

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 заболеваний 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 вируса 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 статистическим 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.

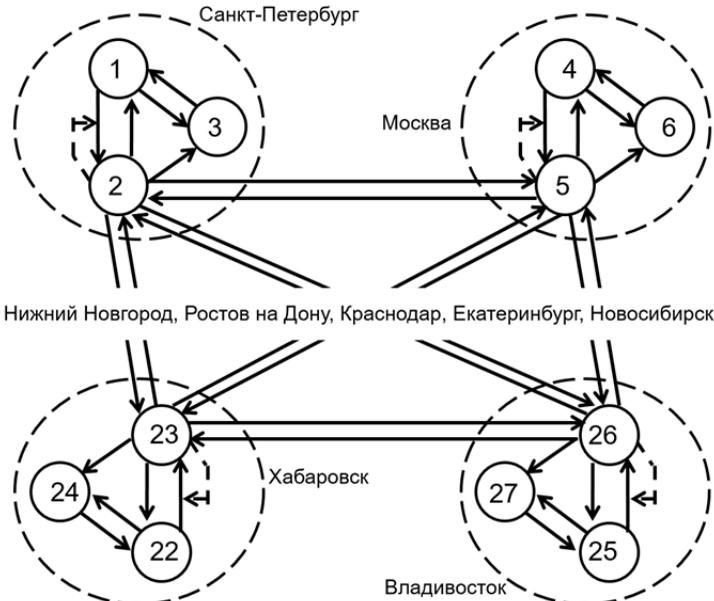


Рис. 1. Граф состояний процесса развития эпидемии в девяти городах

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

$$\begin{aligned} \frac{dP_1(t)}{dt} &= \lambda_{31}P_3(t) + \lambda_{21}P_2(t) - \lambda_{43}P_1(t) - \lambda_{42}P_2(t) \cdot P_1(t), \\ \frac{dP_2(t)}{dt} &= \lambda_{12}P_2(t) \cdot P_1(t) - \lambda_{21}P_2(t) - \lambda_{23}P_2(t) + \lambda_{52}P_5(t) + \lambda_{82}P_8(t) + \lambda_{41,2}P_{11}(t) + \lambda_{44,2}P_{14}(t) + \\ &+ \lambda_{17,2}P_{17}(t) + \lambda_{20,2}P_{20}(t) + \lambda_{23,2}P_{23}(t) + \lambda_{26,2}P_{26}(t) - \\ &- (\lambda_{25} + \lambda_{28} + \lambda_{2,11} + \lambda_{2,14} + \lambda_{2,17} + \lambda_{2,20} + \lambda_{2,23} + \lambda_{2,26}) \cdot P_2(t), \\ \frac{dP_3(t)}{dt} &= \lambda_{13}P_1(t) + \lambda_{23}P_2(t) - \lambda_{31}P_3(t), \\ &..... \\ \frac{dP_{25}(t)}{dt} &= \lambda_{27,25}P_{27}(t) + \lambda_{26,25}P_{26}(t) - \lambda_{25,27}P_{25}(t) - \lambda_{25,26}P_{26}(t) \cdot P_{25}(t), \\ \frac{dP_{26}(t)}{dt} &= \lambda_{25,26}P_{26}(t) \cdot P_{25}(t) - \lambda_{26,25}P_{26}(t) - \lambda_{26,27}P_{26}(t) + \lambda_{2,26}P_2(t) + \lambda_{5,26}P_5(t) + \lambda_{8,26}P_8(t) + \\ &+ \lambda_{11,26}P_{11}(t) + \lambda_{14,26}P_{14}(t) + \lambda_{17,26}P_{17}(t) + \lambda_{20,26}P_{20}(t) + \lambda_{23,26}P_{23}(t) - \\ &- (\lambda_{26,2} + \lambda_{26,5} + \lambda_{26,8} + \lambda_{26,11} + \lambda_{26,14} + \lambda_{26,17} + \lambda_{26,20} + \lambda_{26,23}) \cdot P_{26}(t), \\ \frac{dP_{27}(t)}{dt} &= \lambda_{25,27}P_{25}(t) + \lambda_{26,27}P_{26}(t) - \lambda_{27,25}P_{27}(t). \end{aligned}$$

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 $P_2(t), P_5(t), \dots, P_{23}(t), P_{26}(t)$ 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

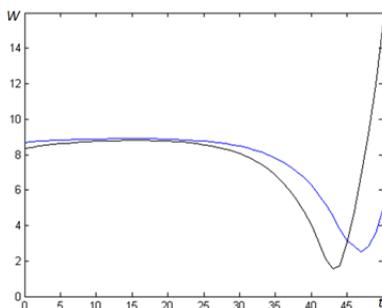
Город	Относительно данных на 30 мая 2020 г. минимум ошибки / момент времени (в неделях)	Относительно данных на 30 июня 2020 г. минимум ошибки / момент времени (в неделях)	Суммарные ошибки	Порядок инфицирования
Санкт-Петербург	1,5804 / 37	2,5819 / 41	4,1623	3
Москва	1,6190 / 28	2,6276 / 32	4,2466	7
Нижний 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

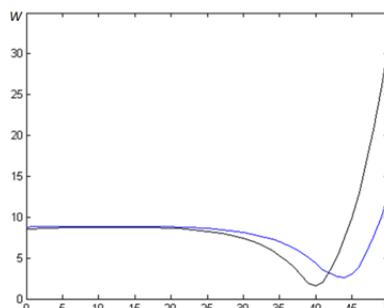
Данная таблица построена при сопоставлении модельных результатов и фактических данных по инфицированию людей в городах на 30 мая и 30 июня 2020 года. Анализ таблицы показывает, что относительно данных на 30 мая 2020 г. на первое место по порядку начала эпидемии выходит Новосибирск, затем Нижний Новгород 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.



