

Биоинспирированный сценарный анализ связи эпидемии COVID-19 и технологий дистанционного обучения

Воронов М.В.*

Московский государственный психолого-педагогический университет
(ФГБОУ ВО МГППУ), г. Москва, Российская Федерация
ORCID: <https://orcid.org/0000-0001-7839-6250>
e-mail: mivoronov@yandex.ru

Давыдовский А.Г.**

Белорусский государственный университет информатики
и радиоэлектроники (БГУИР), г. Минск, Беларусь
ORCID: <https://orcid.org/0000-0003-1806-9674>
e-mail: agd2011@list.ru

Разработана структурная схема и представлено описание биоинспирированного квазимолекулярного алгоритма сценарного анализа динамики вариационных рядов. На основе этого алгоритма разработан и верифицирован комплекс пессимистических, инерционных и оптимистических сценариев связи поисковой активности интернет-пользователей и общей заболеваемости коронавирусной инфекцией. Оценены корреляционные связи между интенсивностями поисковых обращений по теме «дистанционное обучение» относительных частот поисковых обращений по различным коммуникативные информационно-технологическим платформам. Установлено, что распространение поисковых обращений интернет-пользователей по коронавирусной ассоциированной тематике имеет характер цепного разветвленного процесса, что свидетельствует о феномене «инфодемии COVID-19».

Ключевые слова: прогнозирование, сценарный анализ, дистанционное обучение, биоинспирированный алгоритм, COVID-19.

***Воронов Михаил Владимирович**, доктор технических наук, профессор, заведующий кафедрой прикладной математики. Московский государственный психолого-педагогический университет (ФГБОУ ВО МГППУ), г. Москва, Российская Федерация, ORCID: <https://orcid.org/0000-0003-0507-9144>, e-mail: mivoronov@yandex.ru

****Давыдовский Анатолий Григорьевич**, кандидат биологических наук, доцент кафедры, Белорусский государственный университет информатики и радиоэлектроники (БГУИР), г. Минск, Республика Беларусь, ORCID: <https://orcid.org/0000-0003-1806-9674>, e-mail: agd2011@list.ru



Для цитаты:

Воронов М.В., Давыдовский А.Г. Биоинспирированный сценарный анализ связи эпидемии COVID-19 и технологий дистанционного обучения // Моделирование и анализ данных. 2022. Том 12. № 4. С. 25–35. DOI: <https://doi.org/10.17759/mda.2022120402>

1. ПОСТАНОВКА ЗАДАЧИ

Современная жизнь стремительно усложняется. Как следствие, все более актуализируется принятие решений в условиях повышенного уровня неопределенности [1]. Наиболее сложные из такого рода ситуаций стали относить к трудно прогнозируемым и даже непрогнозируемым (была даже разработана теория, получившая название «Чёрный лебедь» [2]).

Для принятия обоснованных решений необходимо иметь представления об альтернативных гипотетически возможных вариантах развития сложившейся ситуации в будущем (сценариев) и оценивать вероятность реализации каждой из них [3, 4].

Глобальная эпидемия COVID-19 обусловила весьма существенное воздействие на социальные, политические, социально-экономические, социотехнические, социокультурные и другие аспекты жизни современного общества, что побуждает к активизации исследований в области прогнозирования развития земной цивилизации в целом и формирования сценариев последствий такого масштаба событий, в частности. Одной из перспективных платформ такого рода исследований является активно развиваемый в рамках системного анализа так называемый сценарный подход [5]. Среди инструментов сценарного подхода выделим, как наиболее подходящий, на наш взгляд, сценарный анализ по Р. Шварцу [6]. Его основные этапы следующие:

- определение фокуса исследования;
- указание ключевых факторов микросреды (внутренних характеристик);
- определение движущих сил макросреды (внешних факторов);
- ранжирование определенных внешних и внутренних факторов по степени их важности;
- выбор «логики»/способа генерации сценариев на основе подготовленных данных;
- конкретизация сценариев;
- рассмотрение последствий построенных сценариев.

Исследователи обратили внимание, что в период эпидемии *COVID-19* произошли существенные трансформации устоявшихся процессов в сфере образования. Связаны ли эти события или нет? Если да, то в какой мере и каковы последствия этого?

