использование носимых устройств также окажет значительную помощь в текущей ситуации с коронавирусом [11-12].

Поддержка принятия решений о госпитализации в условиях пандемии включает анализ регуляторов поведения человека в контексте эпидемий и пандемий. Координация решений сосредоточена вокруг влияния регуляторов поведения в COVID-19, которые включают юридические (жестко фиксированные, общеобязательные), организационные и профессиональные (положения, правила, принятые в опре-

деленном профессиональном сообществе), организационные и культурные нормы (подразумеваются, характеристика отдельных организаций, например, госпиталей и служб), а также индивидуальные нормы (моральные и психологические). Правовые нормы проще всего формализовать в виде теоретико-множественной или логической модели. Анализ организационных и профессиональных норм требует дополнительных исследований, поскольку некоторые из них носят рекомендательный характер и могут иметь различное толкование. Например, Британская медицинская ассоциация (ВМА) опубликовала этические принципы в отношении пандемии COVID-19. Все решения о распределении ресурсов должны быть «разумными при любых обстоятельствах, основанными на наилучших доступных клинических данных и суждениях, основанных на четких этических принципах и обоснованиях, лучше всего согласованных заранее, учитывая, что эти решения могут потребовать срочного пересмотра в меняющихся обстоятельствах» [13].

Использование нечеткой логики и кооперативных игр для описания взаимодействия участников коалиции - относительно новый подход, который, однако, показал свою эффективность в задачах конфигурации цепочки поставок [14] и формирования коалиции [15]. С медицинской точки зрения аппарат нечеткой логики и кооперативных игр может быть использован для оценки эффективности госпиталей [16]. Эффективность этой работы оценивается по большому количеству параметров, в том числе по количеству и качеству персонала (врачи, медсестры, вспомогательный персонал), количеству коек, количеству операций, затратам на лечение и обслуживание.

3. Постановка проблемы. Формальные модели, предназначенные для решения задач распределения ресурсов, на практике часто неприменимы, поскольку не учитывают человеческий фактор, который во многом влияет на результаты работы системы поддержки принятия решений о госпитализации в условиях пандемии.

В исследованиях, проводимых в контексте пандемии COVID-19, отмечается, что основное влияние на людей оказывает незавершенность угрожающей ситуации, а также фиксируется кумулятивный характер стрессового воздействия [1]. Неблагоприятными фоновыми факторами для всех участников процесса принятия решений являются тревожность, страх заражения, вынужденная изоляция от членов семьи [17], нарастание тревожности и депрессивных состояний [18]. На поведение пациентов влияют стресс, страх, депрессия, вызванные самим заболеванием и его тяжестью; а также стрессом, связанным с «кризисом здоровья» [19], который, во многом, обусловлен распро-

странением угрожающей, эмоциональной информации о невидимой угрозе, которая приводит к постоянному обращению людей к ее источникам, вызывает сомнения в том, что в этих условиях можно рассчитывать на полноценное и качественное медицинское обслуживание. Отсутствие средств защиты, проблемы с адекватным лечением, а также чисто человеческий фактор - переутомление, напряжение, тревога и профессиональное выгорание медицинского персонала [18, 20, 21] вызывают негативные эмоциональные состояния.

Процесс принятия решения о госпитализации требует предварительного сбора и обработки большого количества информации. Один из сценариев ситуационного управления во время госпитализации представлен на рисунке 3. На этом рисунке показана схема госпитализации, рекомендованная Департаментом здравоохранения Москвы в случае экстренного вызова пациента с симптомами COVID-19. На диаграмме показан не только протокол взаимодействия с пациентом и путь его обработки от постановки диагноза до госпитализации, но и взаимодействие с потенциально инфицированными людьми.

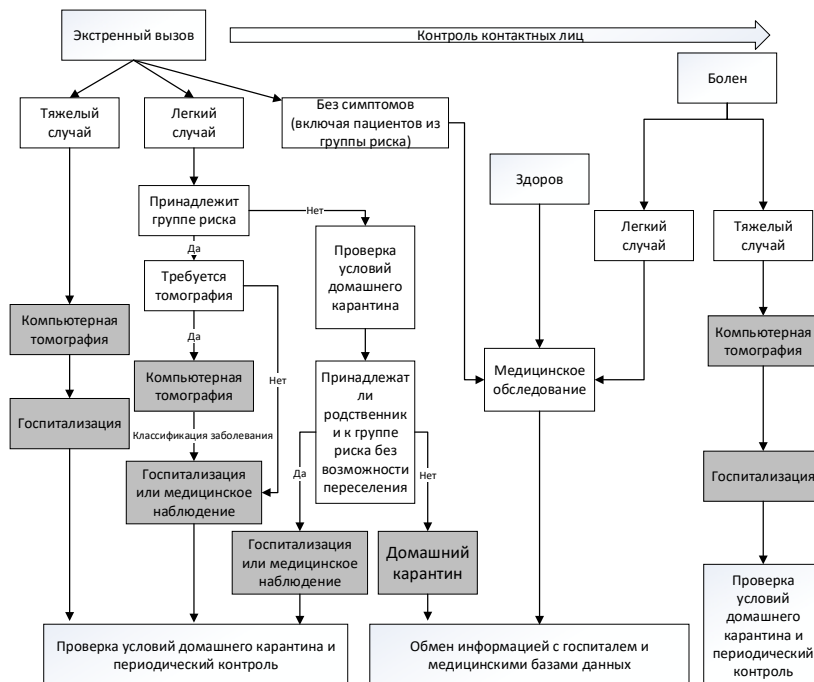


Рис. 3. Схема госпитализации пациентов с коронавирусной инфекцией, согласно рекомендациям Департамента здравоохранения г. Москва [22]

В каждом случае следует принимать взвешенное решение о госпитализации (на рисунке момент принятия решения показан серыми прямоугольниками). При этом, для каждого пациента необходимо учитывать ряд факторов, влияющих на окончательное решение о госпитализации:

1. Неравномерное распределение вызовов в службы экстренной помощи и госпитализаций по дням недели и времени суток, что, во многом, зависит от принятия решений пациентами и членами их семей.

2. Влияние ситуации на поведение медицинского персонала и пациентов в условиях неоднозначности описательных норм, то есть их видения того, как большинство окружающих их людей ведет себя или будет вести себя в таких условиях [23].

3. Особенно важно учитывать временные аспекты при принятии диагностических решений, поскольку при работе с новыми заболеваниями медицина становится плохо структурированной областью знаний. Это создает серьезные трудности при построении систем, связанных с процессами принятия решений. Как отмечают эксперты, «с одной стороны, врачи первичного звена не верят своим диагнозам, с другой - они соответствуют менталитету населения: пациенты требуют, чтобы они получали медицинские услуги в полном объеме и хотят дополнительно обследоваться в компьютерном томографе или быть госпитализированы» [24]. Следует учитывать факторы, влияющие на эффективность совместного принятия решений врачом и пациентом [25].

4. Существенным фактором, определяющим время приема пациентов в приемных покоях госпиталей, являются особенности корпоративной культуры медицинского учреждения, отражающие уровень их адаптации к условиям пандемии. В этом контексте можно предположить, что распределение госпиталей на три категории (три уровня адаптации к пандемии - высокий, средний и низкий) будет достаточным.

5. Особенности принятия решений в зависимости от характеристик пациентов. Врачи службы скорой помощи сталкиваются с «паническими» пациентами, количество которых в зависимости от общей ситуации составляет от 16% до 29%. Также учитываются возрастные особенности. Исследования в США во время первой волны пандемии показали, что молодые люди переоценивают риски COVID-19 для себя и других по сравнению со старшим поколением, что было интерпретировано как шок для молодежи [26].

Для проверки гипотез были собраны данные по госпитализациям с диагнозом COVID-19 в Санкт-Петербурге за период с 1.11.2020 по 31.12.2020. Эти данные включают в себя общую статистику по госпитализациям, а также детализацию по трем станциям скорой помощи. На рисунке 4 представлен график госпитализаций по Санкт-Петербургу, согласно данным, предоставленным СПб МИАЦ. Вертикальные линии отмечают начало недели (понедельник). Наглядно видна неравномерность распределения вызовов по дням недели. Это подтверждается и статистическим анализом, результаты которого представлены на рисунке 5 и в таблице 1. При анализе результатов, представленных в таблице 1 видно, что наименьшее количество госпитализаций в Санкт-Петербурге наблюдается в понедельник и воскресенье. При этом пик госпитализаций приходится на четверг.

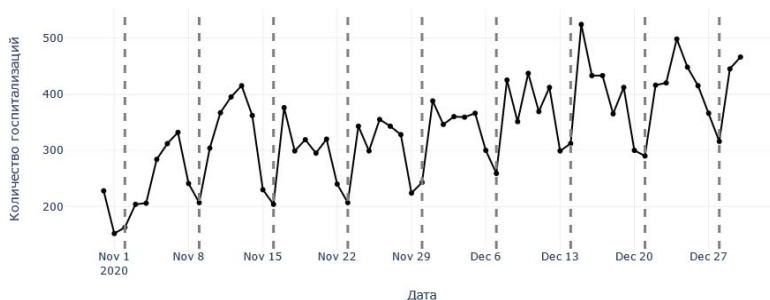


Рис. 4. Количество госпитализаций по дням недели в г. Санкт-Петербург за период 01.11 – 31.12.2020

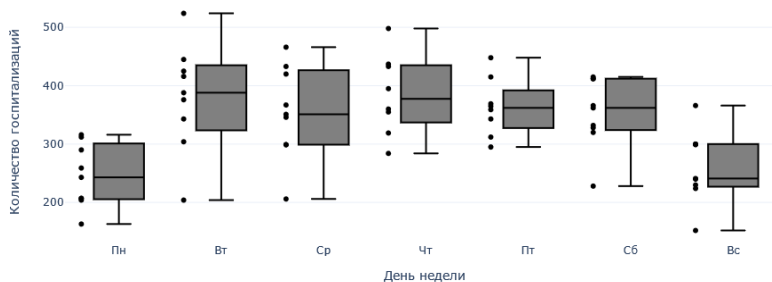


Рис. 5. Количество госпитализаций по дням недели в г. Санкт-Петербург за период 01.11 – 31.12.2020

Данные от станций скорой помощи содержат в себе дополнительную обезличенную информацию о поставленных диагнозах и характеристиках, включая время на принятие решения. Всего было получено 548 записей о гос-

питализациях по трем станциям скорой помощи. Далее представлен анализ зависимости времени принятия решения от пола, возраста и модели пациента (рис. 6-8).

Таблица 1. Статистические показатели госпитализации по дням недели

	Пн	Вт	Ср	Чт	Пт	Сб	Вс
Количество	9.0	9.0	9.0	8.0	8.0	9.0	9.0
Среднее	244.55	380.55	354.11	385.13	363.25	352.78	261.33
СКО	53.67	91.16	80.11	69.65	50.16	60.04	61.81
min значение	163.0	204.0	206.0	284.0	295.0	228.0	152.0
25% квартиль	207.0	343.0	299.0	346.0	335.25	328.0	230.0
50% квартиль	243.0	388.0	351.0	377.5	362.0	362.0	241.0
75% квартиль	290.0	425.0	420.0	434.0	380.5	412.0	300.0
max значение	316.0	524.0	466.0	498.0	448.0	415.0	366.0
% от min знач.	0	55.61	44.80	57.48	48.53	44.25	6.86

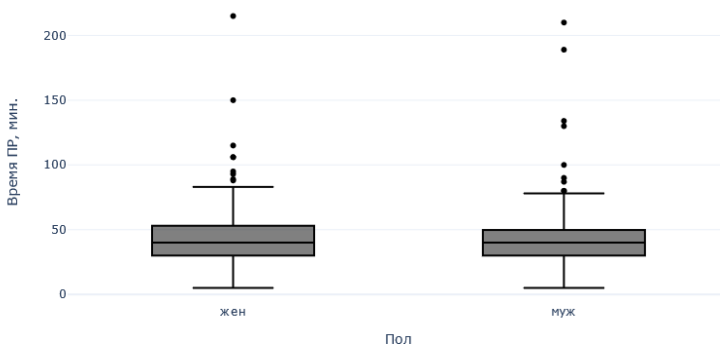


Рис. 6. Распределение времени принятия решения в зависимости от пола пациентов

На графиках видно, что пол и модель пациента практически не влияют на время принятия решения. Под моделью пациента в данном случае понимается комплексное описание состояния пациента, полученное по результатам осмотра [27]. Однако, что касается возрастной группы, то в ней наблюдается связь между временем принятия решения и возрастом. Множество пациентов можно разбить на 6 равных возрастных групп от 20 до 100 лет (0: [20-30), 1: [30-40), 2: [40-50), 3: [50-60), 4: [60-70), 5: [80-90), 6: [90-100)). На рисунке 8, с учетом такого

разбрения видно, что время принятия решения растет с увеличением возраста пациента, что должно быть принято во внимание при формировании описания кооперативной игры.

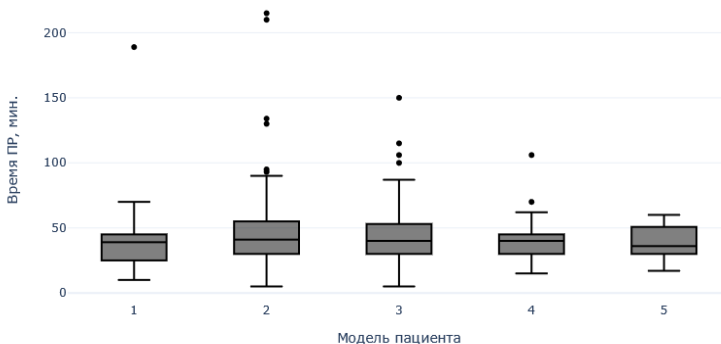


Рис. 7. Распределение времени принятия решения в зависимости от модели пациентов

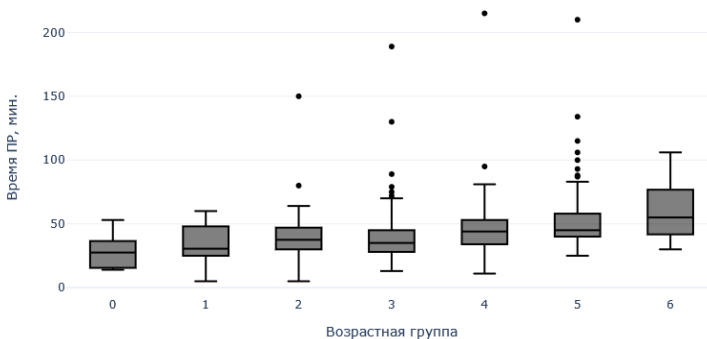


Рис. 8. Распределение времени принятия решения в зависимости от возрастной группы пациентов

Дополнительно был проведен анализ госпиталей относительно времени доставки пациента от места жительства до палаты. Всего было проанализировано 14 госпиталей. Результаты представлены на рисунке 9. По рисунку 9 видно, что госпитали могут быть разделены на три группы в зависимости от среднего времени доставки пациентов: от 0 до 100 минут, от 100 до 200 и от 200 до 300 минут.

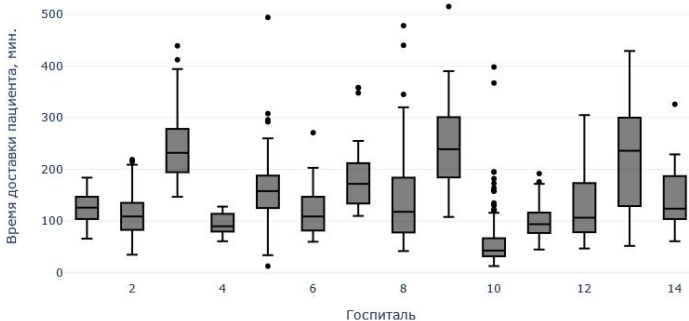


Рис. 9. Распределение госпиталей по времени доставки пациента

С учетом полученных результатов можно сформировать модель влияния человеческих факторов на принятие решения о госпитализации (рис. 10).

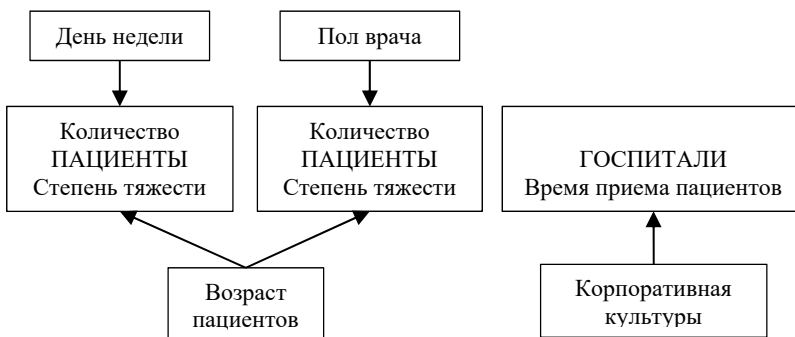


Рис. 10. Влияние человеческого фактора на принятие решения о госпитализации

4. Использование нечеткой кооперативной игры при госпитализации. В этом разделе рассматривается кооперативная игра для поддержки принятия решений при госпитализации в условиях пандемии. В кооперативной игре участвуют пациенты, машины скорой помощи, госпитали и центры компьютерной томографии. При старте расчета формируется набор параметров, которые становятся исходными данными для вычисления характеристической функции коалиции. Решение, полученное в ходе вычисления, будет содержать коалиции игроков, обеспечивающую максимальную выгоду с точки зрения времени и стоимости госпитализации. Также будет определен госпиталь,

в который необходимо доставить пациента, и центр компьютерной томографии, где при необходимости можно будет пройти дополнительное обследование.

Кооперативная игра представляется в виде множества (N, v) , где N – множество игроков, v – характеристическая функция игры (её общий выигрыш). Текущая ситуация описывается при помощи контекста C .

Определим в качестве игроков следующие подмножества: пациенты, нуждающиеся в госпитализации (N^P), персонал машин скорой помощи (N^A), госпитали (N^H), центры компьютерной томографии (КТ) (N^{CT}), $N = N^A \cup N^H \cup N^P \cup N^{CT}$. Для упрощения игры будем считать, что бригада скорой помощи уже прибыла к пациенту и необходимо решить, целесообразно ли проводить компьютерную томографию и выбрать наиболее подходящий центр КТ и госпиталь. У каждого игрока имеется набор стратегий, которые он может выбрать в зависимости от текущей ситуации и собственных поведенческих паттернов, что может рассматриваться как нечеткая функция принадлежности коалиции.

Пациенту соответствуют стратегии (η^P), определяющие согласие или отказ в госпитализации. Для госпиталя определены похожие стратегии (η^H), определяющие согласие или отказ в приеме пациента. Для бригады скорой помощи рассматриваются три стратегии (η^A): оставить пациента дома, транспортировка в центр КТ и транспортировка в госпиталь. Центр КТ описывается двумя стратегиями: (η^{CT}) прием или отказ в приеме пациента. Стратегией коалиции является сочетание стратегий ее членов $\eta_{i,j,k,l} = \{\eta_i, \eta_j, \eta_k, \eta_l\}$, где $i \in P, j \in H, k \in A, l \in CT$.

Функция принадлежности к коалиции (1) определяет вероятность того, будет ли входить участник в состав конкретной коалиции.

$$\eta(x) = \begin{cases} 0, & x < 0.5 \\ 1, & x > 0.5 \end{cases}, \quad (1)$$

где x – вероятность выбора стратегии игроком.

Игра описывается матрицей стоимостей Ω , каждый элемент которой $\omega_{i,j,k,l}$ содержит значение выигрышей участников коалиции i, j, k, l , с заданным набором стратегий.

Введем время обработки пациента $t_{i,j,k,l} = f(C^P, C^H, C^A, C^{CT})$, учитывающее выбранную стратегию $\eta_{i,j,k,l}$. Тогда выигрыш участников коалиции будет вычисляться согласно формуле (2):

$$\omega_{i,j,k,l} = \alpha^P \eta_{i,j,k,l} f^P(C^P, t_{i,j,k,l}, C'^P) + \alpha^H \eta_{i,j,k,l} f^H(C^H, t_{i,j,k,l}, C'^H) + \alpha^A \eta_{i,j,k,l} f^A(C^A, t_{i,j,k,l}, C'^A) + \alpha^{CT} \eta_{i,j,k,l} f^{CT}(C^{CT}, t_{i,j,k,l}, C'^{CT}) \quad (2)$$

где α – весовые коэффициенты. Поясним входящие в (2) параметры. Здесь $\eta_{i,j,k,l} f^P(C^P, t_{i,j,k,l}, C'^P)$ – это функция стоимости для пациента в зависимости от его текущего состояния C^P (физиологического и психологического) и прогнозируемого финального состояния C'^P . В функции также принимается во внимание время принятия решения и работы с пациентом $t_{i,j,k,l}$ при выборе стратегии $\eta_{i,j,k,l}$. Функция (3) расчета стоимости игры для пациента рассчитывается на основе следующих параметров:

$$f^P(C^P, t_{i,j,k,l}, C'^P) = \alpha_1 t + \alpha_2 PHI' + \alpha_3 PSY' \quad (3)$$

где t – прогнозируемое время транспортировки пациента до госпиталя, которое вычисляется как сумма нескольких составляющих: $t = t_{transp_{CT}} + t_{w_{CT}} + t_{proc_{CT}} + t_{transp_H} + t_{w_H} + t_{proc_H}$, среди которых $t_{transp_{CT}}$ – соответствующее времени транспортировки в центр КТ, $t_{w_{CT}}$ – время ожидания в очереди центра КТ, $t_{proc_{CT}}$ – время ожидания заключения специалиста КТ, t_{transp_H} – время транспортировки в госпиталь, t_{w_H} – время ожидания в очереди госпиталя, t_{proc_H} – время, затраченное на прием пациента в госпитале; $PHI' = f^{PHI}(f^{PHI^{-1}}(PHI) - t)$ соответствует прогнозируемому физиологическому состоянию, в котором PHI – обозначает физиологическое состояние, f^{PHI} – эвристическая функция, определяющая физиологическое состояние пациента через некоторое время t , α – весовые коэффициенты функции расчета стоимости. Предлагаемая эвристика оценивает физиологическое состояние пациента на основе основных параметров здоровья, характеризующих течение заболевания (частота сердечных сокращений, частота дыхания,

сатурация крови кислородом, давление, температура тела, возраст), объединяя их в интегральный показатель PHI , определяемый в диапазоне $[0,1]$; $PSY' = f^{PSY}(f^{PSY^{-1}}(PSY, PHI) - t, PHI')$ описывает прогнозируемое психологическое состояние, где PSY – это обозначение психологического состояния, f^{PSY} – эвристическая функция, определяющая психологическое состояние пациента по прошествии времени t . Предлагаемая эвристика оценивает психологическое состояние в диапазоне $[0,1]$. Значение функции зависит от текущего психологического состояния, а также текущего и прогнозируемого физиологического состояния.

Функция $f^H(C^H, t_{i,j,k,l}, C'^H)$ стоимости для госпиталя зависит от текущего состояния C^H (физиологическая и психологическая усталость персонала, запас расходных средств и медикаментов), свободно-го коечного фонда B^H , стоимости содержания и лечения одного пациента $Cost^H$, среднего времени ожидания в очереди на прием t_{wH} и прогнозируемое конечное состояние C'^H . Функция затрат госпиталя (4) определяется с помощью представленных параметров следующим образом:

$$f^H(C^H, t_{i,j,k,l}, C'^H) = Cost^H / B^H (1 + t_{wH}). \quad (4)$$

Также в функции (4) может быть использован тип госпиталя, в зависимости от которого принимается решение о госпитализации или необходимости дополнительного осмотра. Специализированный госпиталь может дообследовать пациента в собственном КТ-центре. В неспециализированный стационар можно принять пациента только при наличии результата компьютерной томографии, сделанной в КТ-центре.

Функция $f^A(C^A, t_{i,j,k,l}, C'^A)$ затрат бригады скорой помощи определяется в зависимости от текущего состояния C^A бригады (психологическая усталость персонала, состояние машины (запас топлива, время до техобслуживания), запас расходных средств), стоимости с учетом времени доставки пациента в КТ центр и в госпиталь $t_{i,j,k,l}$, а также прогнозируемого состояния C'^A . Значение данной функции рассчитывается на основе представленных параметров следующим образом (5):

$$f^A(C^A, t_{i,j,k,l}, C'^A) = \alpha_1 t_{i,j,k,l} + \alpha_2 (t_{i,j,k,l} \cdot Cost^A) + \alpha_3 PSY', \quad (5)$$

где Cost^A - стоимость часа работы бригады скорой помощи (включая затраты на персонал и оборудование); время работы с пациентом: $t_{i,j,k,l} = t_{\text{transpCT}} + t_{wCT} + t_{\text{procCT}} + t_{\text{transpH}} + t_{wH}$; прогнозируемое психологическое состояние бригады скорой помощи: $PSY' = f^{-1}(f(PSY)) - t_{i,j,k,l}$.

Функция $f^{CT}(C^{CT}, t_{i,j,k,l}, C'^{CT})$ стоимости для центра компьютерной томографии показывает зависимость от его текущего состояния C^{CT} (время подготовки к обследованию и время обследования), затраты с учетом времени лечения пациента t_{procCT} , и прогнозируемое состояние C'^{CT} . Функция стоимости центра КТ (6) определяется с помощью следующих параметров:

$$f^{CT}(C^{CT}, t_{i,j,k,l}, C'^{CT}) = \alpha_1(t_{wCT} + t_{\text{procCT}}) + \alpha_2(\text{Cost}^{CT})^{-1}, \quad (6)$$

где t_{wCT} - текущее время ожидания в КТ-центре; t_{procCT} - время обслуживания одного пациента, включая время подготовки к следующему пациенту, Cost^{CT} - стоимость компьютерной томографии.

В целом, для кооперативной игры характеристическая функция v может быть определена следующим образом (7):

$$v = \max_{m \in M} \omega_m \eta_m, M = |N^A| |N^H| |N^P| |N^{CT}|, \quad (7)$$

где M - количество возможных комбинаций игроков в игре, которые могут образовать коалицию (количество потенциальных коалиций). Далее рассмотрим описание функций стоимости для каждого из игроков.

5. Пример коалиционной игры. Для экспериментов были сформированы наборы игроков всех типов (пациенты, бригады скорой помощи, госпитали, центры компьютерной томографии). Для каждого игрока параметры были назначены случайным образом, в пределах оценок, полученных при анализе первой волны пандемии. Так, например, время ожидания бригады скорой помощи - от 10 до 20 минут, время в пути до госпиталя составляло от 10 до 240 минут, время обслуживания в приемном покое госпиталя – от 10 до 20 минут.

Для стационара стоимость лечения пациента определялась с учетом оценки, полученной в среднем для госпиталей, обслуживающих COVID-19, и составила 200 тысяч рублей на пациента [28]. Стоимость

компьютерной томографии легких в среднем по КТ-центрам была оценена в 4750 рублей. Вероятность подтверждения повреждения легких составила 30%.

Поскольку единой оценки физического состояния человека не существует, была построена эвристика, учитывающая восприятие физического состояния в зависимости от параметров пульса, частоты дыхания, насыщения кислородом крови, давления, температуры тела, возраст. При прогнозировании развития состояния использовалась обратная функция (8), позволяющая оценить время, в течение которого осуществится изменение физического состояния пациента до указанного значения:

$$f^{PHI^{-1}}(PHI) = \frac{5000 * (PHI + \frac{1}{16})^2}{21} - \frac{100 * (PHI + \frac{1}{16})}{3} + \frac{8}{7} \quad (8)$$

Функция (8) была сформирована в ходе обсуждения с медицинскими экспертами и является предварительным вариантом оценки физического состояния. При её построении были сформированы узловые точки состояния с учетом времени, которое прошло между изменениями физического состояния. Полученные точки были аппроксимированы с использованием ряда Тейлора до достижения точности 10^{-2} . В дальнейшем данная функция будет уточняться с использованием модели пациента.

Значение функции отображается в пять нечетких множеств, описывающих состояние человека. Эти множества соотнесены с лингвистическими переменными «ужасно», «плохо», «нормально», «хорошо», «идеально» и все вместе составляют диапазон $[0,1]$. Для приведения числового значения к лингвистической переменной используется аппарат нечетких множеств. Носителем каждого множества является значение в соответствующих интервалах ($[0; 0.25]$ - «очень плохо», $(0.15;0.45)$ - «плохо», $(0.35;0.65)$ - «нормально», $(0.55;0.85)$ - «хорошо», $(0.75;1]$ - «очень хорошо»).

Для оценки эмоционального состояния также была построена эвристическая функция (9), оценивающая время до изменения эмоционального состояния в зависимости от исходного эмоционального состояния и изменения физического состояния:

$$f^{PSY^{-1}}(PSY, PHI) = 12 * (PSY^2 + 1) * PHI \quad (9)$$

Функция (9) также является первым приближением на основе имеющихся данных, полученных в ходе обсуждения с экспертами. Аппроксимация осуществлялась с использованием ряда Тейлора и достижения точности 10^{-2} . Обе полученные эвристические функции использовались для оценки состояния пациентов и бригад скорой помощи.

