

С. П. ЛЕВАШКИН, О. И. ЗАХАРОВА
НИЛ ИИ-ПУТИ, ФГБУ ВО, Самара

С. В. КУЛЕШОВ, А. А. ЗАЙЦЕВА
ЛАНИ-СПИИРАН, СПб ФИЦ РАН, Санкт-Петербург

АДАПТИВНО-ПОВЕДЕНЧЕСКАЯ МОДЕЛЬ РАСПРОСТРАНЕНИЯ КОРОНАВИРУСНОЙ ЭПИДЕМИИ И ЕЕ ОПТИМИЗАЦИЯ МЕТОДАМИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Предлагаемый доклад посвящен описанию системного подхода к исследованию новой модели распространения пандемии COVID-19, который имеет конечной целью минимизацию привносимого ею экономического ущерба за счет выбора оптимальных управленческих параметров модели. Наш подход состоит из двух основных частей: 1) адаптивно-поведенческой модели распространения эпидемии, которая является обобщением классической модели SEIR и 2) модуля настройки параметров этой модели методами искусственного интеллекта (сбор, хранение и обработка данных из гетерогенных источников, которые позволяют наиболее точно настраивать параметры нашей модели, превращая ее в интеллектуальную систему для поддержки принятия управленческих решений). Среди параметров модели наиболее важными являются индивидуальные экономико-демографические и психологические характеристики общества и действия властей.

Введение. Сегодня весь мир столкнулся с пандемией COVID-19, которая, по всей вероятности, будет иметь долгосрочные последствия для экономики государств и общества в целом. Данную ситуацию однозначно можно рассматривать как серьезную биогенную угрозу, которая наносит ощутимый ущерб социально-экономическим системам, государственным образованиям и обществу. При этом сверхактуальными научными задачами являются коррекция существующих и разработка новых математических моделей распространения эпидемий и методов цифрового мониторинга не только темпов распространения вируса, но и социально-экономической среды. И те, и другие должны отражать современные реалии и как можно точнее представлять, как саму биогенную угрозу, так и ее последствия. Поэтому своевременный мониторинг экономических, психологических и иных социальных процессов, протекающих внутри такой среды, позволит принимать более оптимальные управленческие решения и минимизировать последствия пандемии за счет более точной настройки прогнозной модели.

Предлагаемый доклад посвящен описанию гибридной модели распространения вирусных заболеваний, сочетающую методы машинного обучения, интеллектуального анализа данных и теории сложных систем. Отталкиваясь от хорошо известной модели распространения эпидемий SEIR [1] мы ее существенно модифицируем, вводя в нее новые процессы и параметры, которые более точно отражают современные реалии (эта новая модель называется адаптивно-поведенческой или ACM-SEIR). Затем эта модель дополняется системой комплексного мониторинга социальных явлений, нацеленных на минимизацию привносимого ею ущерба (таким образом формируется наш конечный продукт – интеллектуальная адаптивно-поведенческая модель или iACM-SEIR для поддержки принятия управленческих решений).

Научная новизна предлагаемого исследования обусловлена:

- 1) интеграцией системы мониторинга и модели распространения вирусных эпидемий;
- 2) выбором в качестве источников для уточнения параметров моделей неструктурированных данных на неограниченном наборе ресурсов, а не ограниченного набора predetermined интернет-источников;
- 3) выявлением слабо прогнозируемых реакций общества на те или иные события, выраженные Интернет контентом с использованием методов машинного обучения.

Адаптивно-поведенческая модель эпидемии. За основу модели распространения эпидемии берется модель «восприимчивый к заражению-подверженный воздействию-зараженный-удаленный» (англ. Susceptible-Exposed-Infectious-Removed: SEIR) с общим размером популяции N и с двумя дополнительными классами (i) «D», имитирующими общественное восприятие риска в отношении числа тяжелых и критических случаев и смертей; и (ii) «C», представляющими количество кумулятивных случаев (как зарегистрированных, так и не зарегистрированных).