Замечено, что условиях глобальной эпидемии COVID-19 резко возрос поток поисковых запросов пользователей по тематике, связанной короновирусной инфекцией. Этот факт побудил выдвинуть гипотезу о том, что полученная при анализе такого рода запросов информация может служить полезным источником для разработки средств прогнозирования медико-биологических, социальных и социально-экономических причин различного рода социальных турбулентностей.

В этой связи была поставлена задача разработки и обоснования биоинспирированного алгоритма анализа сценариев распространения поисковых обращений поль-



зователей, относящихся к области дистанционного обучения, в условиях пандемии COVID-19.

2. БИОИНСПИРИРОВАННЫЙ КВАЗИМОЛЕКУЛЯРНЫЙ АЛГОРИТМ СЦЕНАРНОГО АНАЛИЗА

Для набора исходной информации в русскоязычном сегменте интернет-пространства были выявлены и зафиксированы данные о динамике поисковых обращениях по тематическому кластеру «Дистанционное образование». При этом были использован веб-сервис контекстной рекламы портала Google (<https://ads.google.com>), а также общедоступных сервисов Wordstat.Yandex.by (<https://wordstat.yandex.by>) за период с июля 2020 г. по август 2021 г. и GoogleTrends.com (<https://trends.google.ru>) за период с августа 2016 г. по июль 2021 г. На степень связи «дистанционное обучение», как ключевой темой, были также исследованы относительные показатели поисковых обращений по темам «электронное обучение», «online обучение», «анти-вирус», «Telegram», «Viber», «COVID-19», «Zoom», «Discord», «Moodle», «Skype», «Windows», «Linux», «Википедия», «мемы», «самоизоляция» и др.

Факторами, оказывающими наиболее существенное влияние на динамику поисковых обращений по теме «дистанционное обучение», оказались: «Антивирус», «Википедия», «Самоизоляция», «Мемы», «Moodle», «Skype», «Windows», «Zoom».

С помощью критериев Колмогорова-Смирнова и критерия Лиллиефорса был проведен анализ соответствия построенных вариационных рядов нормальному распределению осуществлен, который показал удовлетворительное соответствие исследуемых вариационных рядов нормальному распределению.

Исходные данные были обработаны с помощью комплексного метода корреляционно-регрессионного анализа с использованием критерия Пирсона, включая построение множественных линейных регрессионных моделей (МЛРМ), а также метода сценарного анализа с формированием оптимистического, базового (инерционного) и пессимистического сценариев. Для прогнозирования тенденций динамики поисковых обращений был использован метод экспоненциального сглаживания, а также основные положения теории аппроксимации [7].

Для прогнозирования связи заболеваемости COVID-19 с поисковой активностью интернет-пользователей был использован как традиционный подход, основанный на возможностях корреляционно-регрессионного аналитического моделирования (КРАМ), а также биоинспирированные методы и алгоритмы, включая впервые разработанный биоинспирированный квазимолекулярный алгоритм (БКМА).

В рамках данного подхода использована «гибридизация сценариев» с конверсией структурных коэффициентов при независимых переменных в МЛРМ из десятичной в бинарную форму и обратно. Полученные таким образом варианты для каждого сценария («гибриды») подвергались селекции по критериям минимума и максимума. Были получены «максимальные» (max, «макси»), «минимальные» (min, «мини»), усредненные варианты сценариев («миди»), а также гибридные варианты



оптимистического (optimistic – O), базового (basic – B) и пессимистического (pessimistic – P) сценариев.

Одним из методологических оснований БКМА является «сценарный крест». Этот методологический инструмент, как показано на рисунке 1 включает четыре различных сценария, базирующихся на взаимодействии двух неопределенностей (А) и (Б) [8].

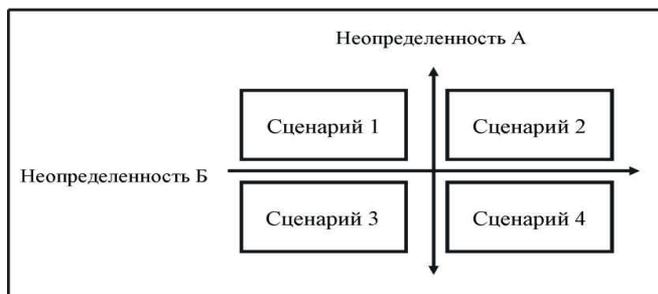


Рис. 1. «Сценарный крест».