Всего было сгенерировано по 10 игроков каждого типа. При решении проблемы использовалась модель C-core. Согласно модели C-core на каждом этапе решения множество возможных составов коалиций рассчитывалось таким образом, чтобы состав коалиции был устойчив к изменениям. Другими словами, в найденном множестве коалиций все выигрыши являются максимальными и любое изменение состава коалиции приводит к снижению выигрыша. По результатам расчета была сформирована большая коалиция, которая дает наибольшую общую пользу при госпитализации пациентов. На рисунке 11 показано распределение выигрышей для различных коалиций.

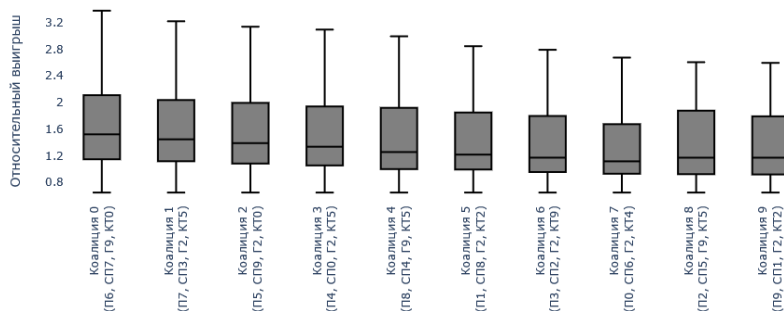


Рис. 11. Результат расчета коалиционной игры

Состав коалиции, дающей максимальную пользу в текущих условиях, указан в круглых скобках на горизонтальной оси, где П - пациент, СП - машина скорой помощи, Г - госпиталь, КТ - центр компьютерной томографии. Расчет коалиций производится итеративно, до тех пор, пока не были сформированы коалиции для госпитализации всех пациентов. Каждая коалиция оценивалась в параметрах относительного выигрыша, для чего были вычислены матрицы затрат для всех комбинаций участников, после чего из них были выбраны комбинации (составы коалиций) с максимальным выигрышем. На рисунке 11 представлено статистическое распределение выигрышей на каждом этапе расчетов и указаны максимальные и минимальные значения выигрышей. Каждый этап, отраженный столбцом на рисунке 11, соответствует расчету состава коалиции для одного пациента, а назначение

машин скорой помощи, пациентов и центров компьютерной томографии осуществляется в соответствии с составом коалиции, дающим максимальную пользу в конкретных условиях. В этом случае участники коалиции блокируются для следующих расчетов до поступления пациента в госпиталь.

Зависимость времени расчета от общего количества участников показана на рисунке 12. В качестве тестового стенда для расчета использовался компьютер на базе Intel Core i9 10900X 3,7 ГГц с 10 физическими ядрами, 64 ГБ оперативной памяти. Было использовано только одно ядро процессора без масштабирования на другие ядра. Код разработан с использованием языка программирования Python, библиотеки Pandas для представления массивов данных и результатов игры, Plotly для отображения результатов. В наихудшем случае, при полном переборе вариантов сложность задачи равна $O\left(\left|N^A\right|\left|N^H\right|\left|N^P\right|\left|N^{CT}\right|\right)$, где каждый N - количество участников соответствующего типа.

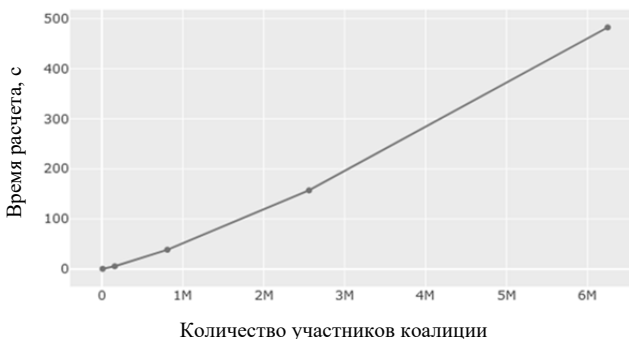


Рис. 12. Зависимость времени расчета от количества игроков

Однако реальная сложность расчетов меньше, поскольку для каждой следующей коалиции количество участников уменьшается за счет фиксации состава предыдущей коалиции и возможно использование эволюционных алгоритмов для расчета состава коалиции.

8. Заключение. В статье рассматриваются влияние различных факторов на процесс принятия решений при госпитализации в условиях пандемии. Всего было выявлено и проверено пять факторов. По результатам их анализа предложена формализация нечеткой коалиционной игры для поддержки принятия решений о госпитализации в условиях пандемии, которая учитывает не только факторы времени и затрат, но и психологический фактор. Представленная формализация

позволяет учитывать основные стратегии действий игроков и оценивать их психологическое состояние при выборе стратегий.

В ходе работы были собраны данные из новостных лент и сводок департаментов здравоохранения Москвы и Санкт-Петербурга. Однако следует отметить, что в настоящее время в открытом доступе недостаточно данных для точной оценки качества принятия решений в ситуационном управлении при использовании коалиционных игр для госпитализации. По этой причине в работе были использованы эвристические модели для оценки психологических и физиологических состояний. Тем не менее, проведенное моделирование расчета коалиционной игры для поддержки принятия решения при госпитализации показало применимость предложенного подхода. В будущем планируется сбор дополнительных данных и улучшение эвристических моделей.

Результаты моделирования показали возможность получения решения кооперативной игры за конечное время. Решением в данном случае являются множества составов коалиций, имеющих близкое распределение выигрышей. Из полученного множества затем выбирается конкретный состав коалиции, дающий максимальный выигрыш. В состав коалиции входит пациент, бригада скорой помощи, госпиталь и центр компьютерной томографии. Преимуществом предложенного подхода является использование метрики эффективности коалиции, учитывающей не только стоимостные и временные, но и социально-экономические характеристики участников при формировании коалиции для госпитализации. При этом в ходе оценки коалиции также формируется распределение выигрыша, который получит каждый из участников в ходе госпитализации. Скорость получения решения может быть повышена путем использования алгоритмов оптимизации, основанных, например, на эволюционных вычислениях. Для сравнения результатов расчета с уже существующими методами поддержки принятия решения может быть проведено сравнение получаемых решений с использованием метрик предложенного и существующих подходов. Данное исследование может быть одним из направлений последующей работы.

В работе также оценивалась вычислительная сложность совместной игры на основе полученных данных. В результате оценки было показано, что предложенная схема расчета состава коалиции обеспечивает получение решения в разумные сроки, что позволит использовать предложенную модель поддержки принятия решений при госпитализации в диспетчерской службе станций скорой помощи.

Последующие направления работы будут ориентированы на разработку функций оценки физиологического и психологического состояния с целью получения более точных результатов. Также планируется расширить возможности использования механизм коалиционных игр с определением наиболее выигрышной стратегии для каждого игрока при использовании для совместного принятия решений «врач-пациент». При получении данных из других стран или регионов может быть проведено дополнительное тестирование предложенной модели, что потребует учета культурных, административных, географических и экономических различий.

Литература

1. Xiong J. et al. Impact of COVID-19 pandemic on mental health in the general population: A systematic review. *Journal of Affective Disorders*. Elsevier B.V. 2020. Vol. 277. pp. 55–64.
2. European Centre for Disease Prevention and Control. Download COVID-19 datasets. 2021. Available: <https://www.ecdc.europa.eu/en/covid-19/data> (accessed: 31.01.2021).
3. Patel U. et al. Early epidemiological indicators, outcomes, and interventions of COVID-19 pandemic: A systematic review. *Journal of Global Health*. NLM (Medline). 2020. Vol. 10. № 2. P. 020506.
4. Mareš M. *Fuzzy Cooperative Games*. Heidelberg: Physica-Verlag HD. 2001. Vol. 72. P. 177.
5. Dalianis H. *Clinical Text Mining*. *Clinical Text Mining: Secondary Use of Electronic Patient Records*. Cham: Springer International Publishing. 2018. pp. 1–181.
6. Topol E.J. High-performance medicine: the convergence of human and artificial intelligence. *Nature Medicine*. Springer US. 2019. Vol. 25. № 1. pp. 44–56.
7. Yu K.-H., Beam A.L., Kohane I.S. Artificial intelligence in healthcare. *Nature Biomedical Engineering*. Springer US, 2018. Vol. 2. № 10. pp. 719–731.
8. Cosoli G., Spinsante S., Scalise L. Wrist-worn and chest-strap wearable devices: Systematic review on accuracy and metrological characteristics. *Measurement*. Elsevier B.V. 2020. Vol. 159. P. 107789.
9. Zhong C.-L., Li Y. Internet of things sensors assisted physical activity recognition and health monitoring of college students. *Measurement*. Elsevier B.V. 2020. Vol. 159. P. 107774.
10. Jamil F. et al. Towards a remote monitoring of patient vital signs based on iot-based blockchain integrity management platforms in smart hospitals. *Sensors (Switzerland)*. 2020. Vol. 20. № 8.
11. Armitage H. Stanford Medicine scientists hope to use data from wearable devices to predict illness, including COVID-19. *Stanford Medicine News Center*. 2020. Available: <http://med.stanford.edu/news/all-news/2020/04/wearable-devices-for-predicting-illness-.html> (accessed: 14-May-2020).
12. Radin J.M. et al. Harnessing wearable device data to improve state-level real-time surveillance of influenza-like illness in the USA: a population-based study. *The Lancet Digital Health*. Elsevier Ltd. 2020. Vol. 2. № 2. pp. e85–e93.
13. British Medical Association. COVID-19-ethical issues. A guidance note In brief. 2020.
14. Sheremetov L.B., Smirnov A.V. Supply Chain Configuration as a Cooperative Game with Fuzzy Coalitions. *Supply Chain Management Under Fuzziness*. Ed. Kahraman

- C., Öztayşi B. Berlin, Heidelberg: Springer Berlin Heidelberg. 2014. Vol. 313. pp. 293–314.
15. Mohebbi S., Li X. Coalitional game theory approach to modeling suppliers' collaboration in supply networks. *International Journal of Production Economics*. Elsevier, 2015. Vol. 169. pp. 333–342.
 16. Omrani H., Shafaat K., Emrouznejad A. An integrated fuzzy clustering cooperative game data envelopment analysis model with application in hospital efficiency. *Expert Systems with Applications*. Elsevier Ltd. 2018. Vol. 114. pp. 615–628.
 17. Shah K. et al. Focus on Mental Health During the Coronavirus (COVID-19) Pandemic: Applying Learnings from the Past Outbreaks. *Cureus*. Cureus, Inc. 2020.
 18. Montemurro N. The emotional impact of COVID-19: From medical staff to common people. *Brain, Behavior, and Immunity*. Academic Press Inc. 2020. Vol. 87. pp. 23–24.
 19. Garfin D.R., Silver R.C., Holman E.A. The novel coronavirus (COVID-2019) outbreak: Amplification of public health consequences by media exposure. *Health Psychology*. American Psychological Association Inc. 2020. Vol. 39. № 5. pp. 355–357.
 20. Li W. et al. Progression of Mental Health Services during the COVID-19 Outbreak in China. *International Journal of Biological Sciences*. NLM (Medline). 2020. Vol. 16. № 10. pp. 1732–1738.
 21. Mo Y. et al. Work stress among Chinese nurses to support Wuhan in fighting against COVID-19 epidemic. *Journal of Nursing Management*. NLM (Medline). 2020. Vol. 28. № 5. pp. 1002–1009.
 22. Департамент здравоохранения города Москвы. Приказ № 230 от 22.03.2020 «Об утверждении регламентов (алгоритмов) работы медицинских организаций, подведомственных Департаменту здравоохранения города Москвы в период с 23 по 30 марта 2020 г. по оказанию медицинской помощи пациентам, заболевшим новой коронавирусной инфекцией (COVID-19) и контактным с ним лицам».
 23. Schultz P.W. et al. The constructive, destructive, and reconstructive power of social norms: Research article. *Psychological Science*. SAGE Publications Sage CA: Los Angeles, CA. 2007. Vol. 18. № 5. pp. 429–434.
 24. Черкашин М. «Вторая волна: «Не надо стремиться в больницу!», РБК. 2020. URL: https://www.rbc.ru/spb_sz/12/12/2020/5fd483019a794782e6ceec26. (Дата обращения: 27.05.2021).
 25. Abrams E.M. et al. The Challenges and Opportunities for Shared Decision Making Highlighted by COVID-19. *Journal of Allergy and Clinical Immunology: In Practice*. American Academy of Allergy, Asthma and Immunology. 2020. Vol. 8. № 8. pp. 2474–2480.e1.
 26. Shleifer A. et al. Older People Are Less Pessimistic about the Health Risks of COVID-19 - HBS Working Knowledge. NBER Working Paper Series. 2020. Vol. 27494. Available: <http://www.nber.org/papers/w27494> (accessed: 31.01.2021).
 27. Шляхто Е.В. и др. Методические рекомендации, алгоритмы действия медицинских работников на различных этапах оказания помощи, чек-листы и типовые документы, разработанные на период наличия и угрозы дальнейшего распространения новой коронавирусной инфекции в Санкт-Петербурге. Версия 3.0 от 22.06.2020. Е.В. Шляхто, А.О. Конради, и др. // Санкт-Петербург, 2020. 187 с. URL: http://zdrav.spb.ru/media/filebrowser/covid-19_22.06.pdf (Дата обращения: 02.09.2021).
 28. Общество и пандемия: Опыт и уроки борьбы с COVID-19 в России // Москва: 2020. 774 с.

Смирнов Александр Викторович — д-р техн. наук, профессор, заслуженный деятель науки РФ, главный научный сотрудник, лаборатория интегрированных систем автоматизации, СПб ФИЦ РАН. Область научных интересов: конфигурирование систем, логистика знаний, поддержка принятия решений, социо-киберфизические системы. Число научных публикаций — 350. smir@iias.spb.su; 14 линия, 39, 199178, Санкт-Петербург, Россия; р.т.: +7 (812) 328 20 73.

Молл Елена Георгиевна — д-р психол. наук, профессор, директор, АНО Европейский центр подготовки руководителей "ЦНПР"ММК". Область научных интересов: психология ИКТ, организационное поведение, управление карьерой, корпоративная культура, лидерство. Число научных публикаций — 90. moll@neva.ru; пр. Просвещения, 33/2, 194295, Санкт-Петербург, Россия; р.т.: +7 (921) 993 37 67.

Тесля Николай Николаевич — канд. техн. наук, старший научный сотрудник, Лаборатория интегрированных систем автоматизации, СПб ФИЦ РАН. Область научных интересов: Управление знаниями, пространственно-временные знания, онтологии, интеллектуальные пространства, геоинформационные системы, технологии распределенных реестров. Число научных публикаций — 75. teslya@iias.spb.su; 14 линия, 39, 199178, Санкт-Петербург, Россия; р.т.: +7 (812) 328 80 71.

Поддержка исследований. Работа выполнена при финансовой поддержке РФФИ в соответствии с исследовательским проектом № 20-04-60054 (все разделы) и бюджетной темы № 0073-2019-0005 (частично по разделу 5). Сбор и анализ данных в разделе 3 проведены в рамках проекта РФФИ № 20-07-00455. Интеграция данных из разнородных источников в общий набор данных проведена в рамках проекта РФФИ № 20-07-00904.

A. SMIRNOV, E. MOLL, N. TESLYA
**USE OF FUZZY COALITION GAMES IN SOCIALLY ORIENTED
DECISION MAKING DURING HOSPITALIZATION IN
PANDEMIC**

Smirnov A., Moll E., Teslya N. Use of Fuzzy Coalition Games in Socially Oriented Decision Making During Hospitalization in Pandemic.

Abstract. The problems of organizing medical care in the context of the COVID-19 pandemic, associated with the uncertainty and limitedness of various resources, led to the need to improve decision-making systems for hospitalization of patients. Situational management can improve the decision-making process to fit the current situation better. At the same time, it becomes important to take into account the influence of psychological factors on decisions made during hospitalization. The paper proposes the use of coalition games for situational management during hospitalization of patients. The players and members of the coalition are hospitals, ambulance teams, patients and computed tomography centers. The goal of the game is to form a coalition of participants that provides the maximum benefit in terms of time and cost of hospitalization at the time of decision making. The general scheme of hospitalization, the main sources of information about the situation, the formulation and formalization of the problem are considered. An experiment was carried out in which the formation of a coalition during hospitalization was tested based on data obtained from analyzing the dynamics of the COVID-19 pandemic. Due to the small amount of data and the lack of approved models of the situation development, when carrying out the calculation, some of the parameters were estimated using heuristic models of the development of the situation, based on the analysis of information from open sources of information. The experiment result contains a set of coalitions that provide the maximum benefit under the specified constraints. At the same time, the calculation time of the coalition game allows using the proposed model of decision-making support during hospitalization in the dispatch service of ambulance stations.

Keywords: hospitalization, decision making, pandemic, fuzzy sets, coalition game.

Smirnov Alexander — Dr.Habil., Professor, Honored scientist of the Russian Federation, Chief Researcher, Laboratory of Computer-Aided Integrated Systems, SPC RAS. Research interests: systems configuration, knowledge logic, decision support, socio-cyber-physical systems. The number of publications — 350. smir@iias.spb.su; 39, 14th line, 199178, St. Petersburg, Russia; office phone: +7 (812) 328 20 73.

Moll Elena — Dr.Habil., Professor, Director, ANO European Management Training Center "ECMD MMK". Research interests: ICT psychology, organizational behavior, career management, corporate culture, leadership. The number of publications — 90. moll@neva.ru; 33/2, Prosveshcheniya Ave., 194295, St. Petersburg, Russia; office phone: +7 (921) 993 37 67.

Teslya Nikolay — Ph.D., Senior researcher, Laboratory of Computer-Aided Integrated Systems, SPC RAS. Research interests: Knowledge management, spatio-temporal knowledge, ontologies, intelligent spaces, geographic information systems, distributed ledger technologies. The number of publications — 75. teslya@iias.spb.su; 39, 14th line, 199178, St. Petersburg, Russia; office phone: +7 (812) 328 80 71.

Acknowledgements. This work was carried out with the financial support of the Russian Foundation for Basic Research in accordance with research project No. 20-04-60054 (all

sections) and budgetary theme No. 0073-2019-0005 (partly under section 5.) Data collection and analysis in section 3 was carried out within the framework of the project RFBR No. 20-07-00455. The integration of data from heterogeneous sources into a common data set was carried out within the framework of the RFBR project No. 20-07-00904.

References

1. Xiong J. et al. Impact of COVID-19 pandemic on mental health in the general population: A systematic review. *Journal of Affective Disorders*. Elsevier B.V. 2020. Vol. 277. pp. 55–64.
2. European Centre for Disease Prevention and Control. Download COVID-19 datasets. 2021. Available: <https://www.ecdc.europa.eu/en/covid-19/data> (accessed: 31.01.2021).
3. Patel U. et al. Early epidemiological indicators, outcomes, and interventions of COVID-19 pandemic: A systematic review. *Journal of Global Health*. NLM (Medline). 2020. Vol. 10. № 2. P. 020506.
4. Mareš M. *Fuzzy Cooperative Games*. Heidelberg: Physica-Verlag HD. 2001. Vol. 72. P. 177.
5. Dalianis H. *Clinical Text Mining*. Clinical Text Mining: Secondary Use of Electronic Patient Records. Cham: Springer International Publishing. 2018. pp. 1–181.
6. Topol E.J. *High-performance medicine: the convergence of human and artificial intelligence*. *Nature Medicine*. Springer US. 2019. Vol. 25. № 1. pp. 44–56.
7. Yu K.-H., Beam A.L., Kohane I.S. *Artificial intelligence in healthcare*. *Nature Biomedical Engineering*. Springer US, 2018. Vol. 2. № 10. pp. 719–731.
8. Cosoli G., Spinsante S., Scalise L. Wrist-worn and chest-strap wearable devices: Systematic review on accuracy and metrological characteristics. *Measurement*. Elsevier B.V. 2020. Vol. 159. P. 107789.
9. Zhong C.-L., Li Y. Internet of things sensors assisted physical activity recognition and health monitoring of college students. *Measurement*. Elsevier B.V. 2020. Vol. 159. P. 107774.
10. Jamil F. et al. Towards a remote monitoring of patient vital signs based on iot-based blockchain integrity management platforms in smart hospitals. *Sensors (Switzerland)*. 2020. Vol. 20. № 8.
11. Armitage H. Stanford Medicine scientists hope to use data from wearable devices to predict illness, including COVID-19. *Stanford Medicine News Center*. 2020. Available: <http://med.stanford.edu/news/all-news/2020/04/wearable-devices-for-predicting-illness-.html> (accessed: 14-May-2020).
12. Radin J.M. et al. Harnessing wearable device data to improve state-level real-time surveillance of influenza-like illness in the USA: a population-based study. *The Lancet Digital Health*. Elsevier Ltd. 2020. Vol. 2. № 2. pp. e85–e93.
13. British Medical Association. COVID-19-ethical issues. A guidance note In brief. 2020.
14. Sheremetov L.B., Smirnov A.V. Supply Chain Configuration as a Cooperative Game with Fuzzy Coalitions. *Supply Chain Management Under Fuzziness*. Ed. Kahraman C., Öztayşi B. Berlin, Heidelberg: Springer Berlin Heidelberg. 2014. Vol. 313. pp. 293–314.
15. Mohebbi S., Li X. Coalitional game theory approach to modeling suppliers' collaboration in supply networks. *International Journal of Production Economics*. Elsevier, 2015. Vol. 169. pp. 333–342.
16. Omrani H., Shafaat K., Emrouznejad A. An integrated fuzzy clustering cooperative game data envelopment analysis model with application in hospital efficiency. *Expert Systems with Applications*. Elsevier Ltd. 2018. Vol. 114. pp. 615–628.

17. Shah K. et al. Focus on Mental Health During the Coronavirus (COVID-19) Pandemic: Applying Learnings from the Past Outbreaks. Cureus. Cureus, Inc. 2020.
18. Montemurro N. The emotional impact of COVID-19: From medical staff to common people. *Brain, Behavior, and Immunity*. Academic Press Inc. 2020. Vol. 87. pp. 23–24.
19. Garfin D.R., Silver R.C., Holman E.A. The novel coronavirus (COVID-2019) outbreak: Amplification of public health consequences by media exposure. *Health Psychology*. American Psychological Association Inc. 2020. Vol. 39. № 5. pp. 355–357.
20. Li W. et al. Progression of Mental Health Services during the COVID-19 Outbreak in China. *International Journal of Biological Sciences*. NLM (Medline). 2020. Vol. 16. № 10. pp. 1732–1738.
21. Mo Y. et al. Work stress among Chinese nurses to support Wuhan in fighting against COVID-19 epidemic. *Journal of Nursing Management*. NLM (Medline). 2020. Vol. 28. № 5. pp. 1002–1009.
22. Departament Zdravoohraneniya goroda Moskvy. Prikaz № 230 ot 22.03.2020 «Ob utverzhenii reglamentov (algoritmov) raboty medicinskih organizacij, podvedomstvennyh Departamentu zdravoohraneniya goroda Moskvy v period s 23 po 30 marta 2020 g. po okazaniyu medicinskoj pomoshhi pacientam, zaboлевshim novoj koronavirusnoj infekciej (COVID-19), i kontaktnym s nim licam» [Department of Health of the city of Moscow. Order No. 230 of March 22, 2020 “On Approval of Regulations (Algorithms) for the Operation of Medical Organizations Subordinated to the Moscow City Health Department in the period from March 23 to March 30, 2020 to provide medical assistance to patients with a new coronavirus infection (COVID-19), and to contact persons”] (In Russ.)
23. Schultz P.W. et al. The constructive, destructive, and reconstructive power of social norms: Research article. *Psychological Science*. SAGE PublicationsSage CA: Los Angeles, CA, 2007. Vol. 18, № 5. P. 429–434.
24. M. Cherkashin. «Vtoraja volna: «Ne nado stremit'sja v bol'nicu!» [«Second wave: «No need to rush to the hospital»] (In Russ.). RBC. 2020. Available: https://www.rbc.ru/spb_sz/12/12/2020/5fd483019a794782e6ceec26. (accessed: 27.05.2021).
25. Abrams E.M. et al. The Challenges and Opportunities for Shared Decision Making Highlighted by COVID-19. *Journal of Allergy and Clinical Immunology: In Practice*. American Academy of Allergy, Asthma and Immunology. 2020. Vol. 8. № 8. pp. 2474–2480.e1.
26. Shleifer A. et al. Older People Are Less Pessimistic about the Health Risks of COVID-19 - HBS Working Knowledge. NBER Working Paper Series. 2020. Vol. 27494. Available: <http://www.nber.org/papers/w27494> (accessed: 31.01.2021).
27. Shlyakhto E.V. et al. Metodicheskie rekomendacii, algoritmy dejstvija medicinskih rabotnikov na razlichnyh jetapah okazaniya pomoshhi, chek-listy i tipovye dokumenty, razrabotannye na period nalichija i ugrozy dal'nejshego rasprostraneniya novoj koronavirusnoj infekcii v Sankt-Peterburge. Versija 3.0 ot 22.06.2020 [Methodological recommendations, algorithms for the actions of medical workers at various stages of assistance, checklists and standard documents developed for the period of the presence and threat of further spread of a new coronavirus infection in St. Petersburg. Version 3.0 from 22.06.2020] (in Russ.). Saint Petersburg, 2020 - 187 pp. Available: http://zdrav.spb.ru/media/filebrowser/covid-19_22.06.pdf (accessed: 02.05.2021).
28. Obshhestvo i pandemija: Opyt i uroki bor'by s COVID-19 v Rossii [Society and Pandemic: Experience and Lessons in Combating COVID-19 in Russia] (in Russ.). Moscow. 2020. P. 774.

А.А. Двойникова, М.В. Маркитантов, Е.В. Рюмина, Д.А. Рюмин,
А.А. Карпов
**АНАЛИТИЧЕСКИЙ ОБЗОР АУДИОВИЗУАЛЬНЫХ СИСТЕМ
ДЛЯ ОПРЕДЕЛЕНИЯ СРЕДСТВ ИНДИВИДУАЛЬНОЙ
ЗАЩИТЫ НА ЛИЦЕ ЧЕЛОВЕКА**

Двойникова А.А., Маркитантов М.В., Рюмина Е.В., Рюмин Д.А., Карпов А.А.
Аналитический обзор аудиовизуальных систем для определения средств индивидуальной защиты на лице человека.

Аннотация. Начиная с 2019 года все страны мира столкнулись со стремительным распространением пандемии, вызванной коронавирусной инфекцией COVID-19, борьба с которой продолжается мировым сообществом и по настоящее время. Несмотря на очевидную эффективность средств индивидуальной защиты органов дыхания от заражения коронавирусной инфекцией, многие люди пренебрегают использованием защитных масок для лица в общественных местах. Поэтому для контроля и своевременного выявления нарушителей общественных правил здравоохранения необходимо применять современные информационные технологии, которые будут детектировать защитные маски на лицах людей по видео- и аудиоинформации. В статье приведен аналитический обзор существующих и разрабатываемых интеллектуальных информационных технологий бимодального анализа голосовых и лицевых характеристик человека в маске. Существует много исследований на тему обнаружения масок по видеозображениям, также в открытом доступе можно найти значительное количество корпусов, содержащих изображения лиц как без масок, так и в масках, полученных различными способами. Исследований и разработок, направленных на детектирование средств индивидуальной защиты органов дыхания по акустическим характеристикам речи человека пока достаточно мало, так как это направление начало развиваться только в период пандемии, вызванной коронавирусной инфекцией COVID-19. Существующие системы позволяют предотвратить распространение коронавирусной инфекции с помощью распознавания наличия/отсутствия масок на лице, также данные системы помогают в дистанционном диагностировании COVID-19 с помощью обнаружения первых симптомов вирусной инфекции по акустическим характеристикам. Однако, на сегодняшний день существует ряд нерешенных проблем в области автоматического диагностирования симптомов COVID-19 и наличия/отсутствия масок на лицах людей. В первую очередь это низкая точность обнаружения масок и коронавирусной инфекции, что не позволяет осуществлять автоматическую диагностику без присутствия экспертов (медицинского персонала). Многие системы не способны работать в режиме реального времени, из-за чего невозможно производить контроль и мониторинг ношения защитных масок в общественных местах. Также большинство существующих систем невозможно встроить в смартфон, чтобы пользователи могли в любом месте произвести диагностирование наличия коронавирусной инфекции. Еще одной основной проблемой является сбор данных пациентов, зараженных COVID-19, так как многие люди не согласны распространять конфиденциальную информацию.

Ключевые слова: определение защитных масок, голосовые характеристики дикторов, COVID-19, средства индивидуальной защиты, обнаружение кашля, лицевые характеристики.

1. Введение. С конца 2019 года все страны столкнулись со стремительным распространением пандемии, вызванной коронавирусной инфекцией COVID-19, борьба с которой продолжается мировым сообществом и по настоящее время. Аномально высокий уровень заражения COVID-19 связан с возможностью его передачи не только воздушно-капельным путем (при кашле, чихании и дыхании инфицированным человеком), но и контактным путем через дверные ручки, поручни в транспорте и другие загрязненные предметы и поверхности. Именно поэтому многие ведущие научные группы и мировые промышленные корпорации из различных научных областей, таких как медицина, биология, информатика и другие, сосредоточились на решении глобальной проблемы уменьшения числа зараженных людей по всему миру [1] и активно ведут исследования и разработки интеллектуальных технологий, которые позволят создать эффективные решения по предотвращению распространения коронавирусной инфекции COVID-19.