ных). Пусть S , E и I обозначают восприимчивые, подверженные и заразные группы населения, а R обозначает удаленную группу (то есть выздоровевших или умерших). В модели определяется степенная функция скорости реакции властей или «эффект от действий правительства». Передача инфекции внешними носителями (зоонозная передача), обозначаемая как F , моделируется как ступенчатая функция, которая принимает нулевое значение после некоторой даты с начала эпидемии. Далее моделируется только устойчивая передача COVID-19 от человека к человеку после этой даты и мобильность населения до того, как был объявлен режим самоизоляции. Таким образом, адаптивная поведенческая модель (англ. Adaptive Compartmental Model: ACM) ACM-SEIR формулируется в виде системы 7 сильно нелинейных обыкновенных дифференциальных уравнений (ОДУ) для определения основных переменных S , E , I , R , N , D , C , z (z – переменная, отвечающая за силу воздействия правительственных органов) их начальных условий, а также множества параметров, подлежащих определению и настройке. Заметим, что, помимо введения новых параметров и функций, базовая модель ACM-SEIR допускает обобщение за счет добавления в систему новых уравнений для других процессов типа S , E , I , R и т.д. Кроме того, для модели количество уравнений системы и количество ее параметров не являются критичными. 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 используется подход, основанный на методах глубокого обучения. Суть этого подхода заключается в следующем. Сначала решение системы ОДУ записывается в виде суммы двух частей. Первая часть удовлетворяет начальным условиям и не содержит настраиваемых параметров. Вторая часть построена так, чтобы не влиять на начальные условия. Эта часть включает в себя искусственную нейронную сеть с прямой связью, содержащую настраиваемые параметры (веса). Таким образом, по построению начальные условия удовлетворяются, и сеть обучается, чтобы удовлетворять поставленной задаче в целом. Причем, при вычислении решения, шаг за шагом минимизируется некая «функция потерь» (решение будет «точным», если эта функция равна 0). Заметим, что для оптимизации процесса вычисления можно заменять на нейронную сеть не все правые части уравнений, а только их нелинейные части, которые доставляют наибольшие проблемы (по сути речь идет об обучении на множестве временных рядов).

Определение и настройка параметров модели. Для настройки параметров и функций модели используются методы искусственного интеллекта на основе цифрового мониторинга Интернет пространства [2], в частности, мониторинга психологического состояния общества, оценки свойств и состояний людей по текстам. Основными параметрами модели ACM-SEIR являются: {начальная численность рассматриваемой популяции N_0 (константа); начальные восприимчивые, подверженные, заразные и удаленные популяции – S_0 , E_0 , I_0 , R_0 (константы); количество зоонозных случаев F (ступенчатая функция); скорость передачи z_0 (ступенчатая функция); правительственная сила действия a (ступенчатая функция); интенсивность реакции населения k (константа); уровень мобильности m (ступенчатая функция); средний латентный период $1/s$ (константа); средний инфекционный период $1/g$ (константа); доля тяжелых случаев d (константа); средняя продолжительность общественной реакции $1/l$ (константа)}. Также элементом настройки параметров модели является анализ настроений – инструмент классификации текста, который анализирует входящее сообщение и определяет, является ли реакция автора по теме сообщения положительной, отрицательной или нейтральной [3, 4, 5]. Анализ настроений может быть дополнен анализом намерений, который анализирует намерения пользователя, стоящие за сообщением, и выявляет, относится ли оно к мнению, новости, действию, жалобе, предложению, оценке или запросу. Эти два вида анализа дополняются контекстным семантическим поиском, который принимает тысячи сообщений и концепт (например, «COVID-19») в качестве входных данных и фильтрует все сообщения, которые точно соответствуют данному концепту. Наконец, полученные результаты оформляются в виде статистических отчетов и визуализируются в виде графических чартов для отражения картины состояния и настроения пользователей по той или иной проблеме и помогают провести тюнинг параметров модели [6, 7].