При этом БКМА формирует пространство сценариев прогнозной динамики временных рядов, например, характеризующих частоту COVID-ассоциированных поисковых обращений интернет-пользователей. После конверсии из десятичной формы в бинарную такие «молекулы-мономеры» подвергаются преобразованию, которое аналогично изомеризации природных органических биомолекул с образованием двух зеркальных изомеров, каждый из которых может быть представлен в цис- или транс форме сцепления. При этом на основе одного исходного структурного коэффициента для каждой независимой переменной могут быть получены четыре варианта-«изомера», что в целом, соответствует методологическому подходу «сценарный крест». При этом пространственная хиральная изомеризация может быть реализована с помощью зеркальной изомеризации – «mirror»-эффекта, а цис-/транс-изомерия может быть осуществлена путем инверсии бинарного кода с заменой «1» на «0» и наоборот. После обратной конверсии – из бинарной формы в десятичную – можно получить квартет значений каждого структурного коэффициента для каждой независимой переменной в модели множественной линейной регрессии. После этого целесообразна селекция наибольших (макси-), минимальных (мини-) и средних, или средневзвешенных (миди-) значений каждого структурного коэффициента для каждой независимой переменной.

Далее путем гибридизации формируется популяция, содержащая три группы МЛРМ: макси-, миди- или мини-коэффициенты при независимых переменных. Тем самым формируется «целевой сценарий» (ЦС), который может включать как мини-, так и миди-, так и макси-коэффициенты при различных структурных переменных в МЛРМ.

На рисунке 2 представлена структурная схема биоинспирированного алгоритма формирования, изомеризации, гибридизации и селекции прогнозных сценариев динамики вариационных рядов.

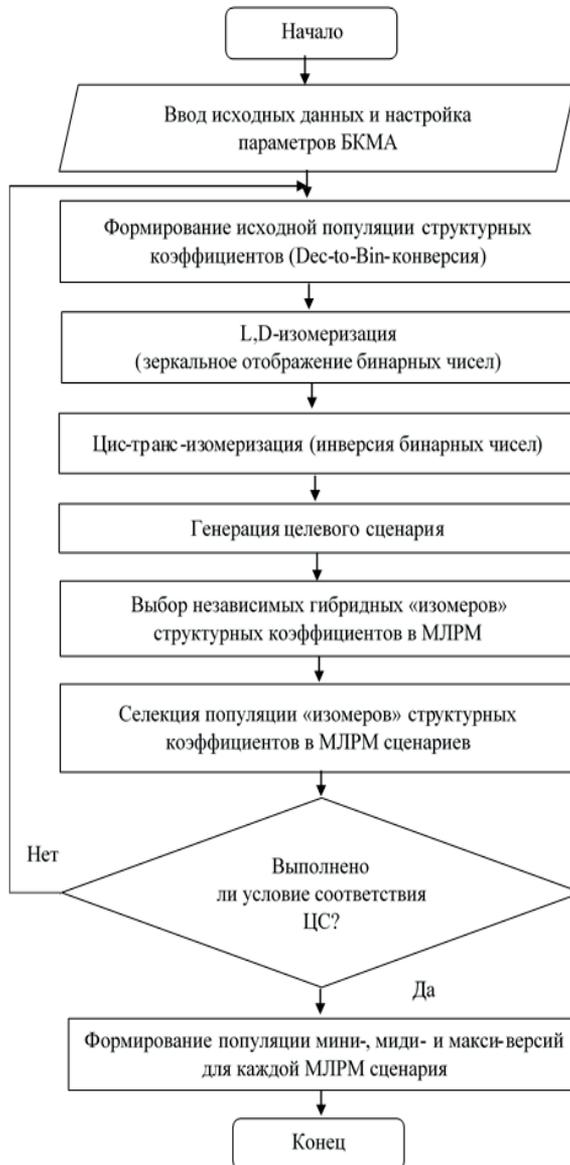


Рис.2. Структурная схема биоинспирированного алгоритма формирования, изомеризации, гибридизации и селекции прогнозных сценариев динамики вариационных рядов.

Популяции МЛРМ могут быть подвергнуты селекции по критерию соответствия ЦС. Соответствующие «полимерные биомолекулы» отбираются, сохраняются и используются для формирования прогностических сценариев в дальнейшем, а не соответствующие – элиминируются из общей популяции. На рисунке 3 представлена



структурная схема формирования прогнозных сценариев на основе метода биоинспирированного сценарного анализа.