Во многих исследованиях, например, [2–4], описывается эффективность использования средств индивидуальной защиты (СИЗ) органов дыхания для предотвращения респираторных заболеваний. Ношение медицинских масок в общественных местах необходимо для обеих категорий людей: для инфицированных, чтобы уменьшить выделения возбудителя инфекции из респираторного тракта и предотвратить дальнейшее распространение болезни, а также для здоровых людей, чтобы избежать попадания вредоносных вирусов и бактерий в организм и минимизировать возможность заражения такими заболеваниями, как туберкулез [5], грипп, острое респираторное заболевание (ОРЗ) или острая респираторная вирусная инфекция (ОРВИ) [6]. Кроме этого, к респираторным заболеваниям относится и коронавирусная инфекция COVID-19 [7], которая легко передается от человека к человеку в воздушной среде, и поэтому для предотвращения пандемии [8] и снижения смертности [9] необходимо обеспечить соблюдение масочного режима в общественных местах. Так, в отечественной работе [10], которая посвящена исследованию по выявлению корреляции между числом заболеваний COVID-19 и плотностью населения, устанавливается линейная зависимость между количеством зараженных людей и численностью определенного субъекта Российской Федерации. Главным способом предотвращения распространения пандемии в странах с высокой плотностью населения является коллективное ношение средств индивидуальной защиты и регулярная их смена [11]. В качестве средств защиты органов дыхания используются маски различных видов: медицинские, тканевые, респираторы и другие. В зависимости

от вида маски уровень их эффективности различается, так, например, группа австралийских ученых в своей работе [12] доказывает, что вероятность попадания различных инфекций в организм человека ниже при ношении специализированных медицинских масок и респираторов относительно тканевых вариаций защитных масок. В другом проведенном исследовании [9] авторы утверждают, что даже малоэффективные маски (уровень задержания вирусов которых не более 20%) могут быть полезны как для профилактики заболеваний, связанных в той или иной степени с COVID-19, у здоровых людей, так и для предотвращения бессимптомной передачи данной коронавирусной инфекции.

Несмотря на эффективность средств индивидуальной защиты органов дыхания от заражения вирусными инфекциями, многие люди пренебрегают использованием защитных масок в общественных местах. На эту тему в 2020 году в Индии проводилось исследование [13], в котором опрашивали людей, пребывавших на домашнем карантине в связи с обнаружением у них респираторного заболевания. Результаты опроса показали, что треть участников (34,1%) не готовы носить маски в общественных местах даже при наличии строгих законов об использовании СИЗ. В свою очередь в работе [14] был произведен анализ и предоставлена статистика ношения масок среди пешеходов на разных улицах городов Ирана, методом наблюдения было установлено, что 45,6% из 10440 людей не используют защитные маски на улицах, а в вечернее время число таких людей увеличивается до 56,1%. В июне 2020 года в штате Висконсин (США) проводилось статистическое наблюдение за покупателями в продуктовых магазинах [15], результаты которого показали, что только 41,5% из 5517 людей правильно носили маски на лицах, остальные 58,5% пренебрегали ношением масок в общественном месте. Во Франции [16] проводили опрос 1012 респондентов, направленный на извлечение мнений людей об обязательных мерах в общественных местах, в том числе ношении масок, 58% опрошенных выступили против наличия в их стране ограничительных мер. Группа исследователей из Института этнологии и антропологии Российской академии наук [17] в первые недели режима самоизоляции провела опрос среди 239 россиян, направленный на получение мнений людей о том, какое процентное соотношение людей носят маски в общественных местах, и результаты опроса показали, что всего 6,3% респондентов считают, что каждый человек надевает защитную маску, 26,8% думают, что защитную маску носит каждый второй, однако большинство респондентов 61,1% говорят, что маски используют только 1-2 человека из толпы, а мнение о том, что маски не носятся

вовсе, разделили между собой 5,9% опрошенных. Во время соблюдения карантинных требований населением для снижения распространения COVID-19 были неизбежны контакты между большим количеством людей, и риск их заражения был достаточно высок, поэтому наличие на лице средств индивидуальной защиты жизненно необходимо. Однако, как показывает статистика [18], только 54,68% водителей такси в Эфиопии использовали маски во время работы. Противоположную статистику можно заметить в медицинских учреждениях, где использование защитных масок является обязательным правилом нахождения на территории помещения. Так, ученые из Малайзии [19] оценили степень распространенности медицинских масок на лицах посетителей районной специализированной больницы во время пандемической вспышки COVID-19. Результаты показали, что среди 1625 человек, включенных в окончательный анализ, почти 97% людей были зарегистрированы в каких-либо масках для лица, из которых 72% носили именно медицинскую маску.

Использование защитных масок помогает уменьшить распространение коронавирусной инфекции COVID-19, однако большинство людей, находясь в общественных местах, пренебрегают средствами индивидуальной защиты. Поэтому для контроля и своевременного выявления нарушителей общественных правил здравоохранения необходимо разрабатывать и применять новые информационные технологии, которые будут детектировать маски на лице человека по видео- и аудиоинформации.

В последние годы в связи с распространением эпидемии, вызванной коронавирусной инфекцией COVID-19, ученые-информатики начали активно заниматься разработкой автоматических систем по детектированию масок на лице человека. Такие системы помогают обеспечить контроль соблюдения людьми законов об обязательном ношении средств индивидуальной защиты органов дыхания в общественных местах. Определить наличие/отсутствие маски на лице человека можно посредством:

1. Видеоинформации (например, видеонаблюдение в общественных местах, камеры смартфонов).
2. Акустической информации (например, телефонные разговоры).

Целью статьи является проведение сравнительного анализа подходов к автоматическому распознаванию наличия/отсутствия масок на лице человека по видео- и аудиоинформации, подходов к распознаванию респираторных заболеваний, в том числе COVID-19 по речи и звукам человека, а также корпусов изображений/видео и аудиоинфор-

мации для решения задач автоматического определения средств индивидуальной защиты. Предложенный сравнительный анализ помогает выделить достоинства и недостатки рассмотренных систем, выявить решенные и нерешенные проблемы в области мониторинга ношения масок в общественных местах, а также выдвинуть универсальные требования к разрабатываемым автоматическим системам определения наличия/отсутствия масок на лице человека.

2. Системы анализа видеoinформации. Одним из самых эффективных способов построения автоматических систем распознавания наличия/отсутствия масок на лицах людей является анализ видеoinформации. Преимуществом таких систем является возможность дистанционного контроля соблюдения масочного режима людей в общественных местах. С помощью видеoinформации можно анализировать ношение масок как конкретных людей (например, сотрудников офиса, клиентов в банке), так и множества людей в целом (например, для сбора статистики ношения масок в общественных местах). Для построения систем детектирования масок на лице человека необходимы корпусы с изображениями людей как в масках, так и без них.

2.1. Корпусы изображений/видео для распознавания наличия/отсутствия масок на лице человека. В последнее время в связи со стремлением многих ученых помочь снизить скорость распространения инфекции COVID-19 было создано несколько корпусов, содержащих в себе изображения людей в масках, для построения систем детектирования наличия/отсутствия СИЗ на лицах людей в общественных местах. Однако еще до возникновения и глобального распространения пандемии коронавирусной инфекции COVID-19 группа китайских ученых [20] собрала и аннотировала корпус MAFA (MAsked FAcEs), который включает изображения лиц с окклюзиями различными предметами, минимальный размер изображений составляет 80 пикселей. Окклюзия лица – перекрытие лица различного характера и степени сложности. Каждое изображение лица аннотировано по трем элементам: тип маски (простая, сложная, перекрытие лица различной областью тела человека, гибридная – комбинация различных типов маски) и степень окклюзий (низкая, средняя, высокая), а также ориентация лица в кадре (слева, слева фронтально, фронтально, справа фронтально, справа).

Medical Masks Dataset (MMD) включает в себя изображения людей в общественных местах. На каждом снимке изображено несколько людей с масками на лице. Face Mask Dataset (FMD) содержит в себе изображения как одного человека, так и нескольких. Разметка изображений производится на 3 класса: «лицо в маске», «лицо без маски»,

«лицо с неправильно надетой маской». Авторы работы [21] собрали корпус изображений людей из регионов Китая, России и Италии под названием MOXAZk. Разметка производилась на 2 класса: «лицо в маске» и «лицо без маски».

Корпус Medical Mask detection содержит изображения лиц людей разной национальности, возрастов и регионов. Изображения имеют разметку 3 классов: «лицо в маске», «лицо без маски» и «лицо с неправильно надетой маской», помимо этого в корпусе также присутствует разметка по 20 объектам, например, шлем, хиджаб, противогаз и другие. Другой корпус – Face mask classification (FMC) включает изображения лиц людей, находящихся на зашумленном фоне. Половина изображений всего объема FMC – это «лица в маске», а другая половина – «лица без маски». Авторы работы [22] собрали корпус Face Mask Detection Video Dataset изображений/видео, содержащих лица людей в масках и без них.

В период пандемии сбор данных, содержащих изображения лиц людей в масках, является затруднительным процессом, так как большинство стран ввели ограничительные меры на свободное передвижение людей, поэтому многие люди соблюдали режим строгой самоизоляции. По этой причине некоторые исследователи прибегали к получению данных изображений лиц в масках синтетическим путем с помощью компьютерного наложения изображения маски на лица людей. Так, авторы работы [23] для построения системы детектирования защитных масок разработали 3 корпуса, которые объединили в один под названием Masked Face Recognition Dataset (RMFD). Он включает в себя корпуса: Masked Face Detection Dataset (MFDD), содержащий в себе 24771 взятых из Интернета изображений лиц в маске, Real-world Masked Face Recognition Dataset (RMFRD) – содержащий 5000 фотографий 525 лиц в масках и 90000 изображений тех же субъектов без маски и Simulated Masked Face Recognition Dataset (SMFRD) – содержащий 500 тыс. изображений лиц, на которые наложили смоделированные медицинские маски.

Simulated Masked Face Dataset (SMFD) содержит в себе фотографии лиц без масок, на которые авторы корпуса наложили изображение обычной медицинской маски. Face Mask ~12K Images Dataset содержит в себе изображения, обрезанные в области лица людей на различных фонах. Корпус Face Mask detection представляет собой изображения лиц в масках и без них, причем маска может быть как реальной, так и наложенной синтетическим путем. Авторы работы [24] использовали корпус CASIA-Webface [25], который содержит фото-

графии лиц знаменитостей, взятых с сайта IMDb¹, для симулирования различных окклюзий на лице. Различные предметы, такие как 6 видов очков и 3 вида маски, в том числе медицинская, с 6 различными текстурами, и их комбинации накладываются на лица людей с учетом ключевых точек изображений лица. Авторы работы [26] создали синтетический корпус путем наложения изображения медицинской маски на лица, используя корпус Labeled Faces in the Wild (LFW) [27], который содержит в себе изображения лиц известных личностей.

На рисунке 1 показаны примеры изображений лиц в реальных и синтетических масках, а также без них из описанных выше корпусов.



Рис. 1. Примеры изображений лиц из различных корпусов

Подробное описание всех упомянутых выше корпусов, а также ссылки для скачивания представлены в таблице 1, из которой видно, что существует достаточно большое количество корпусов изображений лиц как в масках, так и без них. Однако в общей сложности изображений лиц людей, у которых надета маска, значительно меньше, чем лиц без масок, это связано с тем, что сбор данных является сложным процессом из-за введенных во многих странах ограничительных мер. Поэтому некоторые ученые используют синтетический способ получения необходимых данных, а именно компьютерное наложение изображений масок на лица людей. Стоит отметить, что такие данные не всегда выглядят естественно, так как в некоторых случаях маски не перекрывают необходимые области лица, либо, наоборот, перекрывают значительно большую часть лица.

2.2. Подходы к автоматическому распознаванию маски на лице человека. За последние годы было проведено множество исследований, направленных на обнаружение защитных масок на лице человека с использованием видеоинформации. Авторы корпуса MAFA [20] предложили метод LLE-CNN, который состоит из 3 этапов: сначала извлекаются координаты точек лицевых ориентиров, затем применяется локально-линейное встраивание, потом выполняется идентификация потенциальных областей лица и уточнение их позиций путем совместного выполнения задач классификации и регрессии в рамках сверточной нейронной сети (англ. Convolutional Neural Network, CNN).

¹ <https://www.imdb.com/>

В работе [28] предлагается новый подход к детектированию лиц AOFD (Adversarial Occlusion-aware Face Detection), который позволяет одновременно обнаруживать лицо с большим процентом окклюзии и сегментировать его закрытые области. В работе [29] используется более новый корпус Medical Face Mask Detection Dataset (MFMD)², который включает в себя 2 других: MMD и FMD для обнаружения лиц в медицинских масках.

Таблица 1. Корпусы изображений людей в различных защитных масках

Название корпуса (ссылка на источник)	Объем корпуса (изображений)	Количество экземпляров в каждом классе	Способ наложения масок
MAFA ³ [20]	~30 тыс.	~ 35 тыс. лиц в масках	Реальный
MMD ⁴	682	3000 лиц в масках	
FMD ⁵	853	3232 лиц в масках, 717 лиц без масок, 123 с неправильно надетой маской	
MOXA3k ⁶ [21]	3000	3015 лиц в масках, 9161 лиц без масок	
Medical Mask detection ⁷	6000	6769 лиц в масках, 2086 лиц без масок, 227 с неправильно надетой маской	
FMC ⁸	438	219 лиц в масках, 219 лиц без масок	
Face Mask Detection Video Dataset ⁹ [22]	4357	8306 лиц в масках, ~13 тыс. лиц без масок	
SMFD ¹⁰	1570	785 лиц в масках, 785 лиц без масок	Синтетический
CASIA-Webface ¹¹ [24]	~804 тыс.	Нет данных	
LFW ¹² [26]	26 тыс.	13 тыс. лиц в масках 13 тыс. лиц без масок	
RMFD ¹⁴ [23]	~505 тыс.	1025 лиц в масках, 95 тыс. лиц без масок	Реальный и синтетический
Face Mask ~12K Images Dataset ¹⁵	~12 тыс.	5883 лиц в масках, 5909 лиц без масок	Реальный и синтетический
Face Mask detection ¹⁶	4000	1915 лиц в масках, 1918 лиц без масок	

² <https://www.kaggle.com/mloey1/medical-face-mask-detection-dataset>

³ <https://github.com/ThanhNguyenFG/Masked-Face-Detection>

⁴ <https://www.kaggle.com/vtech6/medical-masks-dataset>

⁵ <https://www.kaggle.com/andrewwmvd/face-mask-detection>

⁶ <https://shitty-bots-inc.github.io/MOXA/index.html>

⁷ <https://www.kaggle.com/humansintheloop/medical-mask-detection>

⁸ <https://www.kaggle.com/dhruvmak/face-mask-detection>

⁹ <https://data.mendeley.com/datasets/v3kry8gb59/1>

¹⁰ <https://github.com/prajnasb/observations>

¹¹ <https://github.com/Baojin-Huang/Webface-OCC>

¹² <http://vis-www.cs.umass.edu/lfw/>

¹³ <https://www.kaggle.com/muhammedalkran/lfw-simulated-masked-face-dataset>

¹⁴ <https://github.com/X-zhangyang/Real-World-Masked-Face-Dataset>

¹⁵ <https://www.kaggle.com/ashishjangra27/face-mask-12k-images-dataset>

¹⁶ <https://www.kaggle.com/aneerbanchakraborty/face-mask-detection-data>

Для эффективной работы систем детектирования различных защитных масок на лицах людей необходим большой объем обучающих данных. На сегодняшний день существуют несколько корпусов с изображениями людей в масках. Многие исследователи, работающие в области искусственного интеллекта, для обнаружения СИЗ на лицах используют несколько корпусов для обучения. В статье [26] предлагается применять гибридную модель, использующую глубокое обучение – нейронную сеть ResNet-50 и классическое обучение – метод опорных векторов (англ. Support Vector Machine, SVM), для обнаружения СИЗ на лицах людей. Для обучения, проверки и тестирования такой системы авторы объединяют 2 корпуса: RMFD [23] и SMFD. И дополнительно для тестирования системы авторы используют корпус LFW с изображениями лиц знаменитых с синтетически наложенными масками. Недостатком представленных систем является то, что авторы работ используют корпуса, содержащие лица в медицинских масках, и остается непонятным, какой результат покажут системы, если лицо будет закрыто другим предметом, например смартфоном.

Авторами данной статьи было проведено исследование по обнаружению СИЗ на лицах людей с апробацией системы при окклюзии на лице иным предметом, отличным от защитной маски [30]. Исследование проводилось на трех корпусах: Medical Mask detection была выбрана для обучения и проверки, а RMFRD и MAFA (после дополнительной разметки) – для тестирования. В работе был предложен гибридный метод, в основе которого лежит объединение двух наборов признаков: глубокие признаки, извлеченные с помощью нейронной сети ResNet-50 [31], и характеристики распределения интенсивности пикселей на изображениях. В качестве классификатора выступает простая полносвязная нейросеть (англ. Fully Connected Neural Network, FCNN). Схема предложенного метода представлена на рисунке 2. Для увеличения точности работы метода используется аугментация данных, а именно аффинные преобразования и расширение обучающих данных посредством Интернета. Вероятность ложного срабатывания предложенного метода при окклюзии на лице другим предметом на корпусе MAFA составляет 45,48%. Эксперименты показали, что проблема ложного обнаружения защитной маски при окклюзии на лице другим предметом остается, и для ее решения необходимо использовать сложные экземпляры на этапе обучения, как класс «лицо без маски».



Рис. 2. Функциональная схема метода распознавания наличия/отсутствия защитной маски на лице человека [30]

Системы детектирования защитных масок на лице человека необходимы для того, чтобы в режиме реального времени осуществлять контроль над соблюдением рекомендаций от Всемирной организации здравоохранения (ВОЗ), касаемых ношения средств индивидуальной защиты органов дыхания в общественных местах. Проблемы построения таких систем заключаются в том, что они должны быть небольшого размера из-за ограниченной памяти периферийных устройств, иметь высокую скорость обработки данных и при этом показывать высокую точность обнаружения масок на лицах людей. Для решения данных проблем авторы работы [21] предложили использовать модель обнаружения объектов YOLO v3 Tiny [32] и легкую нейронную сеть MobileNet v2 [33]. С применением предложенных моделей авторам удалось получить максимальную скорость обработки видеоизображений – 138 кадров в секунду. В статье [34] для обнаружения масок в реальном времени предлагается подход под названием SSDMNv2, который включает в себя модель обнаружения объектов Single Shot Multibox Detector (SSD) [35] с базовой архитектурой ResNet-10 [36] и нейронную сеть с архитектурой MobileNet v2 [33] в качестве классификатора. С помощью подхода SSDMNv2 удалось достичь скорости обработки – 15,71 кадров в секунду. Авторы [37] для построения системы обнаружения СИЗ на лице человека использовали 4 различных корпуса: Face Mask ~12K Images Dataset, FMC и 2 корпуса для обучения и тестирования (OpenMV Dataset), собранных авторами статьи [37], содержащих изображения с камеры OpenMV Cam H7, объемами 1979 и 594, соответственно. Для увеличения объема обучающих данных к изображениям применялось масштабирование с различной интерполяцией. В качестве классификатора было предложено использовать CNN, с помощью которой достигается скорость обработки данных – 30 кадров в секунду.

Краткая характеристика рассмотренных работ для задачи распознавания наличия/отсутствия маски на лице отображена в таблице 2. В ней описываются используемые корпуса и методы, а также представлены показатели эффективности предложенных методов.

Таблица 2. Сравнительная характеристика работ по распознаванию наличия/отсутствия масок на лицах с использованием видеoinформации

Работа	Корпусы	Методы	Показатель	Результат
Ge S. et al. [20]	MAFA	LLE-CNN	Average precision	76,14%
Chen Y. et al. [28]	MAFA	AOFD		76,14%
Loey M. et al. [29]	MMD, FMD, MAFA	YOLO v2, ResNet-50		81,00%
Roy B. et al. [21]	MOXA3k	YOLO v3 Tiny, MobileNet v2		56,27%
Nagrath P. et al. [34]	MMD, RMFD	SSDMNV2		94,00%
Loey M. et al. [26]	MMD, FMD, MAFA	ResNet-50, SVM	Accuracy	99,64% (RMFD) 99,49% (SMFD) 100% (LFW)
Mohan P. et al. [37]	Face Mask ~12K Images Dataset, FMC, OpenMV Dataset	CNN		99,83%
Ryumina E. et al. [30]	Medical Mask detection, RMFRD, MAFA	Гибридный метод извлечения признаков, FCNN	UAR	98,12% (RMFRD) 97,68% (MAFA)

Обзор существующих современных исследований показал, что в большинстве случаев используются нейросетевые либо гибридные подходы, с помощью которых можно достичь высоких результатов обнаружения защитных масок у человека. Направление детектирования СИЗ на лицах в области машинного обучения стало востребованным относительно недавно, поэтому на сегодняшний день недостаточно изучены современные подходы и их комбинации, которые показывают высокую эффективность на смежных задачах. Также стоит отметить тот факт, что сравнение эффективности методов не всегда возможно, так как при обучении систем обнаружения масок используются различные корпуса, а также их комбинации. Несмотря на полученные высокие результаты работы систем, предложенных различными исследователями, работающими в области детектирования СИЗ на лицах

людей, эффективность систем при тестировании на лицах с высокой степенью окклюзии различными предметами остается неизвестной.

Несмотря на большое количество корпусов с изображениями лиц для детектирования защитных масок, затруднительно обучить надежную систему определения СИЗ у человека, так как в общей сложности изображений с реальными, а не синтетическими масками на лицах, в несколько раз меньше, чем лиц без масок. Генерация синтетических масок является эффективным способом в случае, когда сбор реальных данных является затруднительным процессом, однако наложение синтетических масок на изображения лиц зачастую выглядит неестественно и самое главное не всегда правильно перекрывает необходимые части лица.

3. Системы анализа аудиоинформации. Акустические характеристики речевых высказываний могут отражать полезную информацию, например о физиологическом [38] и психологическом [39] здоровье человека, о проявлениях различных эмоциональных экспрессий [40]. Для предотвращения борьбы с пандемией COVID-19 также можно использовать анализ акустической информации для решения различных задач [41], например, распознавания респираторных заболеваний, в том числе COVID-19, распознавания наличия/отсутствия маски и другие. Распознавание наличия/отсутствия маски на лице человека по акустическим данным является более сложной задачей, но не менее актуальной на сегодняшний день по сравнению с детектированием СИЗ по видеоданным. Кроме того, выявление различных респираторных заболеваний, кашля и коронавирусной инфекции по голосовым характеристикам человека позволяют значительно снизить скорость распространения COVID-19. Для построения эффективных систем анализа аудиоинформации в различных задачах необходимы наборы обучающих данных.

3.1. Корпусы аудиоданных для задач распознавания заболеваний и наличия/отсутствия масок. Респираторные заболевания передаются от человека к человеку воздушно-капельным путем. Одним из основных признаков таких заболеваний является кашель. С помощью анализа аудиоинформации речевых высказываний можно детектировать у человека наличие/отсутствие кашля. Такие системы могут применяться в области здравоохранения, помогая врачам дистанционно наблюдать за состоянием пациента, а также для незамедлительного оказания медицинской помощи при проявлении первых признаков заболеваний. Для разработки таких автоматических систем нужны новые корпусы, позволяющие решать задачу детектирования кашля и респираторных заболеваний. Так, авторы работы [42] создали 2 речевых

корпуса для обнаружения кашля. Первый корпус – Synthetic Cough Sounds содержит полусинтетические данные. Авторы записывали речь с кашлем и без, а также различные фоновые шумы (звуки ветра, дыхания, одышки, различных помещений, офисов, ресторанов, вокзала, многолюдной улицы), которые впоследствии накладывались на записанные речевые высказывания. Помимо кашля корпус содержит также звуки прочистки горла, сопения, чихания, отрывки, дыхания, отдышки, смеха, храпа, глотания и речи. Данный корпус содержит 26 высказываний с длительностью от 15 до 155 секунд. Количество информантов, их возраст и соотношение полов не уточняются. Информант – человек, служащий источником информации для различных исследовательских целей. Второй корпус – Real Cough Sounds содержит записи пациентов Королевской клиники Эдинбурга (Великобритания), которые болели бронхитом, астмой или хронической обструктивной болезнью легких во время записи корпуса. Корпус записывался в 3 сценариях: запись речи с низким уровнем шума; с фоновым шумом; с шумом, издаваемым информантом во время выполнения различных заданий, например, печатание на клавиатуре, открытие/закрытие окон и так далее. Возраст информантов варьируется от 45 до 72 лет, среди них 3 мужчины и 10 женщин. Всего корпус содержит 78 высказываний.

Отсутствие проявления кашля не может гарантировать здоровое физиологическое состояние человека. При наличии простудных заболеваний у человека может и не проявляться явных респираторных признаков, таких как кашель и чихание, однако акустические характеристики речи человека могут содержать в себе информацию о нездоровом состоянии органов дыхания. Так, в 2017 году проходили международные соревнования по компьютерной паралингвистике INTERSPEECH Computational Paralinguistics Challenge 2017 (ComParE 2017) [43], на которых участникам предлагалось определить наличие/отсутствие простуды по акустическим характеристикам речи. Для этого организаторы предоставили корпус аудиоданных инфекции верхних дыхательных путей (англ. Upper Respiratory Tract Infection Corpus, URTIC), который был разработан Вуппертальским университетом (Германия). Данный корпус содержит высказывания 382 мужчин и 248 женщин. Возраст информантов варьируется от 16 до 84 лет (средний возраст – 29,5 лет). Запись проходила в тихих комнатах. Участники выполняли различные задания: рассказывали истории (последние выходные, отпуск), описывали картинки, произносили различные команды, проговаривали числа, читали короткие рассказы. Сессии длились от 15 минут до 2-х часов, при этом количество заданий варьировалось для каждого информанта. Всего корпус содержит 28652

высказывания с длительностью от 3 до 10 секунд. В обучающую и валидационную выборки включались записи только 210 информантов, из которых 37 были простужены.

В период пандемии анализ акустической информации для обнаружения коронавирусной инфекции является особенно востребованной задачей. В начале 2020 года индийские ученые [44] в целях обнаружения COVID-19 по кашлю, дыханию и голосу разработали новый речевой корпус Coswara, который был собран с помощью платформы краудсорсинга с использованием веб-приложения. С помощью данного приложения информанты добровольно записывали речь, кашель и дыхание на свои электронные устройства, а также проходили анкетирование, в котором указывали параметры записанной речи (тип аудио: кашель, речь, дыхание; отсутствие или присутствие шумов на заднем плане), а также наличие/отсутствие коронавируса у информанта. Данный корпус¹⁷ содержит высказывания 1123 мужчин и 363 женщин, возрастом от 7 до 80 лет.

Ученые из Кембриджского университета (Великобритания) [45] также разработали подобный речевой корпус с использованием приложения для мобильных устройств, записывая информантов с помощью их собственных смартфонов и планшетов. Информанты также предоставляли результаты тестирования на COVID-19, основную демографическую и медицинскую информацию. Кембриджский корпус Cambridge COVID-19 Sound содержит около 10000 аудиозаписей 6613 информантов (4525 мужчин и 2069 женщин). У 406 из них были документально подтвержденные положительные тесты на COVID-19. Кроме того, 691 информант записывались более одного раза. Возрастной диапазон информантов варьировался от 16 до 80 лет. В начале 2021 года количество аудиозаписей и информантов возросло¹⁸. В записи данного корпуса участвовали жители из Греции, Великобритании, Италии, Германии, Испании, Ирана, США, Бангладеша, Индии и Франции.

Корпус Sonde Health COVID-19 2020 (SHC) [46] состоит из записей речи, собранных из бесплатного приложения Sonde Health¹⁹ на смартфоне. Запись данных происходила в неклинических условиях, например дома, в автомобилях, на рабочих и тому подобных местах. Каждый участник записывал по 3 речевых высказывания, заранее подготовленных организаторами, а также заполнял анкету, в которой ука-

¹⁷ https://iiscleap.github.io/coswara-blog/coswara/2020/11/23/visualize_coswara_data_metadata.html

¹⁸ https://www.covid-19-sounds.org/en/blog/voice_covid_icassp.html

¹⁹ <https://www.sondehealth.com/sondeone-page>

зывались пол, возраст, субъективная оценка качества записей, результат теста на COVID-19.

Для предотвращения передачи инфекции между людьми воздушно-капельным путем необходимо использовать средства индивидуальной защиты, например, медицинские маски, как здоровым людям, так и инфицированным. Для задачи обнаружения наличия/отсутствия маски на лице информантов по голосовым характеристикам существуют несколько речевых корпусов. Одним из них является речевой корпус Хельсинкского университета (Финляндия) (англ. *Speech corpus of University of Helsinki*) [47]. Он был разработан для распознавания информантов в масках с целью сокращения преступлений. Корпус содержит речевые высказывания 4 мужчин и 4 женщин, возрастом от 21 до 28 лет. Все они являются носителями финского языка и студентами университета. Корпус записывался в звукоизолированной комнате площадью около 5 квадратных метров. Данные записывались одновременно с 3 микрофонов. Информанты в 2 сеансах зачитывали ряд предложений, а также имитировали спонтанную речь, описывая картинки из комиксов. Каждым информантом было записано 60 аудиофайлов, длительность которых составляет от 60 до 90 секунд.