а)



б)

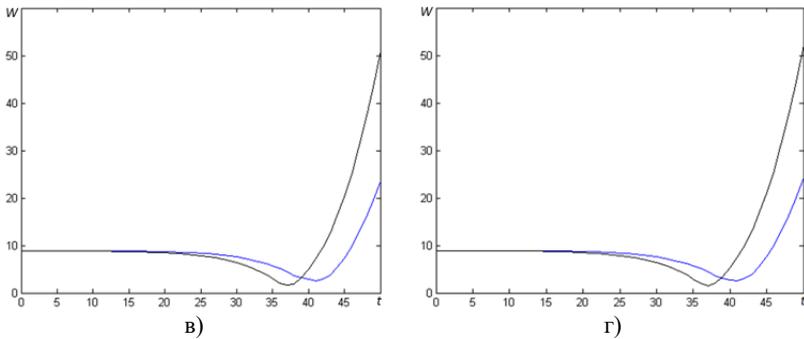
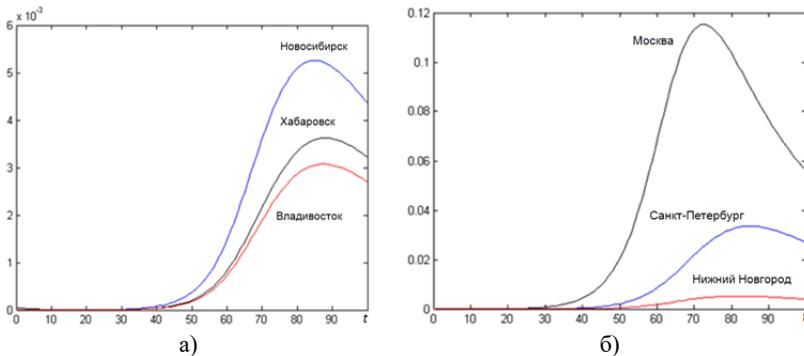


Рис. 2. Зависимости от времени (в неделях) суммарных относительных ошибок совпадения модельных результатов с фактическими данными относительно 30 мая и 30 июня 2020 года: а) при начале эпидемии в Хабаровске; б) в Новосибирске; в) Санкт-Петербурге; г) в Нижнем Новгороде

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.



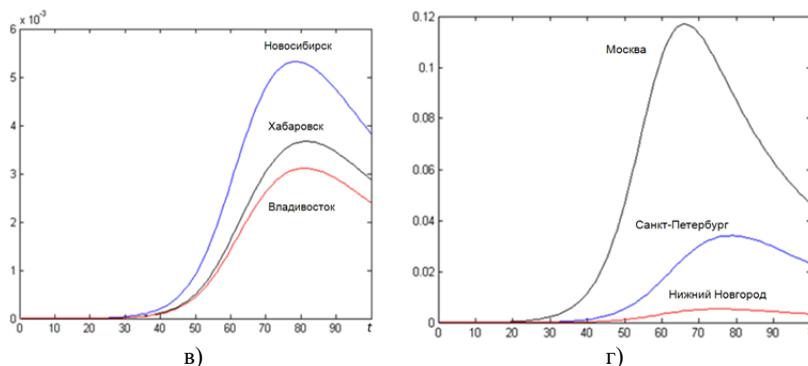


Рис. 3. Зависимости от времени (в неделях) вероятностей инфицированности людей в городах: а), б) - при начале эпидемии в Хабаровске; в), г) – в Санкт-Петербурге

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-svoiy-put-ves-god-mir-nablyudal-za-strategiy-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.cbsafety.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., Björnstad 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. а_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. Svoiy put': ves' god mir nablyudal za strategiyey Shvetsii po bor'be s koronavirusom. [Their path: The world has watched Sweden's strategy to combat Coronavirus all year]. Available at: <https://tjournal.ru/analysis/249657-svoiy-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. No2. 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).
37. 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).
38. Das D. Regional disparities of growth and internal migrant workers in informal sectors in the age of COVID-19.
39. Kojaku S., Hébert-Dufresne L., Mones E. et al. The effectiveness of backward contact tracing in networks. *Nat. Phys.* 2021. 17, pp. 652–658.
40. Klinkenberg D., Fraser C., Heesterbeek H. The effectiveness of contact tracing in emerging epidemics. *PLoS one*. 2006. Vol. 1. No. 1. pp. e12.
41. Müller J., Kretzschmar M., Dietz K. Contact tracing in stochastic and deterministic epidemic models. *Mathematical biosciences*. 2000. Vol. 164. No. 1. pp. 39-64.
42. 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.
43. 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.
44. Okolie A., Müller J. Exact and approximate formulas for contact tracing on random trees. *Mathematical biosciences*. 2020. Vol. 321. pp. 108320.
45. Bianconi G. et al. Message-passing approach to epidemic tracing and mitigation with apps. *Physical Review Research*. 2021. Vol. 3. No. 1. pp. L012014.
46. 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).
47. Coronavirus. URL: <https://coronavirus-monitor.info> (дата обращения: 07.06.2021).