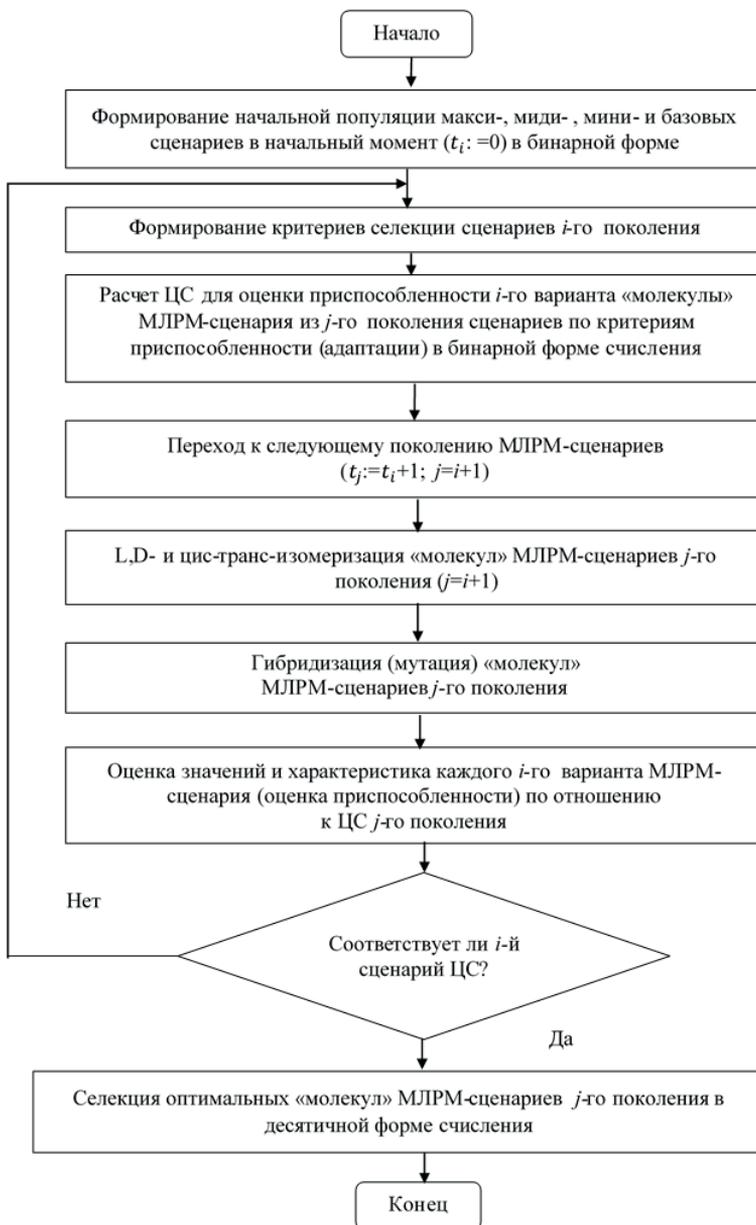


Рис. 3. Структурная схема формирования прогнозных сценариев на основе множественных линейных регрессионных моделей.



Для реализации сценарного анализа на основе БКМА разработано программное обеспечение (ПО), включающее на основе объектно-ориентированной структуры с соблюдением принципов инкапсуляции наследования и абстракции.

При практической реализации ПО БКМА для обеспечения его совместимости с языком программирования Python реализована библиотека PyQt5. Данное ПО включает класс `Ui_MainWindow`, `Controller`, классы `ConverterBase` и `Converter`, `ConverterBase`, `ExcelIO`, `ExcelIO`, классы `Correlation` и `ForwardSelection`, а также прочие классы модели, которые находятся в пакете `model`. В качестве основной библиотеки для работы с многомерными массивами была выбрана библиотека `NumPy`.

Для вычисления коэффициентов сценариев при анализе временных рядов использовался механизм минимизации скалярной функции нескольких переменных, где переменные – коэффициенты модели Холта-Винтерса, а кросс-валидация выступает в качестве непосредственно функции. По нашему мнению, для этих целей вполне подходит функция `optimize.minimize` из библиотеки `SciPy`.