В рамках международной конференции INTERSPEECH 2020 на соревновании ComParE 2020 [48] был впервые представлен Аугсбургский речевой корпус (англ. *Mask Augsburg Speech Corpus, MASC*). Он состоит из речевых высказываний людей, находящихся в медицинской одноразовой маске и без нее. Корпус содержит аудиозаписи носителей немецкого языка (16 женщин и 16 мужчин, возрастом от 20 до 41 года, средний возраст 25,6). Записи были сделаны в звукоизоляционной аудиостудии с использованием конденсаторного микрофона AKG C4500 BC. Исходные аудиозаписи (48 кГц, 24 бит на отсчет) в дальнейшем преобразовывались в аудиофайлы с частотой дискретизации 16 кГц, 16 бит на отсчет. Во время записи участники выполняли разные речевые задания в маске и без: они отвечали на вопросы, читали текст с различными медицинскими терминами, а также описывали словами рисунки со спортивными мероприятиями, семьями, детьми и едой. Аудиофайлы были разделены на фрагменты продолжительностью в 1 секунду без наложения.

Сбор естественной речи является затруднительным процессом, поэтому некоторые ученые прибегают к сбору синтезированной речи. Речевой корпус *Mask Sorbonne Speech Corpus (MSSC)* [49] был собран одной из команд участников соревнований ComParE 2020 для увеличения размера обучающего корпуса MASC. Для имитации человеческого голоса авторы использовали акустическую систему Bose Sound-

Link micro. Процесс записи корпуса проходил в два этапа. Сначала с помощью динамика проигрывались 1000 высказываний от 30 информантов (15 мужчин и 15 женщин), которые были выбраны из речевого корпуса German Distant Speech Data Corpus [50], затем на колонку надевалась медицинская маска и снова проигрывались эти же высказывания. Таким образом авторы создали синтетический корпус, который содержит высказывания людей в масках и без них. Процесс записи происходил в безэховой камере.

В таблице 3 представлены основные параметры рассмотренных речевых корпусов: количество информантов, аудиозаписей, длительность, и количество экземпляров в каждом классе.

Таблица 3. Речевые корпуса для детектирования кашля, наличия/отсутствия маски и COVID-19 (н/д – не доступно)

Название	Кол-во информ.	Длительность, часов	Количество экземпляров в каждом классе
<i>Детектирование кашля и простуды</i>			
Synthetic Cough Sounds [42]	н/д	~3	н/д
Real Cough Sounds [42]	13	~26	н/д
Upper Respiratory Tract Infection Corpus, URTIC [43]	630	~45	1987 (речь простуженных людей), 17070 (речь здоровых людей), + 9551 (скрытая выборка)
<i>Детектирование COVID-19 по речи и звукам</i>			
Coswara ²⁰ [44]	1486	н/д	1498 (речь здоровых людей), 157 (речь людей с COVID-19), 117 (речь людей с респ. инфекцией)
Cambridge COVID-19 Sound [45]	32 тыс.	н/д	н/д
Sonde Health COVID-19 2020 (SHC) [46]	66	0,33	66 (речь инфицированных людей), 132 (речь здоровых людей)
<i>Детектирование наличия/отсутствия маски на лице информанта</i>			
Speech corpus of University of Helsinki [47]	8	~12	96 (речь в шлеме), 96 (речь в резиновой маске), 96 (речь в медицинской маске), 96 (речь в шарфе), 96 (речь без маски)
Mask Augsburg Speech Corpus, MASC [48]	32	~13	~18 тыс. (речь с маской) ~19 982 (речь без маски)
Mask Sorbonne Speech Corpus, MSSC [49]	30	н/д	1000 (речь без маски) 1000 (речь с маской)

²⁰ <https://github.com/iiscleap/Coswara-Data>

Наличие корпусов, содержащих акустические характеристики речевых высказываний людей, говорит о том, что распознавание кашля, наличия/отсутствия маски на лице человека, а также обнаружение коронавирусной инфекции COVID-19 являются актуальными задачами в области анализа речи и аудиосигналов, особенно в условиях пандемии.

3.2. Подходы к автоматическому распознаванию респираторных заболеваний. Еще задолго до наступления пандемии COVID-19 ученые начали проводить исследования по выявлению различных респираторных заболеваний по голосовым характеристикам человека. Так, в работе [51] представлена система, построенная с применением скрытых марковских моделей (англ. Hidden Markov model, HMM), обнаружения кашля у амбулаторных больных, которые носили на груди звукозаписывающее устройство и микрофон. Но ношение специального периферийного устройства для записи кашля для больных пациентов становится затруднительно, поэтому предложено детектировать наличие кашля из аудиозаписей со смартфона. Авторы статьи [42] предложили использовать локальные моментные характеристики, полученные из мел-частотных кепстральных коэффициентов (англ. Mel-Frequency Cepstral Coefficients, MFCCs), и классификатор на основе метода k -ближайших соседей (k NN) для обнаружения кашля по записям со смартфона. Зачастую запись на смартфоны происходит в достаточно зашумленных условиях, поэтому обнаружение кашля по аудиозаписям становится затруднительным. В работе [52] также использовался k NN для обнаружения кашля в зашумленных звуковых сигналах. Для проверки эффективности предложенного метода авторы собрали корпус аудиозаписей, содержащих кашель, и наложили на них различные шумовые звуки.

Победители соревнований INTERSPEECH ComParE-2017 в своей статье [53] достигли наилучших результатов, используя комбинацию нескольких классификаторов: глубокие нейронные сети (англ. Deep Neural Networks, DNN) для извлечения акустических признаков и SVM в качестве классификатора. Системы, направленные на распознавание заболеваний дыхательных путей по аудиосигналам человека, находят активное применение в телемедицине. Такая система, работающая на смартфоне или на любом другом портативном устройстве, может помочь врачам осуществлять дистанционный мониторинг пациентов, находящихся как на территории медицинских учреждений, так и в домашних условиях.

3.3. Подходы к автоматическому распознаванию COVID-19 по речи и звукам человека. В период пандемии анализ акустической

информации для обнаружения коронавирусной инфекции является особенно востребованной задачей. Поэтому в 2021 году в рамках международной конференции INTERSPEECH 2021 было анонсировано несколько соревнований по определению COVID-19 по кашлю, дыханию и речи. На ежегодном соревновании по компьютерной паралингвистике ComParE 2021 [54] организаторы предложили исследователям две задачи: обнаружение коронавируса по кашлю и по речи людей (инфицированных COVID-19 и заболевших иными респираторными инфекциями). Участникам предоставлялись около 2000 записей речевого корпуса Cambridge COVID-19 Sound [45]. Базовый подход, с помощью которого удалось достичь точности по показателю $UAR=73,9\%$ и $UAR=72,1\%$ для задач распознавания COVID-19 по кашлю и по речи, соответственно, представлен в статье [54]. В данных соревнованиях приняли участие несколько команд. В работе [55] в качестве признаков используется необработанный звук в сочетании с различными вариациями спектрограмм, в качестве классификаторов используется ансамбль CNN моделей с различной архитектурой. Авторы работы [56] применяют фонетический анализ и модель Гауссовой смеси (англ. Gaussian Mixture Model, GMM) для распознавания COVID-19 по речи. В работе [57] проводятся исследования с тем же корпусом, но уже не в рамках соревнований. Авторы работы отбирают данные с учетом сбалансированности классов и выделяют свои обучающую, валидационную и тестовую выборки. В качестве признаков авторы используют спектрограммы, прошедшие через 128 мел-фильтров, такое представление спектрограммы известно как мел-спектрограмма.

Также в рамках INTERSPEECH-2021 проводится соревнование DiCOVA [58], включающее в себя несколько задач: детектирование COVID-19 по кашлю и по дыханию. Для секции обнаружения COVID-19 по кашлю предоставлен корпус Coswara [44]. Авторы соревнований предложили использовать MFCCs аудиопризнаки совместно с 3 различными классификаторами на основе: логистической регрессии (англ. Logistic Regression, LR), многослойного перцептрона (англ. Multilayered Perceptron, MP) и случайного леса (англ. Random Forest, RF), с помощью которых достигается базовый результат по показателю площади под кривой (англ. Area Under Curve, AUC) в 69,71%.

Краткая характеристика рассмотренных работ для задачи детектирования наличия/отсутствия кашля и простуды и COVID-19 представлена в таблице 4. В ней описываются используемые признаки и классификаторы для построения систем, а также приведены показатели и их числовые значения.

Таблица 4. Сравнительная характеристика работ по анализу аудиоинформации в задачах распознавания респираторных заболеваний

Работа	Аугментация и признаки	Машинный классификатор	Количественный показатель и его значение
<i>Детектирование кашля и респираторных заболеваний</i>			
Matos S. et al. [51]	н/д	HMM	Accuracy = 82,00%
Monge-Álvarez J. et al. [42]	Локальные моментные характеристики	kNN	Accuracy = 95,28%
Monge-Álvarez J. et al. [52]	MFCs	kNN	Accuracy = 85,71%
Gosztolya G. [53]	Признаки из DNN	SVM	UAR=64,00%
<i>Детектирование COVID-19 по речи и звуку</i>			
Schuller B. et al. [54]	openSMILE, BoAW, auDeep, DeepSpectrum	SVM	UAR= 73,90% (кашель), UAR= 72,10% (речь)
Schuller B. et al. [55]	Необработанный звук с различными спектрограммами	CNN	UAR = 76,10% (кашель), UAR = 73,10% (речь)
Klump P. et al. [56]	Спектрограммы	GMM, SVM, LR	UAR = 64,2%
Xia T. et al. [57]	Мел-спектрограммы	CNN	ROC-AUC = 74,00%
Muguli A. et al. [58]	MFCs	LR, RF, MP	AUC = 69,71%

3.4. Подходы к автоматическому распознаванию наличия/отсутствия маски на лице информанта. Помимо детектирования респираторных заболеваний, также немаловажно контролировать наличие/отсутствие защитной маски на лице инфицированного человека. Анализ видеоинформации не всегда может показывать необходимую эффективность в детектировании маски на лице человека, а в некоторых случаях может быть вообще невозможен, например, при плохом освещении в помещении, удаленности камеры от человека, либо в клиентских сервисах, где камера рассчитана на одного человека, а общение происходит с несколькими, также в государственных учреждениях, где видеосъемка запрещена. Поэтому в случаях отсутствия возможности получить видеоинформацию о наличии/отсутствии маски на лице можно анализировать акустические характеристики голоса человека, которые содержат информацию о наличии/отсутствии защитной маски. Было произведено множество исследований, анализирующих влияние защитной маски на качество и разборчивость речи. В 2008 г. была опубликована статья [59] об анализе изменений акустических характеристик речи медицинского персонала при использовании защитной маски. Спектральные характеристики аудиозаписей показывают различия речевого сигнала при использовании маски и без

нее, тем не менее, наличие маски на лице врачей почти никак не отражается на восприятии и понимании его речи другим персоналом при отсутствии проблем со слухом. С наступлением пандемии COVID-19 начали активно проводиться исследования по разборчивости речи человека в маске. Авторы работы [60] выбрали 156 фраз на английском языке с низкой предсказуемостью из корпуса Speech Perception in Noise (SPIN) [61], которые произносили 2 информанта по несколько раз в масках и без маски, и с различным стилем произношения речи: четким, небрежным и позитивно эмоциональным. Затем эти аудиозаписи давали прослушать 63 экспертам-аудиторам, которые должны были идентифицировать слова в речи информанта, далее эти текстовые данные записывались в один файл²¹. Исследования показывают, что говорящие адаптируют свою речь в зависимости от наличия маски на лице. Разборчивость четко произносимой речи даже лучше, когда человек находится в маске, нежели без нее, при эмоциональной речи получились противоположные результаты, с наличием маски у информантов речь становится хуже воспринимаемой, а при небрежной речи, как при наличии маски, так и без нее, значение разборчивости слов и звуков речи не меняется. В работе [62] представлены исследования влияния различных типов масок на разборчивость речи: одноразовых медицинских и многоразовых тканевых масок, а также респираторов. Для экспериментов использовался имитатор человеческой головы и рта, на который надевали различные маски и записывали синтезированную речь как в масках, так и без них. Результаты анализа показали, что все виды масок ухудшают разборчивость речи, более затруднительно речь воспринимается, когда информант носит двухсторонние тканевые маски. Аналогичные результаты показало исследование, описанное в работе [63], при использовании таких типов «масок» как мотоциклетного шлема, резиновой маски, шарфа и одноразовой медицинской маски.

Организаторы соревнований ComParE 2020 [48] для задачи распознавания маски на лице информанта предоставили базовые признаки речевых высказываний, такие как openSMILE ComParE [64], BoAW [65], auDeer [66], DeepSpectrum [67]. В данных международных соревнованиях приняли участие несколько команд. Большинство команд в качестве признаков использовали спектрограммы. Один из предложенных методов [68] для решения данной задачи заключался в использовании нескольких SVM, которые обучались на MFCCs, векторах Фишера (англ. Fisher Vectors, FV) и всех базовых признаках. Затем на полученных вероятностях обучался еще один SVM классификатор. В

²¹ <https://data.mendeley.com/datasets/74p6w8xx5r/1>

работах [69, 49] помимо акустических характеристик речи анализировали также фонетические характеристики. В работе [69] использовали нейронные сети в качестве классификатора, такие как CNN и рекуррентные нейронные сети (англ. Recurrent Neural Network, RNN). Также был представлен ряд исследований [70-72], в которых использовались различные методы аугментации данных. Так, в [71] был рассмотрен нейросетевой подход на основе генеративно-сопоставительных нейронных сетей (англ. Generative Adversarial Networks, GAN) [73], который не смог превзойти по точности простые методы аугментации (SpecAugment и Mixup). В статье [72] в качестве классификаторов использовались предварительно обученные нейронные сети (англ. Pre-trained Audio Neural Networks, PANNs). Победителем оказалась команда Аризонского университета [74]. Американский коллектив в качестве признаков использовал изображения спектрограмм линейного масштаба, а в качестве классификаторов были предложены нейросетевые подходы с использованием предварительно обученных моделей, таких как VGG [75], ResNet [31], DenseNet [76] и AlexNet [77]. Авторы провели анализ специфичных для классов областей изображений спектрограмм с помощью метода создания тепловых карт (англ. Class Activation Mapping, CAM). Основываясь на этом анализе, они предположили, что спектрограммы линейного масштаба работали лучше, потому что они более растянуты в диапазоне частот 3-5 кГц, которые имеют решающее значение для данной задачи классификации, в то время как на спектрограмме в мел-частотном масштабе данный диапазон составляет около 15%.

Авторы настоящей статьи совместно с нидерландскими и немецкими коллегами также принимали участие в данном соревновании [78]. Использование метода кросс-валидации и тщательный выбор предварительно обученных нейросетевых моделей позволили получить самые современные ансамбли DNN, точность распознавания и обобщающая способность которых значительно увеличились в сравнении с базовым существующим подходом. Мы использовали 10 кросс-валидационных выборок, на которых обучалась предобученная модель ResNet18v2 при различных алгоритмах оптимизации: 10 из них с методом адаптивной оценки моментов (англ. Adaptive Moment Estimation, ADAM) и 10 – со стохастическим градиентным спуском (англ. Stochastic gradient descent, SGD). Общая схема данного метода представлена на рисунке 3. В результате мы обучили 20 нейронных сетей ResNet18v2. После этого в каждой кросс-валидационной выборке данных мы взвешивали прогнозы двух сетей (ResNet18v2 с ADAM и ResNet18v2 с SGD). Затем рассчитывалось среднее предсказание меж-

ду полученными значениями на каждой кросс-валидационной выборке. Все нейронные сети минимизировали целевую функцию бинарной кросс-энтропии.

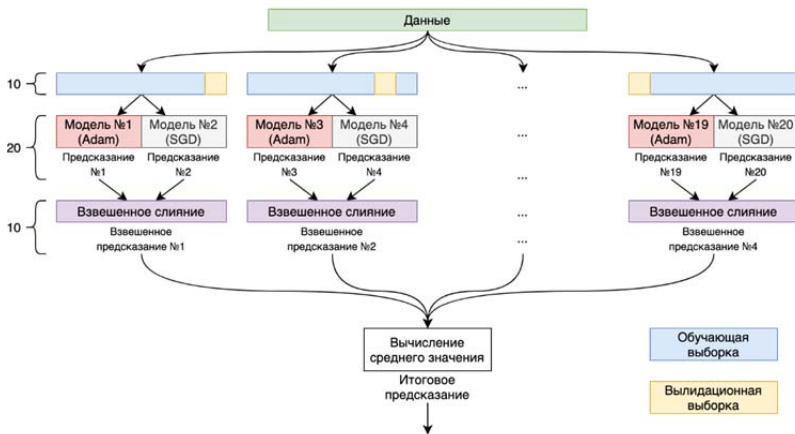


Рис. 3. Схема метода для обнаружения медицинской маски по голосовым характеристикам информанта [78]

Краткая характеристика рассмотренных работ для задачи детектирования наличия/отсутствия маски представлена в таблице 5.

Направление детектирования защитных масок на лицах по акустическим характеристикам речевых высказываний человека является достаточно новой задачей, оно стало востребованным только в период пандемии коронавирусной инфекции COVID-19. Со временем актуальность данного направления только возрастает, так как эффективно работающая система обнаружения наличия/отсутствия инфекции у человека по его кашлю, речи, а также детектирования защитных масок с помощью акустической информации позволит обеспечить контроль состояния человека дистанционно, а также предотвратить процесс распространения вирусной инфекции за счет мониторинга наличия/отсутствия СИЗ на лицах людей. К тому же, данные системы легко можно использовать на устройствах, которыми люди пользуются ежедневно, это поможет осуществлять непрерывное наблюдение за человеком, а также не будет составлять дискомфорт пользователю при ношении данного устройства.

Таблица 5. Сравнительная характеристика работ по анализу аудиоинформации в задаче детектирования маски

Работа	Аугментация и признаки	Машинный классификатор	Количественный показатель и его значение
Plium et al. [70]	Аугментация (скорость, шум, громкость, смещение, SpecAugment). Спектрограммы	CNN	UAR = 71,50%
Базовое решение ComParE 2020 [48]	openSMILE ComParE, BoAW, auDeep, Deep-Spectrum	SVM	UAR = 71,80%
Ristea et al. [71]	Аугментация с помощью GAN. Спектрограммы	ResNet + SVM	UAR = 74,60%
Yang et al. [68]	MFCCs, FV, openSMILE ComParE, BoAW, auDeep, Deep-Spectrum	SVM	UAR = 75,10%
Klumpp et al. [69]	Спектрограммы	CNN, RNN	UAR = 75,40%
Markitantov M. et al. [78]	Спектрограммы	ResNet с разными оптимизаторами	UAR = 75,90%
Koike et al. [72]	SpecAugmen, Mixup. Спектрограммы	PANNs	UAR = 77,50%
Montacié et al. [49]	Признаки из ResNet50 и низкоуровневых дескрипторы (openSMILE)	kNN, SVM	UAR = 77,70%
Szep et al. [74]	Линейные спектрограммы	ResNet, DenseNet	UAR = 80,10%

Помимо детектирования респираторных заболеваний и наличия/отсутствия маски у информанта, аудиоинформация помогает определить у человека его возраст, наличие головной боли, отсутствие аппетита и другое. Все эти метаданные являются вспомогательными признаками коронавируса, например люди пожилого возраста более подвержены заражению, поэтому с помощью таких данных можно осуществлять начальный скрининг заражения вирусной инфекцией. Подробный аналитический обзор по методам, используемым в данных задачах, представлен в статье [79].

4. Аудиовизуальные корпуса. На сегодняшний день известны только 2 многомодальных корпуса для задачи распознавания наличия/отсутствия маски на лице человека. Корпус Audio-Visual Face Cover Corpus (AVFCC) [80] содержит аудиовизуальные записи информан-

тов с различными «масками», такими как мотоциклетный шлем, балаклава, обычная медицинская маска. Одновременно записывались две непрерывные видеозаписи в формате HD (1280×720). Звук данного корпуса был записан на портативный 4-х канальный рекордер Edirol R-4 Pro. Чтобы облегчить синхронизацию двух видеозаписей, в начале каждой записи производились хлопки. Данный корпус состоит из 6120 высказываний, которые были записаны 10 носителями английского языка. Корпус содержит аудиозаписи 5 мужчин и 5 женщин с возрастным диапазоном от 21 до 36 лет (средний возраст – 26,5 лет).

Корпус аудиовизуальных русскоязычных данных людей в защитных масках (англ. Biometric Russian Audio-Visual Extended MASKS corpus, BRAVE-MASKS) [81], созданный авторами настоящей статьи, содержит разноразмерные изображения лиц людей в различных вариациях защитных масок, а также аудиозаписи слитной русской речи людей в масках. Корпус предназначен для анализа лицевых и голосовых характеристик информантов в маске и без, для обучения и тестирования системы автоматического распознавания наличия/отсутствия маски, распознавания типа маски, а также бимодальной верификации и идентификации информантов. Запись корпуса проводилась посредством двух смартфонов Apple iPhone XS Max и планшета Apple iPad Pro в офисных условиях на разнородном фоне. Все видеофайлы были обработаны экспертом и разбиты на группы в зависимости от используемой информантом маски. Записанный корпус имеет следующие параметры: 30 информантов (15 женщин и 15 мужчин) разных возрастов (от 19 до 86 лет, средний возраст – 40 лет). Общее количество вариаций защитных масок и респираторов составляет 33 штуки (одноразовые медицинские маски разных цветов, многоразовые тканевые маски разных цветов с рисунками и без, медицинские и строительные респираторы и другие средства защиты слизистых поверхностей, в которых лицо остается частично видимым). Каждому информанту предоставлялось 5 разных защитных масок. Каждый информант проговаривал 83 высказывания на русском языке в течение 6 сессий, в общей сложности BRAVE-MASKS содержит 498 высказываний. Суммарная длительность всех речевых высказываний с 3-х устройств записи для каждого информанта составляет приблизительно 130 минут. Затем, используя эти же маски, записывались вращения головой с 7 различных точек в комнате (с расстояния 1, 2 и 3 метра под разными углами). Оптическое разрешение видеоданных – 4К 3840×2160 пикселей, частота кадров – 30 (для планшета) и 60 (для смартфонов) кадров в секунду, цветность – 24 бита на пиксель. Аудиофайлы без сжатия в

формате PCM WAV с частотой дискретизации 48 кГц, 16 бит на отсчет.

Сбор многомодальных корпусов является затруднительным процессом, так как он требует наличия большого количества времени, участников, специального оборудования. Также основной проблемой сбора многомодальных корпусов является аннотирование данных, так как важно чтобы метки данных каждой модальности были синхронизированы между собой. К тому же с наступлением коронавирусной пандемии многие люди оказались на самоизоляции, поэтому появилась еще одна проблема сбора корпусов – это сложность в получении однородных данных (с одинаковым освещением, идентичным звуковым качеством записей и тому подобные).

5. Заключение. Возрастающее количество работ, направленное на детектирование масок на лице человека как по видеоинформации, так и по аудиоинформации, показывает актуальность данной задачи в нынешнее время. Существует много исследований на тему обнаружения масок по видеоизображениям, также в открытом доступе можно найти значительное количество корпусов, содержащих изображения лиц как без масок, так и в масках, полученных различными способами. Исследований и разработок, направленных на детектирование средств индивидуальной защиты органов дыхания по акустическим характеристиках речи человека, пока достаточно мало, так как это направление начало развиваться только в период пандемии, вызванной коронавирусной инфекцией COVID-19. Ранее проводились исследования, косвенно относящиеся к проблемам обнаружения масок по речи информанта, например, обнаружение различных респираторных заболеваний по голосу, влияние защитных масок на разборчивость и восприятие речи.

Существующие системы позволяют предотвратить распространение коронавирусной инфекции с помощью распознавания наличия/отсутствия масок на лице, также данные системы помогают в дистанционном диагностировании COVID-19 с помощью обнаружения первых симптомов вирусной инфекции по акустическим характеристикам. Однако на сегодняшний день существует ряд нерешенных проблем в области автоматического диагностирования COVID-19 и наличия/отсутствия масок на лицах людей. В первую очередь это низкая точность обнаружения масок и коронавирусной инфекции, что не позволяет осуществлять автоматическую диагностику без присутствия экспертов (медицинского персонала). Многие системы не способны работать в режиме реального времени, из-за чего невозможно производить контроль и мониторинг ношения защитных масок в обществен-

ных местах. Также большинство существующих систем невозможно встроить в смартфон, чтобы пользователи могли в любом месте произвести диагностирование наличия коронавирусной инфекции. Еще одной основной проблемой является сбор данных пациентов, зараженных COVID-19, так как многие люди не согласны распространять конфиденциальную информацию.

На основе этого можно выделить следующие требования, выдвигаемые к разрабатываемым автоматическим системам определения наличия/отсутствия масок на лице человека:

1. Использование видео- и аудио модальностей. В том случае, когда анализ одной модальности невозможен, можно получить информацию о наличии/отсутствии маски с помощью другой модальности. К тому же анализ обеих модальностей позволяет более точно производить детектирование наличие/отсутствия маски на лицах людей.

2. Высокая точность распознавания наличия/отсутствия маски на лицах. В период пандемии особенно важно максимально точно распознавать СИЗ на лицах людей. Разработанные системы должны детектировать наличия/отсутствия масок с точностью более 98%.

3. Системы видеоанализа помимо детектирования наличия/отсутствия масок должны распознавать тип маски и выявлять окклюзии на лицах различными предметами, отличные от СИЗ. А также уметь различать реальные маски от, например, нарисованных на лицах красками.

4. Системы аудиоанализа должны работать со звуками с различных микрофонов, различной дальности от информантов, учитывать телефонные разговоры, работать как в помещениях, так и в общественных местах, учитывать возможный эффект реверберации. Кроме того, такие системы должны быть независимыми от информантов, то есть система должна быть построена при обучении на одних информантах, а проверка системы должна быть выполнена на других информантах.

5. Также системы должны распознавать неправильно надетые маски. Многие люди носят маски, не закрывая носовую и/или ротовую полость, что представляет риск заражения вирусной инфекцией.

6. Тестирование систем необходимо производить в реальных условиях, учитывая такие параметры, как различное освещение, дальность объекта от камеры, посторонние шумовые звуки, плохое качество связи. Также при обучении использовать изображения лиц и речевые записи людей в реальных масках, а не синтетических.

В дальнейших исследованиях планируется обучить классификаторы распознавания наличия/отсутствия маски на лице человека по

аудио- и видеoinформации с помощью собранного авторами текущей статьи корпуса BRAVE-MASKS. А также на основе этих данных построить системы распознавания типа маски и верификации информантов с масками на лицах.

Литература

1. Habib A. et al. Global Epidemiology of COVID-19 and the Risk of Second Wave. *Journal of Drug Delivery and Therapeutics*. 2021. vol. 11. no. 1. pP. 1–2.
2. Иванов В.А., Часовская Ю.С. Маски-индивидуальные средства защиты от воздушно-капельных инфекций // Интегративные тенденции в медицине и образовании. 2020. Т. 3. С. 47–53.
3. Boškoski I., Gallo C., Wallace M.B., Costamagna G. COVID-19 pandemic and personal protective equipment shortage: protective efficacy comparing masks and scientific methods for respirator reuse. *Gastrointestinal endoscopy*. 2020. vol. 92. no. 3. P. 519–523.
4. Macintyre C.R., Chughtai A.A. Facemasks for the prevention of infection in healthcare and community settings. *Bmj*. 2015. vol. 350.
5. Abdulwhhab M.T. Use of Face-Mask Sampling as a Means of Characterising the Microbiota Exhaled from Human Respiratory Tract in Health and Disease: дис. – University of Leicester. 2020.
6. Нагиев М.Р., Нестерова Н.В. Анализ осведомленности населения об эффективности использования одноразовых медицинских масок в профилактике ОРЗ и ОРВИ, а также перспектива использования лигнина гидролизного в их усовершенствовании // Молодой ученый. 2020. №. 20. С. 207–211.
7. Jiang F. et al. Review of the clinical characteristics of coronavirus disease 2019 (COVID-19). *Journal of general internal medicine*. 2020. vol. 35. no. 5. pp. 1545–1549.
8. Badillo-Goicoechea E., Chang T-H., Kim E., LaRocca S., Morris K., Deng X., Chiu S., Bradford A., Garcia A., Kern C., Cobb C., Kreuter F., Stuart E.A. Global trends and predictors of face mask usage during the COVID-19 pandemic. *arXiv preprint arXiv:2012.11678*. 2020.
9. Eikenberry S.E. et al. To mask or not to mask: Modeling the potential for face mask use by the general public to curtail the COVID-19 pandemic. *Infectious Disease Modelling*. 2020. vol. 5. pp. 293–308.
10. Гольдштейн Э.М. Факторы, влияющие на смертность от новой коронавирусной инфекции в разных субъектах Российской Федерации // Журнал микробиологии, эпидемиологии и иммунобиологии. 2021. Т. 97. №. 6. С. 604–607.
11. Мусихин И.Г. и другие. Ношение медицинских масок как эффективный способ защиты от covid-19 // Современное общество: опыт, проблемы и перспективы развития. 2021. С. 5–17.
12. Chughtai A.A., Seale H., Macintyre C.R. Effectiveness of cloth masks for protection against severe acute respiratory syndrome coronavirus 2. *Emerging infectious diseases*. 2020. vol. 26. no. 10.
13. Singh A. et al. Social perception and practices of households regarding mask use in public places during COVID-19 postquarantine period. *BLDE University Journal of Health Sciences*. 2020. vol. 5. no. 2. P. 209.
14. Rahimi Z. et al. Face mask use among pedestrians during the Covid-19 pandemic in Southwest Iran: an observational study on 10,440 people. *BMC Public Health*. 2021. vol. 21. no. 1. pp. 1–9.