Закключение. В докладе представлена гибридная модель интеллектуальной адаптивно-поведенческой модели распространения короновирусной эпидемии и показана возможность оптимизации привносимого ею экономического ущерба за счет точной настройки параметров

модели. Принятие реальных решений на предотвращение/смягчение угроз требует, чтобы проигрывалось множество возможных вариантов с относительно несложным аппаратом интерпретации, за небольшой отрезок времени. Поэтому мы используем сочетание современных технологий машинного обучения, искусственного интеллекта и методов системной динамики. Такой подход позволяет справиться с основными трудностями поставленной задачи, а именно: 1) будучи реальным феноменом, задача содержит очень большое число параметров, которые, в конце концов, необходимо учесть, чтобы сделать модель как можно более приближенной к реальным процессам; 2) ожидаемая размерность пространства данных, необходимая для достижения той же цели, весьма высока, а количество данных велико. Эти два обстоятельства делают практически неприменимым традиционный математический аппарат статистического анализа, регрессий, сеточных супервычислений и т.п. Однако современные методы глубокого обучения, интеллектуального анализа данных и облачных вычислений, позволяют справиться с многофакторностью задачи, «проклятием размерности» и большим количеством данных необходимых для реалистичной настройки модели. Заметим также, что рассматриваемая задача носит комплексный междисциплинарный характер и связана с разработкой методологии, инструментов и технологий повышения эффективности государственного управления в кризисных ситуациях.

*Исследование выполнено при финансовой поддержке РФФИ в рамках
научного проекта № 20-04-60455*

ЛИТЕРАТУРА

1. **van den Driessche P.** Reproduction numbers of infectious disease models. *Infect Dis Model.* 2017 Jun 29. Vol 2(3). P. 288-303. doi: 10.1016/j.idm.2017.06.002
2. **Kuleshov S.V., Zaytseva A.A., Aksenov A.J.** The tool for the innovation activity ontology creation and visualization. *Advances in Intelligent Systems and Computing.* Springer, Cham, 2019. Vol. 763. pp. 292-301
3. **Raghavendra T. S., Mohan K. G.** Web mining and minimization framework design on sentimental analysis for social tweets using machine learning. *Procedia Computer Science.* 2019. T. 152. P. 230–235.
4. **Patel N. V., Chhinkaniwala H.** Investigating Machine Learning Techniques for User Sentiment Analysis. *International Journal of Decision Support System Technology (IJDSST).* 2019. T. 11. №. 3. P. 1–12.
5. **Priya C., Santhi K., Vincent P. M. D.** Provision of Efficient Sentiment Analysis for Unstructured Data. *Information Systems Design and Intelligent Applications.* Springer, Singapore, 2019. P. 199–207.
6. **John C. A.** Using Twitter to measure behavior patterns. *Epidemiology.* 2012. Iss. 23. Vol. 5. P. 764-765.
7. **Аверкин А.Н., Соболев С.В., Воронцов А.О.** Сравнение различных техник анализа эмоций для решения задачи визуализации индекса настроения. *Мягкие измерения и вычисления.* 2019. №. 11. С. 30-34.

S.P. Levashkin, O.I. Zakharova (LAI-PSUTI, Samara), S.V. Kuleshov, A.A. Zaytseva (LANI-SPIIRAN, SPCRAS, St.-Petersburg)

Adaptive-compartmental Model of Coronavirus Epidemic and its Optimization by the Methods of Artificial Intelligence

In this report, we describe a systemic approach to the study of a new model of COVID-19 pandemic that has the main goal to minimize its economic damage by defining the optimal model parameters. Our approach consists of two main parts: 1) the adaptive-compartmental model of the epidemic (ACM-SEIR) that is a generalization of the classical SEIR model, and 2) the module to tune the ACM-SEIR parameters using artificial intelligence methods (collection, storage and processing of big data from heterogeneous sources that allow the most accurate tuning of the parameters of the ACM-SEIR model and turning it into an intelligent iACM-SEIR system for decision support). Among the parameters of the model, the most important are the individual economic, demographic and psychological characteristics of society and the governmental actions.