3. РЕЗУЛЬТАТЫ РАСЧЁТОВ

На основе применения методологии сценарного подхода и технологии БКМА были исследованы различные аспекты поисковой активности пользователей интернет-сервисов, касающиеся распространения медиаконтента различной направленности в условиях эпидемии COVID-19 в Республике Беларусь [9]. С помощью веб-сервиса `Yandex.WordStat.by` выполнено исследование динамики поисковых обращений, связанных с тематическим локусом «дистанционное обучение» в период второй и третьей волны эпидемии коронавируса. Осуществлен корреляционный анализ частоты поисковых обращений с использованием интернет-пользователей в период с 27.07.2020 г. до 03.08.2021 г. Корреляционный анализ с использованием критерия Пирсона позволил установить следующие коэффициенты корреляции: по Пирсону ($P < 0,05$) между частотами поисковых обращений по тематическому локусу ДО, с одной стороны, и «электронное обучение» ($r=0,56$), «антивирус» ($r=0,36$), «Viber» ($r=0,28$), «COVID-19» ($r=-0,03$), «Zoom» ($r=0,86$), «Discord» ($r=-0,46$), «Moodle» ($r=0,86$), «Skype» ($r=0,65$), «мемы» ($r=0,69$), «Википедия» ($r=0,67$) и «Windows» ($r=0,51$), с другой стороны [10].

Кроме того, установлена достаточно сильная корреляция по частоте поисковых обращений по теме ДО и частотами поисковых обращений «электронное обучение» ($r=0,63$), «Zoom» ($r=0,89$), «Discord» ($r=-0,50$), «Moodle» ($r=0,92$), «Skype» ($r=0,68$), «мемы» ($r=0,85$), «Википедия» ($r=0,75$), «Windows» ($r=0,60$) и «Linux» ($r=0,37$). Показана сильная отрицательная корреляция между поисковыми обращениями в парах «COVID-19»–«Telegram» ($r=-0,55$) и «COVID-19»–«Viber» ($r=-0,76$).

Вместе с тем, проанализирована динамика исследуемых поисковых обращений интернет-пользователей в период с 04.07.2021 г. до 25.07.2022 г. и осуществлено прогнозирование динамики временных рядов поисковых обращений интернет-пользователей с помощью метода тройного экспоненциального сглаживания Холта-Винтерса.



Предложен прогнозный сценарий, основанный на достаточно сильной корреляционной связи поисковых обращениях по тематике ДО, с одной стороны, и поисковыми обращениями «антивирус» ($r=0,45$), «Zoom» ($r=0,89$), «Discord» ($r=-0,50$), «Moodle» ($r=0,92$), «Skype» ($r=0,68$), «мемы» ($r=0,85$), «Википедия» ($r=0,75$), «Windows» ($r=0,60$), с другой, среди русскоязычных интернет-пользователей.

Анализ результатов прогнозирования позволяет предположить возрастание влияния эпидемии коронавирусной инфекции на цифровую трансформацию технологий обучения, в направлении стабильного использования информационно-технологических платформ «Zoom», «Discord», «Moodle», «Skype» ($r=0,60$). Это позволило предположить, что в прогнозируемом периоде COVID-19 является фактором стимуляции заинтересованности интернет-пользователей в использовании для онлайн-обучения таких популярных мессенджеров, как «Viber» ($r=0,61$), «Telegram» ($r=-0,58$), «Википедия» ($r=0,75$), «COVID-19» ($r=0,23$), «Linux» ($r=0,37$). В то же время, повышенная отрицательная корреляция в парах «COVID-19»–«Telegram» ($r=-0,55$), а также «COVID-19»–«Viber» ($r=-0,76$) свидетельствует о возможной роли обращения к популярным мессенджерам «Telegram» и «Viber» в отвлечении внимания интернет-пользователей от проблематики эпидемии COVID-19.

Таким образом, показано, что в прогнозируемом периоде с 04.07.2021 г. до 25.07.2022 г. наиболее значимыми факторами, оказывающими существенное влияние на интенсивность поисковых обращений по тематике ДО, являются такие обращения как «антивирус», «википедия», «мемы», «самоизоляция», «электронное обучение», «COVID-19», «Discord», «Linux», «LMS» (Learning Management System), «Moodle», «online-обучение», «Skype», «Telegram», «Viber», «Windows», «Zoom».