15. Haischer M.H. et al. Who is wearing a mask? Gender-, age-, and location-related differences during the COVID-19 pandemic. *PLoS one*. 2020. vol. 15. no. 10. P. e0240785.
16. Peretti-Watel P. et al. Attitudes about COVID-19 lockdown among general population, France, March 2020. *Emerging infectious diseases*. 2021. vol. 27. no. 1. pp. 301–303.
17. Буркова В.Н., Феденок Ю.Н. Медицинская маска как средство индивидуальной и коллективной защиты в условиях пандемии COVID-19 (кросс-культурные аспекты) // *Вестник антропологии*. (Herald of Anthropology) 2021. Т. 51. №. 3. С. 74–91.
18. Natnael T. et al. Facemask wearing to prevent COVID-19 transmission and associated factors among taxi drivers in Dessie City and Kombolcha Town, Ethiopia. *PLoS one*. 2021. vol. 16. no. 3. P. e0247954.
19. Gunasekaran G.H. et al. Prevalence and acceptance of face mask practice among individuals visiting hospital during COVID-19 pandemic: an observational study. *Preprints 2020*. 2020.
20. Ge S. et al. Detecting masked faces in the wild with lle-cnns. *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017. pp. 2682–2690.
21. Roy B. et al. MOXA: A Deep Learning Based Unmanned Approach For Real-Time Monitoring of People Wearing Medical Masks. *Transactions of the Indian National Academy of Engineering*. 2020. vol. 5. no. 3. pp. 509–518.
22. Faisal N., Wasiq K., Salwa Y., Abir H. Face Mask Detection Video Dataset. *Mendeley Data*. 2020.
23. Wang Z. et al. Masked face recognition dataset and application. *arXiv preprint arXiv:2003.09093*. 2020.
24. Huang B. et al. When Face Recognition Meets Occlusion: A New Benchmark. *ICASSP*. 2021. pp. 4240–4244.
25. Yi D. et al. Learning face representation from scratch. *arXiv preprint arXiv:1411.7923*. 2014.
26. Loey M. et al. A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic. *Measurement*. 2021. vol. 167. P. 108288.
27. Learned-Miller E. et al. Labeled faces in the wild: A survey. *Advances in face detection and facial image analysis*. 2016. pp. 189–248.
28. Chen Y. et al. Adversarial occlusion-aware face detection. 2018 *IEEE 9th International Conference on Biometrics Theory, Applications and Systems (BTAS)*. 2018. pp. 1–9.
29. Loey M. et al. Fighting against COVID-19: A novel deep learning model based on YOLO-v2 with ResNet-50 for medical face mask detection. *Sustainable cities and society*. 2021. vol. 65. P. 102600.
30. Ryumina E., Ryumin D., Ivanko D., Karpov A. Novel Method for Protective Face Mask Detection Using Convolutional Neural Networks and Image Histograms. *International Archives of the Photogrammetry Remote Sensing and Spatial Information Sciences*. 2021. vol. XLIV-2/W1-2021. pp. 177–182.
31. He K., Zhang X., Ren S., Sun J. Deep residual learning for image recognition. *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016. pp. 770–778.
32. Redmon J., Farhadi A. Yolov3: An incremental improvement. *arXiv preprint arXiv:1804.02767*. 2018.
33. Sandler M. et al. Mobilenetv2: Inverted residuals and linear bottlenecks. *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018. pp. 4510–4520.

34. Nagrath P. et al. SSDMV2: A real time DNN-based face mask detection system using single shot multibox detector and MobileNetV2. *Sustainable cities and society*. 2021. vol. 66. P. 102692.
35. Liu W. et al. Ssd: Single shot multibox detector. *Lecture Notes in Computer Science*. 2016. vol. 9905. P. 21–37.
36. Anisimov D., Khanova T. Towards lightweight convolutional neural networks for object detection. 2017 14th IEEE international conference on advanced video and signal based surveillance (AVSS). 2017. pp. 1–8.
37. Mohan P., Paul A.J., Chirania A.A. Tiny CNN Architecture for Medical Face Mask Detection for Resource-Constrained Endpoints. *Innovations in Electrical and Electronic Engineering. Lecture Notes in Electrical Engineering*. 2021. vol. 756.
38. Вашкевич М.И., Азаров И.С. Определение патологии голосового аппарата на основе анализа модуляционного спектра речи в критических полосах. // *Труды СПИИРАН*. 2020. № 2 (19). С. 249–276.
39. Авдеев В.Б., Трушин В.А., Кунгуров М.А. Унифицированная речеподобная помеха для средств активной защиты речевой информации // *Информатика и автоматизация*. 2020. № 5 (19). С. 991–1017.
40. Dvoynikova A., Verkholyak O., Karpov A. Emotion Recognition and Sentiment Analysis of Extemporaneous Speech Transcriptions in Russian. *Lecture Notes in Computer Science*. 2020. vol. 12335 LNAI. pp. 136–144.
41. Deshpande G., Schuller B.W. Audio, Speech, Language, & Signal Processing for COVID-19: A Comprehensive Overview. *arXiv preprint arXiv:2011.14445*. 2020.
42. Monge-Alvarez J. et al. Robust detection of audio-cough events using local hu moments. *IEEE journal of biomedical and health informatics*. 2018. vol. 23. vol. 1. pp. 184–196.
43. Schuller B., et al. The Interspeech 2017 computational paralinguistics challenge: Addressee, cold & snoring. *INTERSPEECH*. 2017. pp. 3442–3446.
44. Sharma N. et al. Coswara A Database of Breathing, Cough, and Voice Sounds for COVID-19 Diagnosis. *INTERSPEECH*. 2020. pp. 4811–4815.
45. Brown C., Chauhan J., Grammenos A. et al. Exploring Automatic Diagnosis of COVID-19 from Crowdsourced Respiratory Sound Data. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD '20)*. 2020. pp. 3474–3484.
46. Stasak B. et al. Automatic Detection of COVID-19 Based on Short-Duration Acoustic Smartphone Speech Analysis. *Journal of Healthcare Informatics Research*. 2021. vol. 5. Is. 2. P. 201–207.
47. Saeidi R., Niemi T., Karppelin H., Pohjalainen J., Kinnunen T., Alku P. Speaker recognition for speech under face cover. *INTERSPEECH*. 2015. pp. 1012–1016.
48. Schuller B., Batliner A., Bergler C., Messner E., Hamilton A., Amiriparian S., Baird A., Rizos G. The INTERSPEECH 2020 Computational paralinguistics challenge: Elderly emotion, Breathing & Masks. *INTERSPEECH*. 2020. pp. 2042–2046.
49. Montacié C., Caraty M. Phonetic, Frame Clustering and Intelligibility Analyses for the INTERSPEECH 2020 ComParE Challenge. *INTERSPEECH*. 2020. pp. 2062–2066.
50. Radeck-Armeth S., Milde B. et al. Open source german distant speech recognition: Corpus and acoustic model. *International Conference on Text, Speech, and Dialogue*. 2015. pp. 480–488.
51. Matos S. et al. Detection of cough signals in continuous audio recordings using hidden Markov models. *IEEE Transactions on Biomedical Engineering*. 2006. vol. 53. vol. 6. pp. 1078–1083.
52. Monge-Alvarez J. et al. Audio-cough event detection based on moment theory. *Applied Acoustics*. 2018. vol. 135. pp. 124–135.

53. Gosztolya G., Busa-Fekete R., Grósz T., Tóth L. DNN-based feature extraction and classifier combination for child-directed speech, cold and snoring identification. *INTERSPEECH*. 2017. pp. 3522–3526.
54. Schuller B., Batliner A., Bergler C., et al. The INTERSPEECH 2021 Computational Paralinguistics Challenge: COVID-19 Cough, COVID-19 Speech, Escalation & Primates. *INTERSPEECH*. 2021. P. 5.
55. Schuller B.W., Coppock H., Gaskell A. Detecting COVID-19 from Breathing and Coughing Sounds using Deep Neural Networks. arXiv preprint arXiv:2012.14553. 2020.
56. Klumpp P., et al The Phonetic Footprint of Covid-19?. *INTERSPEECH*. 2021.
57. Xia T. et al. Uncertainty-Aware COVID-19 Detection from Imbalanced Sound Data. arXiv preprint arXiv:2104.02005. 2021.
58. Muguli A. et al. DiCOVA Challenge: Dataset, task, and baseline system for COVID-19 diagnosis using acoustics. arXiv preprint arXiv:2103.09148. 2021.
59. Mendel L.L., Gardino J.A., Atcherson S.R. Speech understanding using surgical masks: a problem in health care?. *Journal of the American Academy of Audiology*. 2008. vol. 19. no. 9. pp. 686–695.
60. Cohn M., Pycha A., Zellou G. Intelligibility of face-masked speech depends on speaking style: Comparing casual, clear, and emotional speech. *Cognition*. 2021. vol. 210. P. 104570.
61. Kalikow D.N., Stevens K.N., Elliott L.L. Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *The Journal of the acoustical society of America*. 1977. vol. 61. no. 5. pp. 1337–1351.
62. Pörschmann C., Lübeck T., Arend J.M. Impact of face masks on voice radiation. *The Journal of the Acoustical Society of America*. 2020. vol. 148. no. 6. pp. 3663–3670.
63. Saeidi R., Huhtakallio I., Alku P. Analysis of Face Mask Effect on Speaker Recognition. *INTERSPEECH*. 2016. pp. 1800–1804.
64. Weninger F., Eyben F., Schuller B., Mortillaro M., Scherer K. On the Acoustics of Emotion in Audio: What Speech, Music and Sound have in Common. *Frontiers in Emotion Science*. 2013. vol. 4. pp. 1–12.
65. Schmitt M., Schuller B. openXBOW – Introducing the Passau Open-Source Crossmodal Bag-of-Words Toolkit. *Journal of Machine Learning Research*. 2017. vol. 18. pp. 1–5.
66. Freitag M., Amiriparian S., Pugachevskiy S., Cummins N., Schuller B. AuDeep: Unsupervised Learning of Representations from Audio with Deep Recurrent Neural Networks. *Journal of Machine Learning Research*. 2018. vol. 18. pp. 1–5.
67. Amiriparian S., Gerczuk M., Ottl S., Cummins N., Freitag M., Pugachevski S., Schuller B. Snore sound classification using image-based deep spectrum features. *INTERSPEECH*. 2017. pp. 3512–3516.
68. Yang Z., An Z., Fan Z., Jing C., Cao H. Exploration of Acoustic and Lexical Cues for the INTERSPEECH 2020 Computational Paralinguistic Challenge. *INTERSPEECH*. 2020. pp. 2092–2096.
69. Klumpp P., Arias-Vergara T., Vázquez-Correa J., Pérez-Toro P., Hönig F., Nöth E., Orozco-Arroyave J. Surgical Mask Detection with Deep Recurrent Phonetic Models. *INTERSPEECH*. 2020. pp. 2057–2061.
70. Illium S., Müller R., Sedlmeier A., Linnhoff-Popien C. Surgical Mask Detection with Convolutional Neural Networks and Data Augmentations on Spectrograms. *INTERSPEECH*. 2020. pp. 2052–2056.
71. Ristea N., Ionescu R. Are you Wearing a Mask? Improving Mask Detection from Speech Using Augmentation by Cycle-Consistent GANs. *INTERSPEECH*. 2020. pp. 2102–2106.

72. Koike T., Qian K., Schuller B., Yamamoto Y. Learning Higher Representations from Pre-Trained Deep Models with Data Augmentation for the COMPARE 2020 Challenge Mask Task. INTERSPEECH. 2020. pp. 2047–2051.
73. Goodfellow I., Pouget-Abadie J., Mirza M., Xu B., Warde-Farley D., Ozair S., Courville A., Bengio Y. Generative adversarial networks. In Proceedings of the 27th International Conference on Neural Information Processing Systems. 2014. vol. 2. pp. 2672–2680.
74. Szep J., Hariri S. Paralinguistic Classification of Mask Wearing by Image Classifiers and Fusion. INTERSPEECH. 2020. pp. 2087–2091.
75. Simonyan K., Zisserman A. Very Deep Convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556. 2014. P. 14.
76. Huang G., Liu Z., Van Der Maaten L., Weinberger K. Densely connected convolutional networks. Proceedings of the IEEE conference on computer vision and pattern recognition. 2017. pp. 4700–4708.
77. Krizhevsky A., Sutskever I., Hinton G. Imagenet classification with deep convolutional neural networks. Communications of the ACM. 2017. vol. 60. vol. 6. pp. 84–90.
78. Markitantov M. et al. Ensembling end-to-end deep models for computational paralinguistics tasks: ComParE 2020 Mask and Breathing Sub-challenges. INTERSPEECH. 2020. P. 2666.
79. Schuller B.W. et al. Covid-19 and computer audition: An overview on what speech & sound analysis could contribute in the sars-cov-2 corona crisis. arXiv preprint arXiv:2003.11117. 2020.
80. Fecher N. The "audio-visual face cover corpus": investigations into audio-visual speech and speaker recognition when the speaker's face is occluded by facewear. INTERSPEECH. 2012. pp. 2250–2253.
81. Корпус аудиовизуальных русскоязычных данных людей в защитных масках (BRAVE-MASKS - Biometric Russian Audio-Visual Extended MASKS corpus). Свидетельство о государственной регистрации Базы данных № 2021621094 от 26.05.2021, авторы: Маркитантов М.В., Рюмин Д.А., Рюмина Е.В., Карпов А.А., правообладатель: СПб ФИЦ РАН.

Двойникова Анастасия Александровна — младший научный сотрудник, лаборатория речевых и многомодальных интерфейсов, СПб ФИЦ РАН. Область научных интересов: искусственный интеллект, машинное обучение, нейронные сети, распознавание защитных масок по аудиоинформации, сентимент-анализ текстовых данных. Число научных публикаций — 10. dvoynikova.a@iias.spb.su; 14-я линия В.О., 39, 199178, Санкт-Петербург, Россия; р.т.: +7(812) 328 04 21.

Маркитантов Максим Викторович — младший научный сотрудник, лаборатория речевых и многомодальных интерфейсов, СПб ФИЦ РАН. Область научных интересов: искусственный интеллект, машинное обучение, речевые технологии, компьютерная паралингвистика, распознавание характеристик диктора, распознавание пола и возраста диктора, обнаружение защитных масок по аудиоинформации. Число научных публикаций — 10. m.markitantov@yandex.ru; 14-я линия В.О., 39, 199178, Санкт-Петербург, Россия; р.т.: +7 (812) 328 04 21.

Рюмина Елена Витальевна — младший научный сотрудник, лаборатория речевых и многомодальных интерфейсов, СПб ФИЦ РАН. Область научных интересов: аффективные вычисления, цифровая обработка изображений, распознавание визуальных сигналов, автоматическое распознавание паралингвистических явлений, машинное обучение, нейронные сети, биометрические системы, человеко-машинные интерфейсы.

Число научных публикаций — 12. gyumina.e@iiias.spb.su; 14-я линия В.О., 39, 199178, Санкт-Петербург, Россия; р.т.: +7 (812) 328 04 21.

Рюмин Дмитрий Александрович — канд. техн. наук, старший научный сотрудник, лаборатория речевых и многомодальных интерфейсов, СПб ФИЦ РАН. Область научных интересов: цифровая обработка изображений, распознавание образов, автоматическое распознавание визуальной речи, многомодальные интерфейсы, машинное обучение, нейронные сети, биометрия, человеко-машинные интерфейсы. Число научных публикаций — 41. gyumin.d@iiias.spb.su; 14-я линия В.О., 39, 199178, Санкт-Петербург, Россия; р.т.: +7 (812) 328 04 21.

Карпов Алексей Анатольевич — заведующий лабораторией, лаборатория речевых и многомодальных интерфейсов, СПб ФИЦ РАН. Область научных интересов: речевые технологии, автоматическое распознавание речи, обработка аудиовизуальной речи, многомодальные человеко-машинные интерфейсы, компьютерная паралингвистика и другие. Число научных публикаций — 300+. karpov@iiias.spb.su; 14-я линия В.О., 39, 199178, Санкт-Петербург, Россия; р.т.: +7 (812) 328 04 21.

Поддержка исследований. Исследование выполнено при поддержке Российского Фонда Фундаментальных Исследований № 20-04-60529, а также частично в рамках бюджетной темы № 0073-2019-0005.

A. DVOYNIKOVA, M. MARKITANTOV, E. RYUMINA, D. RYUMIN, A. KARPOV
**ANALYTICAL REVIEW OF AUDIOVISUAL SYSTEMS FOR
DETERMINING PERSONAL PROTECTIVE EQUIPMENT ON A
PERSON'S FACE**

Dvoynikova A., Markitantov M., Ryumina E., Ryumin D., Karpov A. Analytical Review of Audiovisual Systems for Determining Personal Protective Equipment on a Person's Face.

Abstract. Since 2019 all countries of the world have faced the rapid spread of the pandemic caused by the COVID-19 coronavirus infection, the fight against which continues to the present day by the world community. Despite the obvious effectiveness of personal respiratory protection equipment against coronavirus infection, many people neglect the use of protective face masks in public places. Therefore, to control and timely identify violators of public health regulations, it is necessary to apply modern information technologies that will detect protective masks on people's faces using video and audio information. The article presents an analytical review of existing and developing intelligent information technologies for bimodal analysis of the voice and facial characteristics of a masked person. There are many studies on the topic of detecting masks from video images, and a significant number of cases containing images of faces both in and without masks obtained by various methods can also be found in the public access. Research and development aimed at detecting personal respiratory protection equipment by the acoustic characteristics of human speech is still quite small, since this direction began to develop only during the pandemic caused by the COVID-19 coronavirus infection. Existing systems allow to prevent the spread of coronavirus infection by recognizing the presence/absence of masks on the face, and these systems also help in remote diagnosis of COVID-19 by detecting the first symptoms of a viral infection by acoustic characteristics. However, to date, there is a number of unresolved problems in the field of automatic diagnosis of COVID-19 and the presence/absence of masks on people's faces. First of all, this is the low accuracy of detecting masks and coronavirus infection, which does not allow for performing automatic diagnosis without the presence of experts (medical personnel). Many systems are not able to operate in real time, which makes it impossible to control and monitor the wearing of protective masks in public places. Also, most of the existing systems cannot be built into a smartphone, so that users be able to diagnose the presence of coronavirus infection anywhere. Another major problem is the collection of data from patients infected with COVID-19, as many people do not agree to distribute confidential information.

Keywords: identification of protective masks, voice characteristics of speakers, facial characteristics, COVID-19, personal protective equipment, cough detection.

Dvoynikova Anastasia — Junior researcher, Laboratory of speech and multimodal interfaces, SPC RAS. Research interests: artificial intelligence, machine learning, neural networks, recognition of protective masks by audio information, sentiment analysis. The number of publications — 10. dvoynikova.a@iias.spb.su; 39, 14-th Line V.O., 199178, St. Petersburg, Russia; office phone: +7(812) 328 04 21.

Markitantov Maxim — Junior researcher, Laboratory of speech and multimodal interfaces, SPC RAS. Research interests: artificial intelligence, machine learning, speech technologies, computational paralinguistics, recognition of the speaker's characteristics, speaker's age and gender recognition, detection of protective masks by audio information. The number of publications — 10. m.markitantov@yandex.ru; 39, 14-th Line V.O., 199178, St. Petersburg, Russia; office phone: +7 (812) 328 04 21.

Ryumina Elena — Junior researcher, Laboratory of speech and multimodal interfaces, SPC RAS. Research interests: Affective computing, digital image processing, visual signal recognition, automatic recognition of paralinguistic phenomena, machine learning, neural networks, biometric systems, human-machine interfaces. The number of publications — 12. ryumina.e@iias.spb.su; 39, 14-th Line V.O., 199178, St. Petersburg, Russia; office phone: +7 (812) 328 04 21.

Ryumin Dmitry — Ph.D., Senior researcher, Laboratory of speech and multimodal interfaces, SPC RAS. Research interests: digital image processing, pattern recognition, automatic visual speech recognition, multimodal interfaces, machine learning, neural networks, biometrics, human-machine interfaces. The number of publications — 41. ryumin.d@iias.spb.su; 39, 14-th Line V.O., 199178, St. Petersburg, Russia; office phone: +7 (812) 328 04 21.

Karpov Alexey — Head of laboratory, Laboratory of speech and multimodal interfaces, SPC RAS. Research interests: speech technology, automatic speech recognition, audio-visual speech processing, multimodal human-computer interfaces, and computational paralinguistics. The number of publications — 300+. karpov@iias.spb.su; 39, 14-th Line V.O., 199178, St. Petersburg, Russia; office phone: +7 (812) 328 04 21.

Acknowledgements. This research was supported by the Russian Foundation for Basic Research № 20-04-60529, as well as partly within the framework of the budget theme No. 0073-2019-0005.

References

1. Habib A. et al. Global Epidemiology of COVID-19 and the Risk of Second Wave. *Journal of Drug Delivery and Therapeutics*. 2021. vol. 11. no. 1. pP. 1–2.
2. Ivanov V.A., Chasovskaya Y.S. [Masks - personal protective equipment against airborne infections]. *Integrativnye tendencii v medicine i obrazovanii* [Integrative trends in medicine and education]. 2020. vol. 3. pP. 47–53. (In Russ).
3. Boškoski I., Gallo C., Wallace M.B., Costamagna G. COVID-19 pandemic and personal protective equipment shortage: protective efficacy comparing masks and scientific methods for respirator reuse. *Gastrointestinal endoscopy*. 2020. vol. 92. no. 3. P. 519–523.
4. Macintyre C.R., Chughtai A.A. Facemasks for the prevention of infection in healthcare and community settings. *Bmj*. 2015. vol. 350.
5. Abdulwhhab M.T. Use of Face-Mask Sampling as a Means of Characterising the Microbiota Exhaled from Human Respiratory Tract in Health and Disease: дис. – University of Leicester. 2020.
6. Nagiev M.R., Nesterova N.V. [Analysis of public awareness about the effectiveness of using disposable medical masks in the prevention of acute respiratory infections and acute respiratory viral infections, as well as the prospect of using hydrolytic lignin in their improvement]. *Molodoj uchenyj* [Young scientist]. 2020. no. 20. pP. 207–211. (In Russ).
7. Jiang F. et al. Review of the clinical characteristics of coronavirus disease 2019 (COVID-19). *Journal of general internal medicine*. 2020. vol. 35. no. 5. pp. 1545–1549.
8. Badillo-Goicoechea E., Chang T-H., Kim E., LaRocca S., Morris K., Deng X., Chiu S., Bradford A., Garcia A., Kern C., Cobb C., Kreuter F., Stuart E.A. Global trends and predictors of face mask usage during the COVID-19 pandemic. *arXiv preprint arXiv:2012.11678*. 2020.

9. Eikenberry S.E. et al. To mask or not to mask: Modeling the potential for face mask use by the general public to curtail the COVID-19 pandemic. *Infectious Disease Modelling*. 2020. vol. 5. pp. 293–308.
10. Goldstein E.M. [Factors affecting mortality for the novel coronavirus infection in different regions of the Russian Federation]. *Zhurnal mikrobiologii, epidemiologii i immunobiologii [Journal of Microbiology, Epidemiology and Immunobiology]*. 2021. vol. 97. no. 6. pp. 604–607. (In Russ).
11. Musikhin I.G. et al. [Wearing medical masks as an effective way of protection against covid-19]. *Sovremennoye obshchestvo: opyt. problemy i perspektivy razvitiya [Modern society: experience, problems and development prospects]*. 2021. pp. 5–17. (In Russ).
12. Chughtai A.A., Seale H., Macintyre C.R. Effectiveness of cloth masks for protection against severe acute respiratory syndrome coronavirus 2. *Emerging infectious diseases*. 2020. vol. 26. no. 10.
13. Singh A. et al. Social perception and practices of households regarding mask use in public places during COVID-19 postquarantine period. *BLDE University Journal of Health Sciences*. 2020. vol. 5. no. 2. P. 209.
14. Rahimi Z. et al. Face mask use among pedestrians during the Covid-19 pandemic in Southwest Iran: an observational study on 10,440 people. *BMC Public Health*. 2021. vol. 21. no. 1. pp. 1–9.
15. Haischer M.H. et al. Who is wearing a mask? Gender-, age-, and location-related differences during the COVID-19 pandemic. *PLoS one*. 2020. vol. 15. no. 10. P. e0240785.
16. Peretti-Watel P. et al. Attitudes about COVID-19 lockdown among general population, France, March 2020. *Emerging infectious diseases*. 2021. vol. 27. no. 1. pp. 301–303.
17. Burkova V.N., Fedenok J.N. [Medical mask as a means of personal and collective protection in the context of the COVID-19 pandemic (cross-cultural aspects)]. *Vestnik antropologii [Herald of Anthropology]*. 2021. vol. 51. no. 3. pp. 74–91. (In Russ).
18. Natnael T. et al. Facemask wearing to prevent COVID-19 transmission and associated factors among taxi drivers in Dessie City and Kombolcha Town, Ethiopia. *PLoS one*. 2021. vol. 16. no. 3. P. e0247954.
19. Gunasekaran G.H. et al. Prevalence and acceptance of face mask practice among individuals visiting hospital during COVID-19 pandemic: an observational study. *Preprints 2020*. 2020.
20. Ge S. et al. Detecting masked faces in the wild with lle-cnns. *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017. pp. 2682–2690.
21. Roy B. et al. MOXA: A Deep Learning Based Unmanned Approach For Real-Time Monitoring of People Wearing Medical Masks. *Transactions of the Indian National Academy of Engineering*. 2020. vol. 5. no. 3. pp. 509–518.
22. Faisal N., Wasiq K., Salwa Y., Abir H. Face Mask Detection Video Dataset. *Mendeley Data*. 2020.
23. Wang Z. et al. Masked face recognition dataset and application. *arXiv preprint arXiv:2003.09093*. 2020.
24. Huang B. et al. When Face Recognition Meets Occlusion: A New Benchmark. *ICASSP*. 2021. pp. 4240–4244.
25. Yi D. et al. Learning face representation from scratch. *arXiv preprint arXiv:1411.7923*. 2014.
26. Loey M. et al. A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic. *Measurement*. 2021. vol. 167. P. 108288.