В дальнейшем с помощью технологии БКМА были разработаны, обоснованы и исследованы пессимистические, инерционные и оптимистические сценарии связи поисковой активности интернет-пользователей и общей заболеваемости коронавирусной инфекцией. Анализ полученных сценариев позволяет предположить наличие сложных причинно-следственных связей между частотой поисковых обращений интернет-пользователей по тематике ДО и технологий социальных интернет-коммуникаций.

4. ЗАКЛЮЧЕНИЕ

На основе проведенных исследований можно сделать следующие выводы.

1. Разработан биоинспирированный квазимолекулярный алгоритм, в котором любой сценарий, представленный уравнением множественной линейной регрессии, рассматривается как полимерная молекула, а структурные коэффициенты при независимых переменных – как отдельные мономеры, которые могут подвергаться изомеризации, гибридизации, модификации и селекции.
2. Анализ полученных результатов свидетельствует об улучшении работы предложенного биоинспирированного квазимолекулярного алгоритма для генерации сценариев прогнозной динамики временных рядов комплексных данных по сравнению с традиционным подходом, основанном на корреляционно-регрессионном



анализе. При этом улучшение качества сценарного анализа для формирования прогнозных сценариев составляет от 24 % до 78 % в зависимости от сложности сценария, а по времени прогнозирования – от 62 % до 387 % анализа и прогнозирования по сравнению с традиционным подходом [10].

3. На основе вышеназванного алгоритма разработаны 287 множественных линейных регрессионных моделей сценариев прогнозной динамики временных рядов поисковых обращений интернет-пользователей с использованием данных веб-сервисов GoogleTrends.com (<https://trends.google.com/>) и WordstatYandex.by (<https://wordstat.yandex.by/>) по тематическим локусам «ДО», «электронное обучение», «online-обучение», «антивирус», «Telegram», «Viber», «COVID-19», «Zoom», «Discord», «Moodle», «Skype», «мемы», «википедия», «Windows», «Linux», «самоизоляция» в русскоязычном сегменте Сети в локации Республики Беларусь в период с 04.07.2021 г. до 25.07.2022.
4. Установлено, что распространение поисковых обращений интернет-пользователей по COVID-19-ассоциированной тематике имеет характер цепного разветвленного процесса, что косвенно указывает на быстрое распространение такого феномена, как «инфодемия COVID-19», в пространстве интернет.

Литература

1. *Вертешев С.М., Воронов М.В., Герасименко П.В., Кремков М.В.* Принятие решений при управлении организационными системами: Монография. – Псков: Псковский государственный университет, 2019. 218 с.
2. *Талёв Н.Н.* Чёрный лебедь. Под знаком непредсказуемости. Пер. с англ. М.: Колибри, 2009. 528 с.
3. *Van Notten Ph.* Writing on the wall: scenario development in times of discontinuity. Florida: Boca Raton; 2005, 209 p.
4. *Van Notten Ph.* Scenario development: a typology of approaches. Think Scenario. 2006: P. 69–84.
5. *Гладков Ю.М., Кононов Д.А. Кранчатв А.И.* Сценарное исследование социально-экономических систем: методология, задачи, практика применения. // Вестник РГГУ. Серия: Экономика, управление, право. 2007. № 12, С. 100–119.
6. *Schwartz P.* The art of the long view: planning for the future in an uncertain world. NY : Currency Doubleday; 1991, 271 p.
7. *Голубинский А.Н.* Методы аппроксимации экспериментальных данных и построения моделей. // Вестник Воронежского института МВД. 2007. № 2, С.138–143.
8. *Переверза Т.В.* Сценарный подход в анализе сложных социальных систем. Системные исследования и информационные технологии. 2011. № 1, С.133–143.
9. *Давыдовский А.Г., Лапицкая Н.В.* Социальные последствия цифровой трансформации образования в условиях коронакризиса. Журнал Белорусского государственного университета. Социология. 2022. № 2, С. 56–65.
10. *Давыдовский А.Г.* Методологические основы сценарного анализа связи заболеваемости COVID-19 и поисковой активности пользователей с помощью биоинспирированных алгоритмов. BIG DATA and Advanced Analytics = BI DATA и анализ высокого уровня: сб. научных статей VIII Междунар. науч.-практ. конф., Минск, 11–12 мая 2022 г. – Минск: Бестпринт, 2022. С. 462–475.



Bioinspired Scenario Analysis of the Connection between the COVID-19 Epidemic and Distance Learning Technologies

Mikhail V. Voronov*

Moscow State University of Psychology and Education (MSUPE), Moscow, Russia

ORCID: <https://orcid.org/0000-0001-7839-6250>

e-mail: mivoronov@yandex.ru

Anatoly G. Davydovsky**

Belarusian State University of Informatics

and Radioelectronics (BSUIR), Minsk, Belarus

ORCID: <https://orcid.org/0000-0003-1806-9674>

e-mail: agd2011@list.ru

The structural scheme is developed and the description of the bioengineered quasimolecular algorithm of scenario analysis of dynamics of variation series is presented. On the basis of this algorithm a complex of pessimistic, inertial and optimistic scenarios of connection of search activity of Internet users and general morbidity coronavirus infection were developed and verified. Correlation relations between the intensity of search queries on the topic «distance learning» of relative frequencies of search queries on various communicative information technology platforms have been evaluated. It has been found that the distribution of Internet users' search queries on coronavirus theme has the character of a chain branching process, which testifies to the phenomenon of «infodemia COVID-19».

Keywords: forecasting, scenario analysis, distance learning, bioinspired algorithm, COVID-19.

For citation:

Voronov M.V., Davydovsky A.G. Bioinspired Scenario Analysis of the connection between the COVID-19 Epidemic and Distance Learning Technologies. *Modelirovanie i analiz dannykh = Modelling and Data Analysis*, 2022. Vol. 12, no. 4, pp. 25–35. DOI: <https://doi.org/10.17759/mda.2022120402> (In Russ., abstr. in Engl.).

References

1. Verteshev S.M., Voronov M.V., Gerasimenko P.V., Kremkov M.V. Decision-making in the management of organizational systems: Monograph. – Pskov: Pskov State University, 2019. 218 p.

***Mikhail V. Voronov**, Doctor of Technical Sciences, Professor, Head of the Department of Applied Mathematics, Moscow State University of Psychology and Education (MSUPE), Moscow, Russia, ORCID: <https://orcid.org/0000-0003-0507-9144>, e-mail: mivoronov@yandex.ru

****Anatoly G. Davydovsky**, Candidate of Biological Sciences, Associate Professor of the Department, Belarusian State University of Informatics and Radioelectronics (BSUIR), Minsk, Republic of Belarus, ORCID: <https://orcid.org/0000-0003-1806-9674>, e-mail: agd2011@list.ru



2. Taleb N.N. Black Swan. Under the sign of unpredictability. Translated from English. M.: Hummingbird; 2009. 528 p.
3. Van Notten Ph. Writing on the wall: scenario development in times of discontinuity. Florida: Boca Raton; 2005, 209 p.
4. Van Notten Ph. Scenario development : a typology of approaches. Think Scenario. 2006. pp. 69–84.
5. Gladkov Yu. M., Kononov D.A. Krapchatv A.I. Scenario study of socio-economic systems: methodology, tasks, application practice. // Bulletin of the Russian State University. Series: Economics, Management, Law. 2007. No. 12, pp. 100–119.
6. Schwartz P. The art of the long view: planning for the future in an uncertain world. NY. : Currency Doubleday. 1991, 271 p.
7. Golubinsky A.N. Methods of approximation of experimental data and model construction. Bulletin of the Voronezh Institute of the Ministry of Internal Affairs. 2007. No. 2, pp. 138–143
8. Pereverza T.V. Scenario approach in the analysis of complex social systems. System research and information technology. 2011. № 1, pp. 133–143.
9. Davydovsky A.G., Lapitskaya N.V. Social consequences of digital transformation of education in the conditions of the coronacrisis. Journal of the Belarusian State University. Sociology. 2022. № 2, pp. 56–65.
10. Davydovsky A.G. Methodological foundations of scenario analysis of the relationship between the incidence of covid-19 and the search activity of users using bioinspired algorithms. BIG DATA and Advanced Analytics = BI DATA and high-level analysis : collection of scientific articles of the VIII International Scientific and Practical Conference, Minsk, May 11–12, 2022 – Minsk: Bestprint, 2022. pp. 462–475.

Получена 15.11.2022

Принята в печать 04.12.2022

Received 15.11.2022

Accepted 04.12.2022