27. Learned-Miller E. et al. Labeled faces in the wild: A survey. *Advances in face detection and facial image analysis*. 2016. pp. 189–248.
28. Chen Y. et al. Adversarial occlusion-aware face detection. 2018 IEEE 9th International Conference on Biometrics Theory, Applications and Systems (BTAS). 2018. pp. 1–9.
29. Loey M. et al. Fighting against COVID-19: A novel deep learning model based on YOLO-v2 with ResNet-50 for medical face mask detection. *Sustainable cities and society*. 2021. vol. 65. P. 102600.
30. Ryumina E., Ryumin D., Ivanko D., Karpov A. Novel Method for Protective Face Mask Detection Using Convolutional Neural Networks and Image Histograms. *International Archives of the Photogrammetry Remote Sensing and Spatial Information Sciences*. 2021. vol. XLIV-2/W1-2021. pp. 177–182.
31. He K., Zhang X., Ren S., Sun J. Deep residual learning for image recognition. *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016. pp. 770–778.
32. Redmon J., Farhadi A. Yolov3: An incremental improvement. *arXiv preprint arXiv:1804.02767*. 2018.
33. Sandler M. et al. Mobilenetv2: Inverted residuals and linear bottlenecks. *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018. pp. 4510–4520.
34. Nagrath P. et al. SSDMNV2: A real time DNN-based face mask detection system using single shot multibox detector and MobileNetV2. *Sustainable cities and society*. 2021. vol. 66. P. 102692.
35. Liu W. et al. Ssd: Single shot multibox detector. *Lecture Notes in Computer Science*. 2016. vol. 9905. P. 21–37.
36. Anisimov D., Khanova T. Towards lightweight convolutional neural networks for object detection. 2017 14th IEEE international conference on advanced video and signal based surveillance (AVSS). 2017. pp. 1–8.
37. Mohan P., Paul A.J., Chirania A.A. Tiny CNN Architecture for Medical Face Mask Detection for Resource-Constrained Endpoints. *Innovations in Electrical and Electronic Engineering. Lecture Notes in Electrical Engineering*. 2021. vol. 756.
38. Vashkevich M., Azarov I. [Voice Pathology Detection based on Analysis of Modulation Spectrum in Critical Bands]. *Trudy SPIIRAN [SPIIRAS Proceedings]*. 2020. no 2 (19). pp. 249–276. (In Russ).
39. Avdeev V., Trushin V., Kungurov M. [Unified Speech-Like Interference for Active Protection of Speech Information]. *Informatika i avtomatizacija [Informatics and Automation]*. 2020. no 5 (19). pp. 991–1017. doi: 10.15622/ia.2020.19.5.4. (In Russ).
40. Dvoynikova A., Verkholyak O., Karpov A. Emotion Recognition and Sentiment Analysis of Extemporaneous Speech Transcriptions in Russian. *Lecture Notes in Computer Science*. 2020. vol. 12335 LNAI. pp. 136–144.
41. Deshpande G., Schuller B.W. Audio, Speech, Language, & Signal Processing for COVID-19: A Comprehensive Overview. *arXiv preprint arXiv:2011.14445*. 2020.
42. Monge-Alvarez J. et al. Robust detection of audio-cough events using local hu moments. *IEEE journal of biomedical and health informatics*. 2018. vol. 23. vol. 1. pp. 184–196.
43. Schuller B., et al. The Interspeech 2017 computational paralinguistics challenge: Addressee, cold & snoring. *INTERSPEECH*. 2017. pp. 3442–3446.
44. Sharma N. et al. Coswara A Database of Breathing, Cough, and Voice Sounds for COVID-19 Diagnosis. *INTERSPEECH*. 2020. pp. 4811–4815.
45. Brown C., Chauhan J., Grammenos A. et al. Exploring Automatic Diagnosis of COVID-19 from Crowdsourced Respiratory Sound Data. In *Proceedings of the 26th*

- ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD '20). 2020. pp. 3474–3484.
46. Stasak B. et al. Automatic Detection of COVID-19 Based on Short-Duration Acoustic Smartphone Speech Analysis. *Journal of Healthcare Informatics Research*. 2021. vol. 5. Is. 2. P. 201–207.
47. Saeidi R., Niemi T., Karppelein H., Pohjalainen J., Kinnunen T., Alku P. Speaker recognition for speech under face cover. *INTERSPEECH*. 2015. pp. 1012–1016.
48. Schuller B., Batliner A., Bergler C., Messner E., Hamilton A., Amiriparian S., Baird A., Rizos G. The INTERSPEECH 2020 Computational paralinguistics challenge: Elderly emotion, Breathing & Masks. *INTERSPEECH*. 2020. pp. 2042–2046.
49. Montacié C., Caraty M. Phonetic, Frame Clustering and Intelligibility Analyses for the INTERSPEECH 2020 ComParE Challenge. *INTERSPEECH*. 2020. pp. 2062–2066.
50. Radeck-Arneth S., Milde B. et al. Open source german distant speech recognition: Corpus and acoustic model. *International Conference on Text, Speech, and Dialogue*. 2015. pp. 480–488.
51. Matos S. et al. Detection of cough signals in continuous audio recordings using hidden Markov models. *IEEE Transactions on Biomedical Engineering*. 2006. vol. 53. vol. 6. pp. 1078–1083.
52. Monge-Alvarez J. et al. Audio-cough event detection based on moment theory. *Applied Acoustics*. 2018. vol. 135. pp. 124–135.
53. Gosztolya G., Busa-Fekete R., Grósz T., Tóth L. DNN-based feature extraction and classifier combination for child-directed speech, cold and snoring identification. *INTERSPEECH*. 2017. pp. 3522–3526.
54. Schuller B., Batliner A., Bergler C., et al. The INTERSPEECH 2021 Computational Paralinguistics Challenge: COVID-19 Cough, COVID-19 Speech, Escalation & Primates. *INTERSPEECH*. 2021. P. 5.
55. Schuller B.W., Coppock H., Gaskell A. Detecting COVID-19 from Breathing and Coughing Sounds using Deep Neural Networks. *arXiv preprint arXiv:2012.14553*. 2020.
56. Klumpp P., et al The Phonetic Footprint of Covid-19?. *INTERSPEECH*. 2021.
57. Xia T. et al. Uncertainty-Aware COVID-19 Detection from Imbalanced Sound Data. *arXiv preprint arXiv:2104.02005*. 2021.
58. Muguli A. et al. DiCOVA Challenge: Dataset, task, and baseline system for COVID-19 diagnosis using acoustics. *arXiv preprint arXiv:2103.09148*. 2021.
59. Mendel L.L., Gardino J.A., Atcherson S.R. Speech understanding using surgical masks: a problem in health care?. *Journal of the American Academy of Audiology*. 2008. vol. 19. vol. 9. pp. 686–695.
60. Cohn M., Pycha A., Zellou G. Intelligibility of face-masked speech depends on speaking style: Comparing casual, clear, and emotional speech. *Cognition*. 2021. vol. 210. P. 104570.
61. Kalikow D.N., Stevens K.N., Elliott L.L. Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *The Journal of the acoustical society of America*. 1977. vol. 61. vol. 5. pp. 1337–1351.
62. Pörschmann C., Lübeck T., Arend J.M. Impact of face masks on voice radiation. *The Journal of the Acoustical Society of America*. 2020. vol. 148. vol. 6. pp. 3663–3670.
63. Saeidi R., Huhtakallio I., Alku P. Analysis of Face Mask Effect on Speaker Recognition. *INTERSPEECH*. 2016. pp. 1800–1804.
64. Weninger F., Eyben F., Schuller B., Mortillaro M., Scherer K. On the Acoustics of Emotion in Audio: What Speech, Music and Sound have in Common. *Frontiers in Emotion Science*. 2013. vol. 4. pp. 1–12.

65. Schmitt M., Schuller B. openXBOW – Introducing the Passau Open-Source Crossmodal Bag-of-Words Toolkit. *Journal of Machine Learning Research*. 2017. vol. 18. pp. 1–5.
66. Freitag M., Amiriparian S., Pugachevskiy S., Cummins N., Schuller B. AuDeep: Unsupervised Learning of Representations from Audio with Deep Recurrent Neural Networks. *Journal of Machine Learning Research*. 2018. vol. 18. pp. 1–5.
67. Amiriparian S., Gerczuk M., Ottl S., Cummins N., Freitag M., Pugachevskiy S., Schuller B. Snore sound classification using image-based deep spectrum features. *INTERSPEECH*. 2017. pp. 3512–3516.
68. Yang Z., An Z., Fan Z., Jing C., Cao H. Exploration of Acoustic and Lexical Cues for the INTERSPEECH 2020 Computational Paralinguistic Challenge. *INTERSPEECH*. 2020. pp. 2092–2096.
69. Klumpp P., Arias-Vergara T., Vásquez-Correa J., Pérez-Toro P., Hönig F., Nöth E., Orozco-Aroyave J. Surgical Mask Detection with Deep Recurrent Phonetic Models. *INTERSPEECH*. 2020. pp. 2057–2061.
70. Illium S., Müller R., Sedlmeier A., Linnhoff-Popien C. Surgical Mask Detection with Convolutional Neural Networks and Data Augmentations on Spectrograms. *INTERSPEECH*. 2020. pp. 2052–2056.
71. Ristea N., Ionescu R. Are you Wearing a Mask? Improving Mask Detection from Speech Using Augmentation by Cycle-Consistent GANs. *INTERSPEECH*. 2020. pp. 2102–2106.
72. Koike T., Qian K., Schuller B., Yamamoto Y. Learning Higher Representations from Pre-Trained Deep Models with Data Augmentation for the COMPARE 2020 Challenge Mask Task. *INTERSPEECH*. 2020. pp. 2047–2051.
73. Goodfellow I., Pouget-Abadie J., Mirza M., Xu B., Warde-Farley D., Ozair S., Courville A., Bengio Y. Generative adversarial networks. In *Proceedings of the 27th International Conference on Neural Information Processing Systems*. 2014. vol. 2. pp. 2672–2680.
74. Szep J., Hariri S. Paralinguistic Classification of Mask Wearing by Image Classifiers and Fusion. *INTERSPEECH*. 2020. pp. 2087–2091.
75. Simonyan K., Zisserman A. Very Deep Convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*. 2014. P. 14.
76. Huang G., Liu Z., Van Der Maaten L., Weinberger K. Densely connected convolutional networks. *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017. pp. 4700–4708.
77. Krizhevsky A., Sutskever I., Hinton G. Imagenet classification with deep convolutional neural networks. *Communications of the ACM*. 2017. vol. 60. vol. 6. pp. 84–90.
78. Markitantov M. et al. Ensembling end-to-end deep models for computational paralinguistics tasks: ComParE 2020 Mask and Breathing Sub-challenges. *INTERSPEECH*. 2020. P. 2666.
79. Schuller B.W. et al. Covid-19 and computer audition: An overview on what speech & sound analysis could contribute in the sars-cov-2 corona crisis. *arXiv preprint arXiv:2003.11117*. 2020.
80. Fecher N. The "audio-visual face cover corpus": investigations into audio-visual speech and speaker recognition when the speaker's face is occluded by facewear. *INTERSPEECH*. 2012. pp. 2250–2253.
81. Корпус аудиовизуальных русскоязычных данных людей в заштитных масках (BRAVE-MASKS - Biometric Russian Audio-Visual Extended MASKS corpus). Свидетельство о государственной регистрации Базы данных № 2021621094. 26.05.2021. Authors: Markitantov M.V., Ryumin D.A., Ryumina E.V., Karpov A.A., pravoobladatel: SPC RAS. (In Russ).

А.В. БОГОМОЛОВ
**ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ ЦИФРОВОЙ
АДАПТАЦИОННОЙ МЕДИЦИНЫ**

Богомолов А.В. Информационные технологии цифровой адаптационной медицины.

Аннотация. В статье дана комплексная характеристика информационных технологий цифровой адаптационной медицины. Акцент сделан на применимость к разработке специализированных автоматизированных комплексов, программных моделей и систем изучения адаптационных возможностей человека к условиям внешней среды. Сформулированы требования к информационным технологиям повышения этих возможностей. Отражены особенности информационных технологий применительно к проведению прикладных системных исследований обеспечения жизнедеятельности, сохранения профессионального здоровья и продления долголетия человека.

Охарактеризованы шесть базовых концепций адаптационной медицины с акцентом на особенности математического обеспечения обработки информации, определены приоритеты совершенствования информационных технологий, применяемых в этих концепциях.

Рассмотрены информационные технологии, применяемые в задачах обеспечения профессиональной работоспособности человека с акцентом на необходимость применения адекватных методов диагностики состояния человека на всех этапах профессиональной деятельности и необходимости разработки технологий цифровых двойников, адекватно моделирующих адаптационные процессы и реакции организма в реальных условиях.

Дана характеристика информационных технологий персонализированного мониторинга рисков здоровью, позволяющих объективизировать воздействия физических факторов условий деятельности и реализовать индивидуальное и коллективное информирование персонала об опасности окружающей среды.

Показана насущная необходимость стандартизации методов обработки информации при разработке информационных технологий цифровой адаптационной медицины в интересах обеспечения физиологической адекватности и математической корректности подходов к получению и обработке информации о состоянии человека.

Сделаны выводы о том, что приоритеты совершенствования информационных технологий цифровой адаптационной медицины связаны с внедрением достижений четвертой промышленной революции, в том числе, концепции социкиберфизических систем.

Ключевые слова: цифровая медицина, цифровое здравоохранение, медицинская информатика, адаптационная медицина, профилактическая медицина, социкиберфизическая система.

1. Введение. Интенсивное развитие цифровых технологий во всех сферах жизнедеятельности человека и их ускоренное внедрение в экономике и социальной сфере отнесено к национальным целям и стратегическим задачам развития Российской Федерации. Одним из ключевых приоритетов достижения этих целей и решения поставленных задач является разработка и реализация технологий цифровой медицины [1-3].

Цифровая медицина – отрасль медицины, ключевыми факторами которой являются данные в цифровом виде, обработка больших объемов и использование результатов анализа которых по сравнению с традиционными формами позволяют существенно повысить эффективность профилактики, диагностики, лечения и реабилитации [4-7].

Анализируя достижения цифровой медицины, следует отметить, что подавляющее большинство из них относятся к созданию:

«умных больниц» (мониторинг траекторий перемещения пациентов и врачей, адресная доставка лекарств, телемедицина, технологии жизнеобеспечения пациентов в клинических условиях и т.п.) [8-12];

технологий индивидуального мониторинга состояния человека («умные часы», фитнес-браслеты, датчики-регистраторы и программное обеспечение для смартфонов и др.) [13-16];

технологий медицинской диагностики (обработка медицинских изображений, медицинские информационные системы, интеллектуальные медицинские системы и т.п.) [17-20];

систем электронного документооборота (электронные медицинские книжки, электронные назначения, технологии «одного окна» и др.) [21-29].

Вместе с тем, в последнее время существенно повышается роль *профилактической медицины*, направленной на предупреждение возникновения заболеваний и устранение (минимизацию) факторов риска их развития. При этом переход от здоровья к болезни рассматривается как стадийно протекающий процесс адаптации организма к условиям среды, в ходе которого возникновению качественно нового состояния – болезни – предшествуют изменения напряжения регуляторных механизмов [30, 31]. В связи с этим акцент с фиксации процесса перехода здорового в больного смещается на контроль резервов здоровья в интересах недопущения их истощения [32-34].

Реализация такого контроля основана на концепции *персонализированной* (персонализированной) медицины, предполагающей реализацию медицинского обеспечения пациентов с учетом индивидуальных особенностей состояния как всего организма, так и его отдельных органов, тканей, клеток (в том числе, на генетическом уровне) с акцентами на прогностичность (мониторинг индивидуальных рисков здоровью), проактивность (индивидуальная профилактика заболеваний) и превентивность (ориентация на лечение пациента, а не болезни).

Для достижения целей профилактической медицины необходимо изучение адаптационных возможностей человека к условиям внешней среды, разработка методов и средств для повышения этих возможностей, а также проведение прикладных системных исследований ме-

дицинских, физиолого-гигиенических, психологических, инженерно-психологических и эргономических аспектов обеспечения жизнедеятельности, эффективности, надежности деятельности, сохранения профессионального здоровья и продления профессионального долголетия человека [35-37]. Решением этих задач занимается *адаптационная медицина*.

В этом аспекте ключевое значение для достижений национальных целей и решения стратегических задач развития Российской Федерации имеет развитие *информационных технологий*, определяющих процессы, методы поиска, сбора, хранения, обработки, предоставления, распространения информации и способы осуществления таких процессов и методов при решении задач адаптационной медицины.

Информационные технологии цифровой адаптационной медицины имеют те же ограничения и требования к реализации, что и информационные технологии цифровой клинической медицины: обеспечение надежной работы, целостности собираемых и хранимых данных, а также доказательности, точности, достоверности и физиологической адекватности алгоритмов обработки информации и их реализации в виде программного и аппаратного обеспечения.

Однако анализ современных отечественных и зарубежных публикаций в области цифровой медицины показал, что информационным технологиям цифровой адаптационной медицины уделяется недостаточное внимание.

Существенное значение для адаптационной медицины имеют информационные технологии персонализированного (персонализированного) мониторинга состояния человека, однако в подавляющем большинстве они не доведены до уровня интеграции в концепции адаптационной медицины, а технические средства и программное обеспечение мониторинга не верифицированы и во многих случаях не обеспечивают адекватное решение задач определения:

функциональных возможностей организма по отношению к тем нагрузкам, в условиях которых осуществляется жизнедеятельность человека;

способностей организма восстанавливать свои резервы в отведенное трудовым регламентом время;

потенциальных возможностей пополнения резервов здоровья.

Основная причина этого очевидна: решение задач клинической медицины приносит существенную экономическую прибыль в ближайшей, а решение задач профилактической медицины – в отдаленной перспективе.

Цель статьи - дать общую характеристику информационных технологий цифровой адаптационной медицины для координации усилий специалистов различных предметных областей по их развитию в соответствии с достижениями научно-технического прогресса и требованиями адаптационной медицины.

2. Базовые концепции адаптационной медицины. Современные информационные технологии цифровой адаптационной медицины реализуют концепции, мониторинга и повышения адаптационных возможностей человека, важнейшими из которых являются [38, 39].

1. *Профессиональное здоровье* как характеристика генетически детерминированного и эволюционного процесса нормально протекающего биологического старения человека в конкретной социально-экономической формации с акцентом на обеспечение надежной профессиональной деятельности, здоровья и долголетия индивида [40].

Обеспечение профессионального здоровья обуславливает необходимость использования соответствующих информационных технологий, а также реализации управления компенсаторными свойствами и резервами организма.

Математическое обеспечение этой концепции основано на методах взвешенных сверток первичных показателей в интегральные, характеризующие различные аспекты профессионального здоровья, а также на технологиях адаптивной индивидуальной нормы его показателей [39, 41].

Примером реализации концепции является компьютерная методика «Навигатор здоровья», позволяющая сформировать индивидуальный паспорт здоровья и рекомендации по повышению уровня функциональных резервов организма [33]. По этой методике обследовано более 650 тысяч человек в возрасте от 6 до 60 лет, разделенных на 17 возрастных групп, построены рейтинговые шкалы для балльной оценки показателей здоровья, используемые при расчете индекса физического здоровья [33, 39]. Показано, что система восстановительных мероприятий, реализованная на базе концепции, позволила снизить общую заболеваемость в 1,7 раза, повысить надежность профессиональной деятельности в 1,6 раза и продлить профессиональное долголетие на 20-30% [39].

2. *Профилактическая каскадная концепция изменения здоровья*, отличительной чертой которой является реализация дозового рискометрического подхода к действию факторов риска здоровью и соответствующая классификация направлений его обеспечения: от медицинского и профессионального отбора до восстановления и реабилитации [41, 42]. На основе каскадной схемы реализуют важнейшие для

персонализированной адаптационной медицины технологии планирования затрат с точки зрения полезности того или иного медицинского вмешательства для изменения качества, продолжительности жизни и «витальной» потребности индивида [42].

Математическое обеспечение этой концепции составляют методы построения кривых «доза-эффект», характеризующих зависимости реакций биологического объекта на воздействия стресс-факторов различных интенсивностей [42].

Концепция успешно апробирована при решении задач управления рисками здоровью операторов атомных электростанций, ликвидаторов последствий радиационных аварий, персонала объектов по уничтожению химического оружия и работников химически опасных производств [41, 42].

3. *Концепция синдромосходных состояний*, основанная на результатах, подтверждающих, что изменение здоровья вследствие воздействия различных физических факторов сопровождается синдромосходным комплексом изменений на молекулярно-клеточном и тканевом уровнях [41, 42].

Математическое обеспечение этой концепции составляют методы деревьев решений, экстраполяции по принципу «модель-фактор» и мягких вычислений, позволяющие получать решающие правила синдромной диагностики, устойчивые как к неточности знания, заключенного в решающем правиле, так и к неточности, связанной с возможной вариабельностью характеристик состояния человека. За счет этого обеспечивается высокое качество синдромной диагностики, позволившее, в том числе, реализовать персонализированные программы профилактики и повышения адаптационных возможностей человека [41].

Концепция синдромосходных состояний успешно применена для решения комплекса практических задач медицинского обеспечения специалистов опасных профессий, подвергающихся воздействию физических факторов экстремальной и субэкстремальной интенсивности, экспериментальное изучение медико-биологических эффектов с участием испытуемых-добровольцев которых не представляется возможным [41, 42].

4. *Биологический возраст и профессиональное долголетие*, в понятиях которых возрастные изменения функциональных систем индивида сопоставляются со средними популяционными изменениями на основе биомаркеров старения (показателей состояния человека, значительно изменяющихся с возрастом и имеющих малый индивидуальный разброс) [39, 41, 43]. Биологический возраст, помимо наследственно-

сти, в большой степени зависит от условий жизнедеятельности и образа жизни индивида. Выявляемые различия хронологического и биологического возраста позволяют оценить интенсивность старения и функциональные возможности индивида [44, 45]. Концепция ориентирована на выработку стратегии коррекции наиболее вероятных или уже наступивших неблагоприятных изменения здоровья с учетом диапазона возможных изменений активности функциональных систем организма, который может быть обеспечен его активационными механизмами [46].

Математическое обеспечение этой концепции составляют методы сокращения размерности пространства признаков, методы восстановления зависимостей и методы синтеза интегральных показателей [41]. Для определения биологического возраста и профессионального долголетия разработано большое число специальных программно-аппаратных комплексов, ориентированных на представителей различных социо-профессиональных групп населения.

Концепция биологического возраста и профессионального долголетия успешно применена при решении комплекса задач мониторинга адаптационного потенциала летного состава государственной авиации, населения территорий, неблагоприятных в экологическом отношении, и персонала, работающего на таких территориях [42-46].

5. Концепция *качества жизни*, связанного со здоровьем, как интегральной характеристики физического, эмоционального, психологического и социального благополучия человека, основанная на его субъективной оценке, позволяющая дать многоплановый анализ важнейших составляющих здоровья [41, 42, 47]. Известно более 200 методик квалиметрии жизни, среди которых наиболее распространены опросники, рекомендованные Всемирной организацией здравоохранения, а также опросники SF-36 и EQ-5D.

Особенностью автоматизированных систем исследования качества жизни является ориентированность на выявление скрытых факторов риска здоровью, что позволяет обеспечить «расшифровку» причин риска и необходимые фактографические данные. Такие технологии эффективны для реализации в качестве интернет-приложений, предназначенных для самоконтроля здоровья с персонифицированным управлением его рисками.

Математическое обеспечение этой концепции ориентировано на обработку результатов анкетирования с использованием компьютерных вопросников. Широкое распространение при обработке результатов анкетирования получили методы многофакторного анализа, методы аналитических сетей (анализа иерархий), а также метод анализа и

синтеза показателей состояния сложных систем при информационном дефиците [41, 48, 49].

Концепция качества жизни успешно применена для решения задач анализа изменения здоровья населения во взаимосвязи с состоянием окружающей среды в интересах разработки и реализации технологий управления рисками здоровью, исследования соотношений между затратами на лечение и его эффективностью (фармакоэкономика), стандартизации методов лечения, разработки прогностических моделей развития болезней и т.п. [50, 51]. Показатели качества жизни населения являются ключевыми показателями эффективности реализации приоритетных национальных проектов долгосрочного социально-экономического развития Российской Федерации.

6. Концепция *рискометрии здоровья здоровых*, позволяющая обеспечить поддержку принятия решений о принадлежности текущего состояния человека системы к одному из альтернативных классов состояний с определением выраженности проявления признаков состояния определенного класса [41].

Концепция ориентирована на анализ, наряду со среднестатистическими (популяционными) показателями, необходимыми для планирования медицинской помощи, единичных отклонений в показателях состояния. Такие отклонения не считаются артефактами, а подлежат анализу, поскольку могут свидетельствовать о наличии донозологических изменений, которые способны перейти в патологические изменения. Причем аномальные значения показателей состояния анализируются системно – во взаимосвязи с показателями состояния других функциональных систем организма, множество которых определяется, основываясь на теории доминанты и теории функциональных систем организма [41, 51].

Математическое обеспечение этой концепции составляют методы многоуровневого характера, включающие: математические методы компьютерного анкетирования и анализа его результатов на первом уровне, математические методы обработки результатов лабораторных и инструментальных исследований – на втором уровне и математические методы анализа многомерных данных на заключительном уровне позволяют эффективно решать задачи персонализированного мониторинга резервов здоровья.

Концепция рискометрии здоровья здоровых эффективно применена при решении комплекса практических задач обеспечения работоспособности и сохранения здоровья операторов эргатических систем, в том числе, летного состава государственной авиации [41].

Несмотря на то, что базовые медико-биологические концепции адаптационной медицины созданы более 30 лет назад, подтверждены практикой и позволяют решать задачи цифровой адаптационной медицины, информационные технологии, реализующие эти концепции, нуждаются в совершенствовании. Приоритетом такого совершенствования должно быть внедрение в медико-биологические концепции современных:

технологий искусственного интеллекта, когнитивных вычислений и математической логики;

алгоритмов обработки больших массивов («BigData») структурных и иных физиологических данных, анализа временных рядов, сигналов, изображений и видеоданных, а также текстов и символьных последовательностей;

моделей молекулярно-генетических, биофизических, экосистемных и биосферных процессов, а также физиологической организации поведения при интеграции сенсорных, когнитивных и управляющих процессов;

технологий интернета вещей, разработки, сопровождения и анализа программ и информационно-коммуникационных систем;

вычислительных методов и алгоритмов для компьютерных систем высокой производительности, глобальных и интегрированных информационно-вычислительных и телекоммуникационных систем и сетей, облачных и грид-технологий;

достижений высокопроизводительных вычислений, нано- и микросистемной техники;

технологий управления распределенными вычислительными средами на основе технологий распределенного реестра (блокчейн и смарт-контрактов).

В свою очередь новые возможности сбора и обработки информации, открывающиеся за счет реализации современных информационных технологий, обусловят развитие изложенных и создание новых базовых медико-биологических концепций адаптационной медицины. Прогресс в реализации фундаментальных научных достижений в прикладные теории, технологии и методы, которые позволяют связать воедино достижения различных предметных областей для решения задач адаптационной медицины, связывается с прогрессом *трансляционной медицины*.

Специфической особенностью трансляционной медицины является реализации двухкомпонентного замкнутого контура обмена информацией [52, 53]:

от исследования к пациенту, направленного на повышение эффективности испытаний новых технологий, разработанных в результате научных исследований,

от пациента к исследованию, предполагающего обратную связь относительно применения технологий и определения приоритетов их совершенствования.

Прогресс трансляционной медицины существенным образом зависит от применяемых информационных технологий, ключевой значение имеют методы технологии системного анализа (абстрагирование и конкретизация, анализ и синтез, индукция и дедукция, формализация и конкретизация, реинжиниринг и другие).

3. Информационные технологии в задачах обеспечения профессиональной работоспособности человека. *Профессиональная работоспособность* человека отражает характеристики систем и функций организма, необходимые для выполнения профессиональных задач. Исходя из этого информационные технологии обеспечения профессиональной работоспособности человека должны обеспечивать диагностику состояния человека на всех этапах профессиональной деятельности и подготовки к ней.

Исследования показали, что встречающиеся на практике ситуации, требующие диагностики состояния человека в аспекте обеспечения его профессиональной работоспособности, подразделяются на пять типов [54]:

1) выявление устойчивых свойств организма человека, необходимых для надёжного выполнения задач профессиональной деятельности или препятствующих этому (профессиональный отбор);

2) решение о допуске конкретных специалистов к выполнению задач профессиональной деятельности (предсменный контроль);

3) определение максимально возможной продолжительности успешного решения профессиональных задач при воздействии неблагоприятных факторов условий деятельности;

4) выявление изменений состояния человека в процессе деятельности;

5) объективизация психофизиологических явлений, не имеющих выраженного прикладного характера (исследование активации внимания, мотивации и т.п.).

Одним из ключевых аспектов обеспечения надёжной деятельности, реализуемый в рамках подготовки к ней, является профессиональный отбор. Технологии профессионального отбора специалистов различных социо-профессиональных групп отработаны десятилетиями и включают определение множества профессионально важных качеств,

формирование системы тестов, позволяющих оценить выраженность этих качеств, и методы синтеза интегральных показателей, характеризующих профессиональную пригодность индивида.

Среди современных трендов совершенствования профессионального отбора – применение для решения его задач информационных технологий анализа полиморфизма генов, определяющих успешность адаптации индивида к нагрузкам, специфичным для конкретной профессиональной деятельности. Так, в результате экспериментальных исследований доказана актуальность скринингового исследования полиморфизма генов ACTN3, TFAM, PPARA и PPARGC1A при проведении военно-профессионального отбора для прохождения службы в подразделениях, выполняющих специальные задачи, связанные с высокими физическими нагрузками [55], известны результаты эффективного применения таких технологий в спортивной медицине [56].

Состояние человека в процессе деятельности изменяется под воздействием внешних и внутренних факторов ее условий. Переносимость человеком воздействия внешних неблагоприятных факторов условий деятельности зависит не только от их вида, продолжительности, интенсивности воздействия, но и от внутренних факторов («исходное» состояние человека, его опытность, общая физическая тренированность, психофизиологическая готовность и др.). Общие и специфические эффекты изменения состояния человека при воздействии факторов условий деятельности говорят о существовании закономерностей, в том числе [41, 42]:

немонотонная нелинейная зависимость изменения состояния человека от интенсивности и продолжительности внешних воздействий;

фазность, цикличность изменений характеристик состояния при длительном внешнем воздействии;

отсутствие однозначности и параллелизма физиологических и психологических сдвигов состояния (часто важную роль играет целевая установка).

Следует также учитывать, что в отличие от клинической медицины адаптационная медицина оперирует с состояниями человека, характеризующимися индивидуальной специфичностью, невыраженностью и обратимостью.

Ключевой особенностью изучения состояний человека является понимание их как реакций, формируемых организмом. Одним из наиболее важных моментов при этом является наличие комплекса причин, определяющих состояние человека в каждой конкретной ситуации, и выделение множества факторов, вносящих определённый

вклад в формирование ответной реакции организма. Качественная неоднородность разных состояний обуславливается, прежде всего, различиями в основных причинах, их вызывающих, и в условиях, в которых осуществляется воздействие факторов условий деятельности в каждом конкретном случае.

Таким образом, *состояние человека* – это сложная системная реакция, являющаяся результатом взаимодействия функциональных систем его организма и определяющая профессиональную работоспособность человека. Однако многообразие возможных состояний человека приводит к тому, что построить их универсальную классификацию не удаётся [41].

Поэтому вопрос выбора адекватного подхода к диагностике состояния в адаптационной медицине рассматривается в том смысле, насколько тот или иной подход является перспективным для дальнейшего продвижения, углубления в понимании проблемы и возможностях поиска средств контроля, прогнозирования и управления состоянием человека.

Несмотря на различие изложенных типовых постановок задач диагностики состояний человека – статичность или динамичность, оценивание или прогнозирование состояния – в формальном математическом смысле они являются тождественными, позволяющими свести задачу диагностики состояния человека к задаче *распознавания образов* [41]. В общем случае решение задачи распознавания образов заключается в разработке на основании исходных данных (набора описаний образов с указанием некоторой идентифицирующей характеристики) процедуры, позволяющей идентифицировать образ по его описанию с минимальным риском неправильной идентификации и, при необходимости, сравнить различные образы по некоторому критерию. При таком подходе к постановке задачи диагностики состояний человека целесообразно использовать их разделение на экстенсивные и интенсивные.

Экстенсивные состояния имеют разную основу, выраженное качественное своеобразие: сопоставление их внутри некоторой единой шкалы невозможно. К классу экстенсивных состояний относятся, например, диагнозы, имеющиеся у человека (согласно расписанию болезней и международной классификации болезней). Экстенсивные состояния являются «фоном» для развития интенсивных состояний, определяя устойчивые свойства организма человека. Информационные технологии распознавания экстенсивных состояний ориентированы на отнесение текущего состояния к одному из типов (классов), как правило, с указанием количественной оценки степени (меры) соответствия.

Интенсивные состояния имеют принципиальное сходство по своему содержанию: их можно упорядочить посредством единой шкалы, а изменение состояния может быть представлено в виде подвижной точки внутри этой шкалы. К классу интенсивных состояний человека относятся, например, функциональные состояния (утомление, эмоциональная и операционная напряженность, монотония и т.п.). Информационные технологии распознавания интенсивных состояний ориентированы на получение количественной оценки выраженности текущего состояния применительно к заранее заданным классам.

Таким образом, целью диагностики состояния является получение его оценки, под которой понимают заранее описанный класс (вид, тип), к которому отнесено текущее состояние в результате его идентификации.

При этом существенные различия в используемых на практике способах описания состояний человека не позволяют говорить о возможности решения всех типовых задач диагностики состояния с помощью какого-либо одного математического метода и требуют специального рассмотрения.

Формальное описание состояния человека задаётся совокупностью (кортежем) его характеристик (параметров, индексов, показателей). Используемые на практике способы описания состояния можно разделить на два типа: описания динамическим рядом характеристик и описания набором характеристик (табл. 1).

Таблица 1. Сопоставление типов описания состояния человека

Описание	Динамическим рядом характеристик	Набором характеристик
Размерность описания	До нескольких тысяч характеристик	До нескольких десятков характеристик
Физиологическая интерпретация характеристик	Одинаковая	Различная

Выбор математического метода диагностики состояния определяется особенностями его описания (параметрами или показателями/индексами) состояния и типом исходных данных (гетероассоциативные или автоассоциативные).

Для обеспечения корректности математического обеспечения диагностики состояния человека при сборе исходных данных необходимо гарантировать их однородность, репрезентативность, достаточность и отсутствие аномальных наблюдений, а выбор показателей качества созданного математического обеспечения диагностики состоя-

ния человека следует производить исходя из того, к какому типу – экстенсивных или интенсивных – относится оцениваемое состояние.

При апробации математического метода диагностики экстенсивных состояний в качестве характеристик качества следует использовать: чувствительность, специфичность, прогностическую ценность положительного и отрицательного результата, отношение правдоподобия положительного и отрицательного результата, а при апробации математического метода диагностики интенсивных состояний – число допущенных гиподиагностических (первого рода) и гипердиагностических (второго рода) ошибок.

При этом в любом случае необходимо обеспечивать возможность интерпретации результатов диагностики состояния человека медицинским персоналом – понимания «внутреннего содержания» математического подхода к обработке информации о состоянии. С этих позиций, в частности, обуславливается сложность применения для построения интегральных показателей и решающих правил диагностики состояния человека на основе искусственных нейронных сетей. Решающие правила, реализованные на основе искусственных нейронных сетей, крайне сложны для интерпретации – поэтому такие результаты диагностики состояния человека подвергаются сомнению и воспринимаются с недоверием. Кроме того, имеет место проблема переобучения искусственной нейронной сети, которая состоит в неадекватной реакции на примеры, не применявшиеся при ее настройке: формируются существенно разные выходные данные для близких входных данных.

Реализация технологий обеспечения профессиональной работоспособности тесно связана с разработкой информационных технологий решения задач мониторинга состояния человека в условиях комбинированного воздействия факторов условий деятельности. При решении этих задач необходимо учитывать специфические особенности:

одинаковый медико-биологический эффект могут дать суммация или аддитивное взаимодействие факторов, потенциация, антагонизм и их комбинирование (потенциация при малых величинах и антагонизм – при больших и наоборот);

медико-биологические эффекты и зависимости равных эффектов соотношений «интенсивность-длительность» для большинства комбинаций факторов условий деятельности человека недостаточно изучены.

Решение задач диагностики состояния человека в условиях комбинированного воздействия факторов условий деятельности традиционно связывается с применением информационных технологий экс-

траполяции медико-биологических эффектов воздействия с животных на человека и методов планирования экспериментов [42, 57].

Прогресс в решении этих задач, важных для адапционной медицины, связывается с развитием технологий *цифровых двойников* – программных комплексов, адекватно моделирующих адаптационные процессы и реакции организма в реальных условиях. С применением технологий цифровых двойников появится возможность комплексных исследований реакций функциональных систем организма на изолированное и комбинированное воздействие факторов условий деятельности (в том числе, экстремальной интенсивности) без проведения экспериментов с участием человека. С развитием этих технологий также связывают прогресс в развитии средств и методов проектирования защитного снаряжения, эргономической экспертизы рабочих мест и оптимизации условий деятельности человека.

4. Персонализированный мониторинг рисков здоровью. Достижения информационных технологий открывают новые возможности мониторинга рисков здоровью, обусловленных воздействием факторов условий деятельности, в интересах решения задач адапционной медицины. Цифровые технологии обеспечивают прогресс в двух направлениях:

объективизация воздействия физических факторов условий деятельности;

реализация индивидуального и коллективного информирования персонала об опасности окружающей среды.

Объективизация воздействия физических факторов условий деятельности предполагает получение информации о дозах факторов условий деятельности, полученных индивидом в процессе рабочих смен и/или за любой период выполнения деятельности. Это достигается за счет реализации персонализированного гигиенического мониторинга. С этой целью на предприятиях промышленности должно быть создано рабочее место мониторинга рисков, на обмундировании (снаряжении) каждого работника перед началом выполнения профессиональной деятельности должен закрепляться транспондер (метка радиочастотной идентификации – *RFID-метка*), а каждый объект – источник потенциально опасных физических факторов условий профессиональной деятельности должен быть оборудован параметрическим регистратором [58, 59].

Рабочее место мониторинга рисков должно обеспечивать ведение медико-дозиметрического гигиенического регистра, содержащего информацию о дозах физических факторов, получаемых каждым работником в процессе профессиональной деятельности.

Тип применяемых транспондеров (RFID-меток) по рабочей частоте, по источнику питания, по типу памяти, по исполнению и другим характеристикам выбирают в зависимости от специфики организаций (структурных подразделений) и от специфики профессиональной деятельности работника. В частности, за работником может быть закреплен уникальный идентификатор (для использования работником изготавливается один или несколько транспондеров с уникальным идентификатором) или же транспондер может выдаваться работнику перед началом деятельности (в этом случае идентификатор транспондера каждый раз связывается с конкретным работником) [60].

Параметрический регистратор, устанавливаемый на каждом объекте – источнике потенциально опасных физических факторов условий профессиональной деятельности, должен обеспечивать:

фиксацию в реальном времени перемещения нескольких маркированных объектов (работников, имеющих транспондер), либо идентификацию их положения в пространстве;

расчет интенсивности физического фактора, образующегося при функционировании оборудования (в зависимости от режима функционирования) в точке нахождения каждого фиксируемого объекта в реальном времени;

беспроводную передачу комбинации «идентификатор объекта – интенсивность физического фактора» на рабочее место мониторинга рисков.

Таким образом, при включении промышленного оборудования активируется параметрический регистратор, радиус считывания транспондеров (радиус идентификации) которого должен превышать расстояние, на котором вероятность превышения максимально возможной интенсивности физических факторов, сопровождающих функционирование оборудования, предельно допустимых уровней отлична от нуля.

При идентификации работника (попадании транспондера в зону идентификации) параметрический регистратор обеспечивает расчет в реальном времени интенсивности физического фактора в точке нахождения каждого фиксируемого объекта (работника). Расчет может осуществляться как на основании верифицированных математических моделей физических полей, так и на результатах интерполяции прямых измерений полей физических факторов, полученных при санитарно-гигиенической паспортизации объектов. Частота (дискретность) расчетов определяется для каждого объекта (или для каждого физического фактора, если объект является источником нескольких потенци-

ально опасных физических факторов) в зависимости от специфики решаемых задач [61, 62].

Информационные технологии, применяемые при решении этих задач, обеспечивают сопоставление траектории работника с расположением и режимами работы источников физических факторов, определяя дозы факторов, реально полученные работником в процессе деятельности. Полученная информация учитывается в *медико-дозиметрическом регистре*, обеспечивающем определение текущих и кумулятивных доз физических факторов, полученных работником, с привязкой к информации об индивидуальных особенностях здоровья. Это обеспечивает выработку персонифицированных рекомендаций по сохранению профессионального здоровья и принятие управленческих решений, направленных на его сохранение [60].

Реализация *индивидуального информирования* персонала об опасности окружающей среды осуществляется с помощью применения персонифицированных индикаторов опасности окружающей среды. Такие индикаторы могут быть построены по традиционной схеме – как миниатюрные дозиметры. Однако такой подход существенно усложняет конструктивное исполнение индикаторов (особенно при необходимости одновременного мониторинга нескольких физических факторов).

Перспективным является реализация индикаторов в изложенной системе объективизации воздействия физических факторов. Текущие риски работоспособности и здоровью рассчитываются в реальном времени по результатам сопоставления траектории работника с расположением и режимом работы источников физических факторов с учетом эффекта «защита временем», характеристик защитного снаряжения, применяемого работником, и индивидуальных особенностей его здоровья.

При получении сигнала об опасности окружающей обстановки, в зависимости от особенностей деятельности, работник должен зайти в средство коллективной защиты, перейти на другой участок работы и т.п.

Реализация *коллективного информирования* персонала об опасности окружающей среды обеспечивается с помощью специальных свето/звукосигнальных табло, размещаемых в местах, видимых большому количеству работников. В состав табло входят измерители интенсивностей и времени экспозиции физических факторов условий деятельности и вычислитель, рассчитывающий уровень опасности окружающей обстановки. В случае сигнала опасности работник должен применить средство индивидуальной защиты, зайти в средство

коллективной защиты либо реализовать другие меры, определенные инструкцией по охране труда [60].

Изложенный подход к мониторингу рисков здоровью успешно апробирован в процессе обеспечения надежной деятельности инженерно-технического состава государственной авиации [60] и может быть распространен на любую социо-профессиональную группу работников, подвергающихся воздействию потенциально опасных физических факторов условий деятельности.

5. Стандартизация методов обработки информации при разработке информационных технологий цифровой адаптационной медицины. Важнейшим аспектом эффективного применения информационных технологий цифровой адаптационной медицины является обеспечение адекватности подходов к сбору и обработке информации [63]. При этом необходимо учитывать:

1. Физиологическую адекватность подходов к получению информации о состоянии человека. Информационные технологии диагностики состояния человека для решения задач адаптационной медицины должны быть ориентированы на рассмотрение его как особого психофизиологического явления со специфическими закономерностями, которое заложено в архитектуре функциональных систем организма и проявляется на биохимическом, физиологическом, психологическом и поведенческом уровнях с учетом изменения эффективности выполняемой деятельности, проявляющейся в ее результативности, внутренней цене и содержании [41, 64-67].

Методы диагностики состояний человека в процессе деятельности должны ориентироваться на то, что каждому состоянию соответствует специфическая доминанта, т. е. для обеспечения формирования и развития этого состояния мобилизуются различные функциональные системы организма, причем степень их вовлечения в формирование состояния различная у каждого индивида (этим, собственно, и объясняется индивидуальность реакций человека на воздействие факторов условий деятельности).

Однако анализ опыта диагностики состояния человека, особенно с применением свободно распространяемых приложений для мобильных устройств, показывает, что разработчики таких приложений часто не учитывают специфику формирования диагностируемых состояний человека. Например, нервно-эмоциональное напряжение диагностируется только по изменению частоты пульса без учета информации о состоянии центральной нервной системы.

2. Математическую корректность подходов к сбору и обработке информации о состоянии человека. Разработчики информационных

технологий в большинстве случаев считают информацию об особенностях применяемого математического обеспечения коммерческой тайной, что не позволяет проверить его корректность и обеспечить сопоставимость идентичных (по сути) результатов, полученных различными программно-аппаратными комплексами. Например, при расчете оценок одних и тех же спектральных показателей используют различные окна, процедуры заполнения нулями и другие техники спектрального анализа, приводящие к несопоставимости его результатов [68-72].

В частности, исследования [73], проведенные путем ввода в альтернативные образцы приборов, выполняющих спектральный анализ ритмокардиограмм, 23 идентичных тестовых последовательностей QRS-комплексов с различными характеристиками мощности в стандартных диапазонах сердечного ритма показал, что все приборы не обеспечивают повторяемых результатов исследований.

Результаты решения задачи стандартизации методики расчета оценок спектральных показателей квазипериодических низкочастотных неэквилидистантно квантованных физиологических сигналов и исследования их точности на основе имитационного моделирования представлены в [72, 74, 75].

Изложенное обуславливает насущную актуальность разработки и реализации единых подходов к стандартизации методов обработки информации при разработке информационных технологий цифровой адаптационной медицины.

6. Заключение. Цифровая адаптационная медицина пока находится в начальной стадии развития, что обусловлено комплексом причин: от неготовности материально-технической базы организаций, приоритетом которых является решение задач адаптационной медицины, до отсутствия нормативно-правового регулирования решения таких задач.

Результаты анализа информационных технологий цифровой адаптационной медицины свидетельствуют о том, что их математическое и программно-аппаратное обеспечение нуждаются в совершенствовании. Приоритеты такого совершенствования связываются, прежде всего, с внедрением достижений четвёртой промышленной революции (Индустрии 4.0) и концепции социкиберфизических систем. Перспективные информационные технологии цифровой адаптационной медицины должны обеспечить решение задач адаптационной медицины с максимально полным применением для этого достижений научно-технического прогресса в интересах обеспечения:

увеличения объемов доступной для анализа информации о состоянии человека с возможностью ее получения без каких-либо помех

деятельности за счет реализации миниатюрных (в том числе, бесконтактных) датчиков и источников энергии, а также миниатюрных высокопроизводительных процессоров;

получения новых знаний о механизмах формирования адаптационных реакций человека на условия жизнедеятельности за счет применения цифровых двойников;

обеспечения сохранения конфиденциальности персональной информации о здоровье и защищенности информации от искажений.

За счет этого будет обеспечен прогресс в развитии потенциальных и адаптационных возможностей организма человека и увеличения продолжительности его активной жизни с помощью направленного пациент-ассоциированного воздействия на основе учета влияний генетических факторов и среды, что является приоритетом адаптационной медицины.

Литература

1. *Заболотная Н.В., Гатилова И.Н., Заболотный А.Т.* Цифровизация здравоохранения: достижения и перспективы развития // Экономика. Информатика. 2020. Т. 47. № 2. С. 380-389.
2. *Зарубина Т.В.* Актуальные вопросы внедрения информационных технологий в здравоохранении // Вестник Росздравнадзора. 2018. № 3. С. 20-25.
3. *Карпов О.Э., Субботин С.А., Шишкинов Д.В., Замятин М.Н.* Цифровое здравоохранение: необходимость и предпосылки // Врач и информационные технологии. 2017. № 3. С. 6-22.
4. *Гусев А.В., Плисс М.А., Левин М.Б., Новицкий Р.Э.* Тренды и прогнозы развития медицинских информационных систем в России // Врач и информационные технологии. 2019. № 2. С. 38-49.
5. *Mathews S.C., McShea M.J., Hanley C.L., Ravitz A., Labrique A.B., Cohen A.B.* Digital health: a path to validation // Digital Medicine. 2019. No. 2. P. 38. Doi: 10.1038/s41746-019-0111-3.
6. *Meldo A., Utkin L., Kovalev M., Kasimov E.* The natural language explanation algorithms for the lung cancer computer-aided diagnosis system // Artificial Intelligence in Medicine. 2020. Vol. 108. Pp. 101952. Doi: 10.1016/j.artmed.2020.101952
7. *Столяр В.П., Крайнюков П.Е., Калачёв О.В.* Цифровая трансформация здравоохранения и ведомственной медицины. М.: Планета, 2020. 200 с.
8. *Nazarenko G.I., Kleimenova E.B., Yashina L.P., Payushchik S.A., Konstantinova M.V., Mokin M.V., Otdelenov V.A., Molodchenkov A.I., Sychev D.A.* Development of the ontology of patient management technological records for modeling of clinical workflows in a general hospital // Scientific and Technical Information Processing. 2015. Vol. 42. No. 6. Pp. 455-462.
9. *Atreja A., Francis S., Kurra S., Kabra R.* Digital Medicine and Evolution of Remote Patient Monitoring in Cardiac Electrophysiology: A State-of-the-Art Perspective // Current Treatment Options in Cardiovascular Medicine. 2019. Vol. 21. No. 12. P. 92. Doi: 10.1007/s11936-019-0787-3.
10. *Györfly Z., Radó N., Mesko B.* Digitally engaged physicians about the digital health transition // PLoS One. 2020. Vol. 15. No. 9, e0238658. Doi: 10.1371/journal.pone.0238658.

11. *Kvedar J.C.* Evidence for the effectiveness of digital health // *Digital Medicine*. 2020. No. 3. P. 34. Doi: 10.1038/s41746-020-0231-9.
12. *Стефанова Н.А., Андропова И.В.* Проблемы цифровизации сферы здравоохранения: российский и зарубежный опыт // *Вестник Самарского университета. Экономика и управление*. 2018. Т. 9. № 3. С. 31-35.
13. *Mentsiev A.U., Yunaeva S.M.* Advantages and disadvantages of digital technology in healthcare system // *Medical Institute Bulletin*. 2019. Vol. 16. No. 2. Pp. 122-124.
14. *Orlov O.I., Perevedentsev O.V., Mamonova E.Y., Levanov V.M.* An integrated automated method for analyzing occupational health and medical provision under the extreme conditions of industrial activity // *Human Physiology*. 2018. Vol. 44. No. 7. Pp. 819-823.
15. *Engelhard M.M., Oliver J.A., McClernon F.J.* Digital envirotyping: quantifying environmental determinants of health and behavior // *Digital Medicine*. 2020. Vol. 12. No. 3. P. 36. Doi: 10.1038/s41746-020-0245-3.
16. *Labrique A., Agarwal S., Tamrat T., Mehl G.* WHO Digital Health Guidelines: a milestone for global health // *Digital Medicine*. 2020. No. 3. P. 120. Doi: 10.1038/s41746-020-00330-2.
17. *Смазулов С.М., Смазулова В.К.* Новейшие технологические тренды в медицине // *Инновации и инвестиции*. 2019. № 3. С. 289-290.
18. *Аллабергенов Р.Д., Махмуджанов С.У.* Интернет вещей в медицине: цифровые больницы // *Интернаука*. 2019. Т. 11. № 93. С. 6-7.
19. *Sandle T., Chesca A., Abdulina G.* Digital advances in modern pathology // *Bulletin of the Karaganda university. Biology. Medicine. Geography Series*. 2018. Vol. 90. No. 2. Pp. 86-94.
20. *Zhuravlev Yu.I., Nazarenko G.I., Vinogradov A.P., Dokukin A.A., Katerinochkina N.N., Kleimenova E.B., Konstantinova M.V., Ryazanov V.V., Sen'ko O.V., Cherkashov A.M.* Methods for discrete analysis of medical data on the basis of recognition theory and some of their applications // *Pattern Recognition and Image Analysis (Advances in Mathematical Theory and Applications)*. 2016. Vol. 26. No. 3. Pp. 643-664.
21. *Howarth A., Quesada J., Silva J., Judycki S., Mills P.R.* The impact of digital health interventions on health-related outcomes in the workplace: A systematic review // *Digit Health*. 2018. No. 4, 2055207618770861. Doi: 10.1177/2055207618770861.
22. *Kobrinskii B.A., Grigoriev O.G., Molodchenkov A.I., Smirnov I.V., Blagosklonov N.A.* Artificial intelligence technologies application for personal health management. IFAC-PapersOnLine // 19th IFAC Conference on Technology, Culture and International Stability, TECIS 2019. Pp. 70-74.
23. *Кобринский Б.А.* Системы искусственного интеллекта в медицинской практике: состояние и перспективы // *Вестник Росздравнадзора*. 2020. № 3. С. 37-43.
24. *Истомина Т.В.* Современное состояние и перспективы применения инфокоммуникационных технологий в российской медицине // *Медицинская техника*. 2021. № 1. С. 30-33.
25. *Гальцова О.М., Пронькин Н.Н.* Современные информационные технологии для здравоохранения будущего // *International Journal of Professional Science*. 2020. № 1. С. 39-44.
26. *Сертакова О.В.* Цифровые технологии Индустрии 4.0 в системе повышения качества медицинской помощи: телемедицина // *Экономика и социум: современные модели развития*. 2020. Т. 10. № 4. С. 367-380.
27. *Шепель Р.Н., Кутчер А.В., Ваховская Т.В., Дранкина О.М.* История развития телемедицины в Российской Федерации // *Неотложная кардиология и кардиоваскулярные риски*. 2019. Т. 3. № 2. С. 765-771.
28. *Уткин Л.В., Мелдо А.А., Ковалев М.С., Касимов Э.М.* Обзор методов объяснения и интерпретации принятия решений в интеллектуальных системах диагностики

- онкологических заболеваний // Искусственный интеллект и принятие решений. 2020. № 4. С. 55-65.
29. *Морозова Ю.А.* Цифровая трансформация российского здравоохранения как фактор развития отрасли // Интеллект. Инновации. Инвестиции. 2020. № 2. С. 36-47.
30. *Баевский Р.М.* Прогнозирование состояний на грани нормы и патологии. М.: Медицина, 1979. 298 с.
31. *Ushakov I.B., Chernikova A.G., Baevskiy R.M., Simakova T.G., Zipa O.M.* Prenosological approach to attaining the objectives of aviation medical expert assessment // *Human Physiology*. 2016. Vol. 42. No. 7. Pp. 724-730.
32. *Баранов В.М., Баевский Р.М., Берсенева А.П., Михайлов В.М.* Оценка адаптационных возможностей организма и задачи повышения эффективности здравоохранения // *Экология человека*. 2004. № 6. С. 25-29.
33. *Baevsky R.M., Baevsky A.R., Bersenev E.Y., Isaeva O.N., Pougachev V.I., Chernikova A.G.* Development of space technologies and problems of "home medicine" // *Cardiometry*. 2015. No. 6. Pp. 30-40.
34. *Fogel A.L., Kvedar J.C.* Artificial intelligence powers digital medicine // *Digital Medicine*. 2018. No. 1. P. 5. Doi: 10.1038/s41746-017-0012-2.
35. *Монахов Д.Н., Прончев Г.Б.* Политика цифровизации здравоохранения // Вопросы национальных и федеративных отношений. 2020. Т. 10. № 6 (63). С. 1582-1592.
36. *Суоаков К.В., Кукес В.Г., Хоманов К.З., Жестовская А.С.* Возможность восстановления функционального состояния адаптационных систем организма с применением методологии персонализированной медицины // *Лекарственные препараты и рациональная фармакотерапия*. 2013. № 2. С. 3-5.
37. *Труханов А.И., Скакун С.Г., Гречко А.В.* Современная роль персонифицированной цифровой медицины в развитии медицинской реабилитации // *Вестник восстановительной медицины*. 2018. № 1 (83). С. 2-13.
38. *Ушаков И.Б., Богомолов А.В.* Информатизация программ персонифицированной адаптационной медицины // *Вестник Российской академии медицинских наук*. 2014. Т. 69. № 5-6. С. 124-128.
39. *Пономаренко В.А., Разумов А.Н., Пискунов В.А.* Здоровье здорового человека. М.: Медицина, 1997. 205 с.
40. *Разумов А.Н., Пономаренко В.А.* Концепция «здоровье здорового человека»: интеграция медицины, психологии и религии // *Психологический журнал*. 2015. Т. 36. № 6. С. 88-93.
41. *Богомолов А.В., Гридин Л.А., Кукушкин Ю.А., Ушаков И.Б.* Диагностика состояния человека: математические подходы. М.: Медицина, 2003. 464 с.
42. *Ушаков И.Б.* Комбинированные воздействия в экологии человека и экстремальной медицине. М.: Издатцентр, 2013. 442 с.
43. *Прохоров Н.И., Донцов В.И., Крутько В.Н., Ходыкина Т.М.* Биологический возраст как метод оценки уровня здоровья при наличии экологических рисков (обзор литературы) // *Гигиена и санитария*. 2019. Т. 98. № 7. С. 761-765.
44. *Dontsov V.I., Krut'ko V.N.* Biological age as a method for systematic assessment of ontogenetic changes in the state of an organism // *Russian Journal of Developmental Biology*. 2015. Vol. 46. № 5. Pp. 246-253.
45. *Prilipko N.S., Turbinsky V.V., Bobrovniky I.P.* Hygienic evaluation of personalized health risk for prevention of environmentally diseased diseases in the primary health care system: overview // *Russian Journal of Rehabilitation Medicine*. 2020. No.3. Pp. 5-35.
46. *Крутько В.Н., Донцов В.И., Митрохин О.В., Матвеев А.А., Ермакова Н.А., Потемкина Н.С.* Искусственный интеллект для здоровьесбережения и развития личностного потенциала (обзор) // *Труды Института системного анализа Российской академии наук*. 2020. Т. 70. № 3. С. 86-100.

47. Павлова Л.Л., Барбаков О.М. Качество жизни человека в регионе под влиянием процессов цифровизации: программы реализации, оценка // Известия высших учебных заведений. Социология. Экономика. Политика. 2021. № 2. С. 86-101.
48. Хованов Н.В. Анализ и синтез показателей при информационном дефиците. СПб: СПбГУ, 1996. 196 с.
49. Саати Т.Л. Принятие решений при зависимостях и обратных связях: аналитические сети. М.: URSS, 2021. 360 с.
50. Иванова Г.Н., Окрепилов В.В., Окрепилова И.Г. Развитие теории управления качеством с целью устойчивого развития и повышения качества жизни // Качество и жизнь. 2016. № 1 (9). С. 3-9.
51. Rakhmanin Yu.A., Bobrovniksky I.P. Scientific and organizational-methodological bases of environmental medicine as the integrative direction of medical science and practical health care // Hygiene and sanitation. 2017. Vol. 96. No.10. Pp. 917-921.
52. Pereginya O.V., Lutsenko T.M. Translation medicine, biomedicine and medical biotechnology: the transition to personalized medicine // Biotechnologia Acta. 2020. Vol. 13. No. 2. Pp. 5-11.
53. Клеймёнова Е.Б., Назаренко Г.И., Яшина Л.П., Пающик С.А. Проблема трансляции знаний в здравоохранении: инструменты для ее решения в области безопасности пациентов // Вестник Российской академии медицинских наук. 2018. Т. 73. № 2. С. 105-114.
54. Генкин А.А., Медведев В.И. Прогнозирование психофизиологических состояний. Л.: Наука, 1973. 78 с.
55. Пятибрат А.О., Мельнов С.Б., Козлова А.С., Пятибрат Е.Д. Физиологическая оценка наследственной предрасположенности к экстремальным видам профессиональной деятельности // Медико-биологические проблемы жизнедеятельности. 2015. № 2 (14). С. 73-78.
56. Трифонова О.П., Балашова Е.Е., Маслов Д.Л., Григорьев А.И., Лисица А.В., Пономаренко Е.А., Арчаков А.И. Метаболомный анализ крови для создания цифрового образа здорового человека // Биомедицинская химия. 2020. Т. 66. № 3. С. 216-223.
57. Иванов И.В., Ушаков И.Б. Принципы экстраполяции экспериментальных данных с лабораторных животных на человека // Военно-медицинский журнал. 2019. № 12. С. 50-56.
58. Ушаков И.Б., Богомолов А.В., Драган С.П., Солдатов С.К. Методологические основы персонализированного гигиенического мониторинга // Авиакосмическая и экологическая медицина. 2017. Т. 51. № 6. С. 53-56.
59. Ронжин А.Л., Соколов Б.В., Джао В.Ю.Д., Миронова Е.Г., Стыскин М.М. Применение технологии радиочастотной идентификации для построения системы контроля оборота бортового оборудования // Вопросы радиоэлектроники. Серия: Техника телевидения. 2020. № 1. С. 13-20.
60. Ушаков И.Б., Богомолов А.В., Драган С.П., Солдатов С.К. Методологические основы персонализированного акустического мониторинга // Безопасность труда в промышленности. 2020. № 10. С. 33-39. Doi: 10.24000/0409-2961-2020-10-33-39
61. Kuleshov S.V., Zaytseva A.A., Ronzhin A.L. The development of soft defined distributed infocommunication systems architecture based on the active data technology // Lecture Notes in Networks and Systems. 2020. Vol. 95. Pp. 257-265.
62. Алёхин М.Д., Богомолов А.В., Кукушкин Ю.А. Методики анализа паттернов дыхания при бесконтактном мониторинге психофизиологических состояний операторов эргатических систем // Авиакосмическая и экологическая медицина. 2019. Т. 53. № 2. С. 99-101.
63. Максимов И.Б., Столяр В.П., Богомолов А.В. Прикладная теория информационного обеспечения медико-биологических исследований. М.: Бином, 2013. 312 с.

64. Ушаков И.Б., Богомолов А.В. Диагностика функциональных состояний человека в приоритетных исследованиях отечественных физиологических школ // Медико-биологические и социально-психологические проблемы безопасности в чрезвычайных ситуациях. 2021. № 3. С. 91-100. Doi: 10.25016/2541-7487-2021-0-3-91-100

65. Luzhkov P.V., Pika T.O., Shamaev D.M. Developing the structure of a hardware and software system for quantitative diagnosis of microhemodynamics // International Journal of Biomedicine. 2015. Vol. 5. No. 4. Pp. 228-230.

66. Ильин Е.П. Психофизиология состояний человека. СПб: Питер, 2005. 412 с.

67. Ушаков И.Б., Богомолов А.В., Кукушкин Ю.А. Психофизиологические механизмы формирования и развития функциональных состояний // Российский физиологический журнал им. И.М. Сеченова. 2014. Т. 100. № 10. С. 1130-1137.

68. Sandercock G., Shelton C., Bromley P., Brodie D. Agreement between three commercially available instruments for measuring short-term heart rate variability // Physiological Measurement. 2004. Vol. 25. No. 5. Pp. 1115-1124.

69. Kukushkin Y.A., Bogomolov A.V., Maistrov A.I. Rhythmocardiogram approximation methods for calculation of spectral parameters of cardiac rhythm variability // Biomedical Engineering. 2010. Vol. 44. No. 3. Pp. 92-103.

70. Nunan D., Jakovljevic D., Donovan G. Levels of agreement for RR-intervals and short-term heart rate variability obtained from the Polar S810 and an alternative system // European Journal of Applied Physiology. 2008. Vol. 103. No. 5. Pp. 529-537.

71. Lytaev S. Modeling and estimation of physiological, psychological and sensory indicators for working capacity // Advances in Intelligent Systems and Computing. 2021. Vol. 1201 AISC. Pp. 207-213. Doi: 10.1007/978-3-030-51041-1_28

72. Bogomolov A.V., Maistrov A.I. Theoretical-experimental analysis of convergence of heart rate variability spectral measures estimated via heart rate and heart period signals // Biomedical Engineering. 2009. Vol. 43. No. 2. Pp. 75-80.

73. Воробьев К.П., Паламарчук Е.А. Результаты независимого тестирования трех программ вычисления показателей variability сердечного ритма // Украинский медицинский часопис. 2007. Т. 3. № 59. С. 45-51.

74. Майстров А.И., Богомолов А.В., Алехин М.Д., Зарецкий А.П. Математическое моделирование ритмокардиографических сигналов для стандартизации методов их спектрального анализа // Труды Московского физико-технического института (национально-исследовательского университета). 2015. Т. 7. № 3 (27). С. 116-130.

75. Maistrov A.I., Alekhin M.D., Bogomolov A.V., Zaretskiy A.P. Mathematical modeling of rhythmocardiographic signals spectrum // Proceedings - The Second International Conference on Engineering and Telecommunication. En&T-2015. 2015. Pp. 85-86. Doi: 10.1109/EnT.2015.27

Богомолов Алексей Валерьевич — д-р техн. наук, профессор, главный научный сотрудник, лаборатория технологий больших данных социкиберфизических систем, СПб ФИЦ РАН. Область научных интересов: информационные технологии, компьютерные технологии, медицинская информатика, биофизика, применение методов математики и информатики в медико-биологических исследованиях. Число научных публикаций — 200. a.v.bogomolov@gmail.com; 14-я линия В.О., д. 39, г. Санкт-Петербург, 199178, РФ; р.т. +7(812)328-3311, факс +7(812)328-4450.

Поддержка исследований. Работа выполнена при финансовой поддержке РФФИ (проект № 20-17-50056).

A. BOGOMOLOV
**INFORMATION TECHNOLOGIES OF DIGITAL ADAPTIVE
MEDICINE**

Bogomolov A. Information technologies of digital adaptive medicine.

Abstract. The article provides a comprehensive description of information technologies of digital adaptive medicine. The emphasis is on the applicability to the development of specialized automated complexes, software models and systems for studying the adaptive capabilities of a person to environmental conditions. Requirements for information technologies to enhance these capabilities are formulated. The features of information technologies are reflected in relation to the implementation of applied systemic studies of life support, preservation of professional health and prolongation of human longevity.

Six basic concepts of adaptive medicine with an emphasis on the features of the mathematical support for information processing are characterized, priorities for improving information technologies used in these concepts are determined.

The information technologies used in the tasks of ensuring the professional performance of a person with an emphasis on the need to use adequate methods for diagnosing the state of a person at all stages of professional activity and the need to develop technologies for digital twins that adequately simulate the adaptation processes and reactions of the body in real conditions are considered.

The characteristics of information technologies for personalized monitoring of health risks are given, which make it possible to objectify the effects of physical factors of the conditions of activity and to implement individual and collective informing of personnel about environmental hazards.

The urgent need to standardize information processing methods in the development of information technologies for digital adaptive medicine in the interests of ensuring physiological adequacy and mathematical correctness of approaches to obtaining and processing information about a person's state is shown.

It is concluded that the priorities for improving information technologies of digital adaptive medicine are associated with the implementation of the achievements of the fourth industrial revolution, including the concept of sociocyberphysical systems.

Keywords: digital medicine, digital health, medical informatics, adaptive medicine, preventive medicine, sociocyberphysical systems.

Bogomolov Alexey — Ph.D., Dr. Sci., Professor, Chief Researcher, Laboratory of Big Data Technologies of Sociocyberphysical Systems, SPC RAS. Research interests: information technology, computer technology, medical informatics, biophysics, application of methods of mathematics and informatics in biomedical research. The number of publications — 200. a.v.bogomolov@gmail.com; 39, 14-th Line V.O., St. Petersburg, 199178, Russia; office phone +7(812)328-3311, fax +7(812)328-4450.

Acknowledgements. This research is supported by RFBR (grant 20-17-50056).

References

1. Zabolotnaya N.V., Gatilova I.N., Zabolotnyy A.T. [Digitalization of healthcare: achievements and development prospects]. *Ekonomika. Informatika*. [Economy. Computer science]. 2020, vol. 47, no. 2, pp. 380-389. (In Russ.).

2. Zarubina T.V. [Topical issues of the implementation of information technologies in health care]. *Vestnik Roszdravnadzora*. [Roszdravnadzor Bulletin]. 2018, no. 3, pp. 20-25. (In Russ.).
3. Karpov O.E., Subbotin S.A., Shishkanov D.V., Zamyatin M.N. [Digital health care: necessity and prerequisites]. *Vrach i informatsionnyye tekhnologii*. [Physician and information technology]. 2017, no. 3, p. 6-22. (In Russ.).
4. Gusev A.V., Pliss M.A., Levin M.B., Novitsky R.E. [Trends and forecasts for the development of medical information systems in Russia]. *Vrach i informatsionnyye tekhnologii*. [Physician and information technology]. 2019, no. 2, pp. 38-49. (In Russ.).
5. Mathews S.C., McShea M.J., Hanley C.L., Ravitz A., Labrique A.B., Cohen A.B. Digital health: a path to validation. *Digital Medicine*. 2019, no. 2, p. 38. doi: 10.1038/s41746-019-0111-3.
6. Meldo A., Utkin L., Kovalev M., Kasimov E. The natural language explanation algorithms for the lung cancer computer-aided diagnosis system. *Artificial Intelligence in Medicine*. 2020, vol. 108, pp. 101952. Doi: 10.1016/j.artmed.2020.101952
7. Stolyar V.P., Kraynyukov P.Ye., Kalachov O.V. *Tsifrovaya transformatsiya zdra-vookhraneniya i vedomstvennoy meditsiny*. [Digital transformation of healthcare and departmental medicine]. Moscow: Planeta, 2020. 200 p. (In Russ.).
8. Nazarenko G.I., Kleimenova E.B., Yashina L.P., Payushchik S.A., Konstantinova M.V., Mokin M.V., Otdelenov V.A., Molodchenkov A.I., Sychev D.A. Development of the ontology of patient management technological records for modeling of clinical workflows in a general hospital. *Scientific and Technical Information Processing*. 2015, vol. 42, no. 6, pp. 455-462.
9. Atreja A., Francis S., Kurra S., Kabra R. Digital Medicine and Evolution of Remote Patient Monitoring in Cardiac Electrophysiology: A State-of-the-Art Perspective. *Current Treatment Options in Cardiovascular Medicine*. 2019, vol. 21, no. 12, p. 92. Doi: 10.1007/s11936-019-0787-3.
10. Györfly Z., Radó N., Mesko B. Digitally engaged physicians about the digital health transition. *PLoS One*. 2020, vol. 15, no. 9, e0238658. Doi: 10.1371/journal.pone.0238658.
11. Kvedar J.C. Evidence for the effectiveness of digital health. *Digital Medicine*. 2020, no. 3, p. 34. Doi: 10.1038/s41746-020-0231-9.
12. Stefanova N.A., Andronova I.V. Problems of digitalization of the healthcare sector: Russian and foreign experience. *Vestnik Samarskogo universiteta. Ekonomika i upravleniye*. [Samara University Bulletin. Economics and Management]. 2018, vol. 9, no. 3, pp. 31-35. (In Russ.).
13. Mentsiev A.U., Yunaeva S.M. Advantages and disadvantages of digital technology in healthcare system. *Medical Institute Bulletin*. 2019, vol. 16, no. 2, pp. 122-124.
14. Orlov O.I., Perevedentsev O.V., Mamonova E.Y., Levanov V.M. An integrated automated method for analyzing occupational health and medical provision under the extreme conditions of industrial activity. *Human Physiology*. 2018, vol. 44, no. 7, pp. 819-823.
15. Engelhard M.M., Oliver J.A., McClernon F.J. Digital envirotyping: quantifying environmental determinants of health and behavior. *Digital Medicine*. 2020, vol. 12, no. 3, p. 36. Doi: 10.1038/s41746-020-0245-3.
16. Labrique A., Agarwal S., Tamrat T., Mehl G. WHO Digital Health Guidelines: a milestone for global health. *Digital Medicine*. 2020, no. 3, p. 120. Doi: 10.1038/s41746-020-00330-2.
17. Smagulov S.M., Smagulova V.K. [The latest technological trends in medicine]. *Innovatsii i investitsii*. [Innovation and investment]. 2019, no. 3, pp. 289-290.
18. Allabergenov R.D., Makhmudzhanov S.U. [The Internet of Things in Medicine: Digital Hospitals]. *Internauka*. [Internauka]. 2019, vol. 11, no. 93, pp. 6-7.

19. Sandle T., Chesca A., Abdulina G. Digital advances in modern pathology. Bulletin of the Karaganda university. Biology. Medicine. Geography Series. 2018, vol. 90, no. 2, pp. 86-94.
20. Zhuravlev Yu.I., Nazarenko G.I., Vinogradov A.P., Dokukin A.A., Katerinochkina N.N., Kleimenova E.B., Konstantinova M.V., Ryazanov V.V., Sen'ko O.V., Cherkashov A.M. Methods for discrete analysis of medical data on the basis of recognition theory and some of their applications. Pattern Recognition and Image Analysis (Advances in Mathematical Theory and Applications). 2016, vol. 26, no. 3, pp. 643-664.
21. Howarth A., Quesada J., Silva J., Judycki S., Mills P.R. The impact of digital health interventions on health-related outcomes in the workplace: A systematic review. Digit Health. 2018, no. 4, 2055207618770861. Doi: 10.1177/2055207618770861.
22. Kobrinskii B.A., Grigoriev O.G., Molodchenkov A.I., Smirnov I.V., Blagosklonov N.A. Artificial intelligence technologies application for personal health management. IFAC-PapersOnLine. 19th IFAC Conference on Technology, Culture and International Stability, TECIS 2019, pp. 70-74.
23. Kobrinskii B.A. [Artificial intelligence systems in medical practice: state and prospects]. *Vestnik Roszdravnadzora*. [Roszdravnadzor Bulletin]. 2020, no. 3, pp. 37-43. (In Russ.).
24. Istomina T.V. [Current state and prospects of using infocommunication technologies in Russian medicine]. *Meditsinskaya tekhnika*. [Medical equipment]. 2021, no. 1, pp. 30-33. (In Russ.).
25. Galtsova O.M., Pronkin N.N. [Modern information technologies for the healthcare of the future]. *International Journal of Professional Science*. 2020, no. 1, pp. 39-44. (In Russ.).
26. Sertakova O.V. [Digital technologies of Industry 4.0 in the system of improving the quality of medical care: telemedicine]. *Ekonomika i sotsium: sovremennyye modeli razvitiya*. [Economy and society: modern development models]. 2020, vol. 10, no. 4, pp. 367-380. (In Russ.).
27. Shepel R.N., Kutcher A.V., Vakhovskaya T.V., Drapkina O.M. [The history of the development of telemedicine in the Russian Federation]. *Neotlozhnaya kardiologiya i kardiovaskulyarnyye riski*. [Emergency cardiology and cardiovascular risks]. 2019, vol. 3, no. 2, pp. 765-771. (In Russ.).
28. Utkin L.V., Meldo A.A., Kovalev M.S., Kasimov E.M. [Review of methods for explaining and interpreting decision-making in intelligent systems for diagnosing oncological diseases]. *Iskusstvennyy intellekt i prinyatiye resheniy*. [Artificial intelligence and decision making]. 2020, no. 4, pp. 55-65. (In Russ.).
29. Morozova Yu.A. [Digital transformation of Russian healthcare as a factor in the development of the industry]. *Intellekt. Innovatsii. Investitsii*. [Intelligence. Innovation. Investments]. 2020, no. 2, pp. 36-47. (In Russ.).
30. Baevsky R.M. *Prognozirovaniye sostoyaniy na grani normy i patologii*. [Prediction of conditions on the verge of norm and pathology]. Moscow: Medicine, 1979. 298 p. (In Russ.).
31. Ushakov I.B., Chernikova A.G., Baevskiy R.M., Simakova T.G., Zipa O.M. Prenosological approach to attaining the objectives of aviation medical expert assessment. *Human Physiology*. 2016, vol. 42, no. 7, pp. 724-730.
32. Baranov V.M., Baevsky R.M., Berseneva A.P., Mikhailov V.M. [Assessment of the body's adaptive capabilities and the tasks of improving the efficiency of health care]. *Ekologiya cheloveka*. [Human ecology]. 2004, no. 6, pp. 25-29. (In Russ.).
33. Baevsky R.M., Baevsky A.R., Bersenev E.Y., Isaeva O.N., Pougachev V.I., Chernikova A.G. Development of space technologies and problems of "home medicine". *Cardiometry*. 2015, no. 6, pp. 30-40.
34. Fogel A.L., Kvedar J.C. Artificial intelligence powers digital medicine. *Digital Medicine*. 2018, no. 1, p. 5. Doi: 10.1038/s41746-017-0012-2.

35. Monakhov D.N., Pronchev G.B. [Healthcare digitalization policy]. *Voprosy natsional'nykh i federativnykh otnosheniy*. [Questions of national and federal relations]. 2020, vol. 10, no. 6 (63), pp. 1582-1592. (In Russ.).
36. Sudakov K.V., Kukes V.G., Khomanov K.Z., Zhestovskaya A.S. [The possibility of restoring the functional state of the body's adaptive systems using the methodology of personalized medicine]. *Lekarstvennyye preparaty i ratsional'naya farmakoterapiya*. [Medicines and rational pharmacotherapy]. 2013, No. 2, pp. 3-5. (In Russ.).
37. Trukhanov A.I., Skakun S.G., Grechko A.V. [The modern role of personalized digital medicine in the development of medical rehabilitation]. *Vestnik vosstanovitel'noy meditsiny*. [Herald of restorative medicine]. 2018, no. 1 (83), pp. 2-13. (In Russ.).
38. Ushakov I.B., Bogomolov A.V. [Computerization of programs of personalized adaptive medicine]. *Vestnik Rossiyskoy akademii meditsinskikh nauk*. [Bulletin of the Russian Academy of Medical Sciences]. 2014, vol. 69, no. 5-6, pp. 124-128. (In Russ.).
39. Ponomarenko V.A., Razumov A.N., Piskunov V.A. *Zdorov'ye zdorovogo cheloveka*. [Health of a healthy person]. Moscow: Medicine, 1997.205 p. (In Russ.).
40. Razumov A.N., Ponomarenko V.A. [The concept of "health of a healthy person": the integration of medicine, psychology and religion]. *Psikhologicheskii zhurnal*. [Psychological journal]. 2015, vol. 36, no. 6, pp. 88-93. (In Russ.).
41. Bogomolov A.V., Gridin L.A., Kukushkin Yu.A., Ushakov I.B. *Diagnostika sostoyaniya cheloveka: matematicheskiye podkhody*. [Diagnostics of the human condition: mathematical approaches]. Moscow: Medicine, 2003. 464 p. (In Russ.).
42. Ushakov I.B. *Kombinirovannyye vozdeystviya v ekologii cheloveka i ekstremal'noy meditsine*. [Combined effects in human ecology and extreme medicine]. Moscow: Izdatcentr, 2013. 442 p. (In Russ.).
43. Prokhorov N.I., Dontsov V.I., Krutko V.N., Khodykina T.M. [Biological age as a method for assessing the level of health in the presence of environmental risks (literature review)]. *Gigiyena i sanitariya*. [Hygiene and sanitation]. 2019, vol. 98, no. 7, p. 761-765. (In Russ.).
44. *Dontsov V.I., Krut'ko V.N.* Biological age as a method for systematic assessment of ontogenetic changes in the state of an organism. *Russian Journal of Developmental Biology*. 2015, vol. 46, № 5, pp. 246-253.
45. *Prilipko N.S., Turbinsky V.V., Bobrovniy I.P.* Hygienic evaluation of personalized health risk for prevention of environmentally diseased diseases in the primary health care system: overview. *Russian Journal of Rehabilitation Medicine*. 2020, no.3, pp. 5-35.
46. Krut'ko V.N., Dontsov V.I., Mitrokhin O.V., Matveev A.A., Ermakova N.A., Potemkina N.S. [Artificial Intelligence for Health Preservation and Development of Personal Potential (Review)]. *Trudy Instituta sistemnogo analiza Rossiyskoy akademii nauk*. [Proceedings of the Institute for Systems Analysis of the Russian Academy of Sciences]. 2020, vol. 70, no. 3, pp. 86-100. (In Russ.).
47. Pavlova L.L., Barbakov O.M. [The quality of human life in the region under the influence of digitalization processes: implementation programs, assessment]. *Izvestiya vysshikh uchebnykh zavedeniy. Sotsiologiya. Ekonomika. Politika*. [Proceedings of higher educational institutions. Sociology. Economy. Politics]. 2021, no. 2, pp. 86-101. (In Russ.).
48. Khovanov N.V. *Analiz i sintez pokazateley pri informatsionnom defitsite*. [Analysis and synthesis of indicators for information deficit]. SPb: SPbGU, 1996. 196 p. (In Russ.).
49. Saati T.L. *Prinyatiye resheniy pri zavisimostyakh i obratnykh svyazyakh: analiticheskiye seti*. [Dependency and Feedback Decision Making: Analytical Networks]. Moscow: URSS, 2021. 360 p. (In Russ.).
50. Ivanova G.N., Okrepilov V.V., Okrepilova I.G. [Development of the theory of quality management with the aim of sustainable development and improving the quality of life]. *Kachestvo i zhizn'*. [Quality and life]. 2016, no. 1 (9), pp. 3-9. (In Russ.).

51. Rakhmanin Yu.A., Bobrovniksky I.P. Scientific and organizational-methodological bases of environmental medicine as the integrative direction of medical science and practical health care. Hygiene and sanitation. 2017, vol. 96, no.10, pp. 917-921.
52. Pereginya O.V., Lutsenko T.M. Translation medicine, biomedicine and medical biotechnology: the transition to personalized medicine. Biotechnologia Acta. 2020, vol. 13, no. 2, pp. 5-11.
53. Kleymonova Ye.B., Nazarenko G.I., Yashina L.P., Payushchik S.A. [The problem of knowledge translation in healthcare: tools for solving it in the field of patient safety]. *Vestnik Rossiyskoy akademii meditsinskikh nauk*. [Bulletin of the Russian Academy of Medical Sciences]. 2018, vol. 73, no. 2, pp. 105-114. (In Russ.).
54. Genkin A.A., Medvedev V.I. *Prognozirovaniye psikhofiziologicheskikh sostoyaniy*. [Prediction of psychophysiological states]. Leningrad: Nauka, 1973. 78 p. (In Russ.).
55. Pyatibrat A.O., Melnov S.B., Kozlova A.S., Pyatibrat E.D. [Physiological assessment of hereditary predisposition to extreme types of professional activity]. *Mediko-biologicheskiye problemy zhiznedeyatel'nosti*. [Biomedical problems of life]. 2015, no. 2 (14), pp. 73-78. (In Russ.).
56. Trifonova O.P., Balashova E.E., Maslov D.L., Grigoriev A.I., Lisitsa A.V., Ponomarenko E.A., Archakov A.I. [Metabolic blood test to create a digital image of a healthy person]. *Biomeditsinskaya khimiya*. [Biomedical Chemistry]. 2020, vol. 66, no. 3, pp. 216-223. (In Russ.).
57. Ivanov I.V., Ushakov I.B. [Principles of extrapolation of experimental data from laboratory animals to humans]. *Voyenno-meditsinskiy zhurnal*. [Military Medical Journal]. 2019, no. 12, pp. 50-56. (In Russ.).
58. Ushakov I.B., Bogomolov A.V., Dragan S.P., Soldatov S.K. [Methodological foundations of personalized hygienic monitoring]. *Aviakosmicheskaya i ekologicheskaya meditsina*. [Aviation and Space and Environmental Medicine]. 2017, vol. 51, no. 6, pp. 53-56. (In Russ.).
59. Ronzhin A.L., Sokolov B.V., Dzhaio V.Yu.D., Mironova E.G., Styskin M.M. [Application of radio frequency identification technology to build a control system for the turnover of onboard equipment]. *Voprosy radioelektroniki. Seriya: Tekhnika teledeniya*. [Radio electronics issues. Series: Television technology]. 2020, no. 1, pp. 13-20. (In Russ.).
60. Ushakov I.B., Bogomolov A.V., Dragan S.P., Soldatov S.K. [Methodological foundations of personalized acoustic monitoring]. *Bezopasnost' truda v promyshlennosti*. [Labor safety in industry]. 2020, no. 10, pp. 33-39. Doi: 10.24000 / 0409-2961-2020-10-33-39 (In Russ.).
61. Kuleshov S.V., Zaytseva A.A., Ronzhin A.L. The development of soft defined distributed infocommunication systems architecture based on the active data technology. Lecture Notes in Networks and Systems. 2020, vol. 95, pp. 257-265.
62. Alekhin M.D., Bogomolov A.V., Kukushkin Yu.A. [Methods for analyzing breathing patterns during non-contact monitoring of psychophysiological states of operators of ergatic systems]. *Aviakosmicheskaya i ekologicheskaya meditsina*. [Aerospace and Environmental Medicine]. 2019, vol. 53, no. 2, pp. 99-101. (In Russ.).
63. Maksimov I.B., Stolyar V.P., Bogomolov A.V. *Prikladnaya teoriya informatsionnogo obespecheniya mediko-biologicheskikh issledovaniy*. [Applied theory of information support for biomedical research]. Moscow: Binom, 2013. 312 p. (In Russ.).
64. Ushakov I.B., Bogomolov A.V. [Diagnostics of human functional states in priority studies of domestic physiological schools]. *Mediko-biologicheskiye i sotsial'no-psikhologicheskiye problemy bezopasnosti v chrezvychaynykh situatsiyakh*. [Medico-biological and socio-psychological problems of safety in emergency situations]. 2021, no. 3, pp. 91-100. Doi: 10.25016 / 2541-7487-2021-0-3-91-100 (In Russ.).

65. Luzhnov P.V., Pika T.O., Shamaev D.M. Developing the structure of a hardware and software system for quantitative diagnosis of microhemodynamics. *International Journal of Biomedicine*. 2015, vol. 5, no. 4, pp. 228-230.
66. Ilyin E.P. *Psikhofiziologiya sostoyaniy cheloveka*. [Psychophysiology of human states]. Saint Petersburg: Piter, 2005. 412 p. (In Russ.).
67. Ushakov I.B., Bogomolov A.V., Kukushkin Yu.A. [Psychophysiological mechanisms of formation and development of functional states]. *Rossiyskiy fiziologicheskiy zhurnal im. I.M. Sechenova*. [Russian physiological journal named after I.M. Sechenov]. 2014, vol. 100, no. 10, pp. 1130-1137. (In Russ.).
68. Sandercock G., Shelton C., Bromley P., Brodie D. Agreement between three commercially available instruments for measuring short-term heart rate variability. *Physiological Measurement*. 2004, vol. 25, no. 5, pp. 1115-1124.
69. Kukushkin Y.A., Bogomolov A.V., Maistrov A.I. Rhythmocardiogram approximation methods for calculation of spectral parameters of cardiac rhythm variability. *Biomedical Engineering*. 2010, vol. 44, no. 3, pp. 92-103.
70. Nunan D., Jakovljevic D., Donovan G. Levels of agreement for RR-intervals and short-term heart rate variability obtained from the Polar S810 and an alternative system. *European Journal of Applied Physiology*. 2008, vol. 103, no. 5, pp. 529-537.
71. Lytaev S. Modeling and estimation of physiological, psychological and sensory indicators for working capacity. *Advances in Intelligent Systems and Computing*. 2021, vol. 1201 AISC, pp. 207-213. Doi: 10.1007/978-3-030-51041-1_28
72. Bogomolov A.V., Maistrov A.I. Theoretical-experimental analysis of convergence of heart rate variability spectral measures estimated via heart rate and heart period signals. *Biomedical Engineering*. 2009, vol. 43, no. 2, pp. 75-80.
73. Vorobiev K.P., Palamarchuk E.A. [Results of independent testing of three programs for calculating heart rate variability indices]. *Ukrains'kiy medichniy chasopis*. [Ukrainian medical clock]. 2007, vol. 3, no. 59, pp. 45-51. (In Russ.).
74. Maistrov A.I., Bogomolov A.V., Alekhin M.D., Zaretsky A.P. [Mathematical modeling of rhythmocardiographic signals to standardize the methods of their spectral analysis]. *Trudy Moskovskogo fiziko-tehnicheskogo instituta (natsional'nogo issledovatel'skogo universiteta)*. [Proceedings of the Moscow Institute of Physics and Technology (National Research University)]. 2015, vol. 7, no. 3 (27), pp. 116-130. (In Russ.).
75. Maistrov A.I., Alekhin M.D., Bogomolov A.V., Zaretskiy A.P. Mathematical modeling of rhythmocardiographic signals spectrum. *Proceedings - The Second International Conference on Engineering and Telecommunication. En&T-2015*. 2015, pp. 85-86. Doi: 10.1109/EnT.2015.27

Руководство для авторов

Взаимодействие автора с редакцией осуществляется через личный кабинет на сайте журнала «Информатика и автоматизация» <http://ia.spcras.ru/>. При регистрации авторам рекомендуется заполнить все предложенные поля данных. Подготовка статьи ведется с помощью текстовых редакторов MS Word 2007 и выше или LaTeX. Объем основного текста (до раздела Литература) - от 20 до 30 страниц включительно. Переносы разрешены. Номера страниц не проставляются. Основная часть текста статьи разбивается на разделы, среди которых являются обязательными: введение, хотя бы один «содержательный» раздел и заключение. Допускается также мотивированное содержанием и структурой материал а выделение подразделов. В основную часть опускается помещать рисунки, таблицы, листинги и формулы. Правила их оформления подробно рассмотрены на нашем сайте в разделе «Руководство для авторов».

Author guidelines

Interaction between each potential author and the Editorial board is realized through the pesoal account on the website of the journal “Informatics and Automation” <http://ia.spcras.ru/>. At the registration the authors are requested to fill out all data fields in the proposed form. The submissions should be prepared using MS Word 2007, LaTeX. The text of the paper in the main part should not exceed 30 pages. Pages are not numbered; hyphenations are allowed. Certain figures, tables, listings and formulas are allowed in the main section, and their typography is considered in more detail at the journal web.

Signed to print 01.10.2021

Printed in Publishing center GUAP, 67, B. Morskaya, St. Petersburg, 190000, Russia

The journal is registered in the Russian Federal Agency for Communications and Mass-Media Supervision, certificate ПИ № ФС77-79228 dated September 25, 2020
Subscription Index П5513, Russian Post Catalog

Подписано к печати 01.10.2021. Формат 60×90 1/16. Усл. печ. л. 10,52 Заказ № 421.

Тираж 300 экз., цена свободная.

Отпечатано в Редакционно-издательском центре ГУАП, 190000, Санкт-Петербург, Б. Морская, д. 67

Журнал зарегистрирован Федеральной службой по надзору в сфере связи и массовых коммуникаций, свидетельство ПИ № ФС77-79228 от 25 сентября 2020 г.

Подписной индекс П5513 по каталогу «Почта России»