

ОЦЕНКА ПРОСТРАНСТВЕННО-ВРЕМЕННЫХ ЗАКОНОМЕРНОСТЕЙ ДОСТАВКИ ЕДЫ ИЗ СЕТЕВЫХ РЕСТОРАНОВ БЫСТРОГО ПИТАНИЯ (на примере г. Минска)

И. С. Князев

кафедра почвоведения и ГИС факультета географии и геоинформатики Белорусского государственного университета, г. Минск, geo.knyazev@bsu.by

А. А. Сазонов

кафедра почвоведения и ГИС факультета географии и геоинформатики Белорусского государственного университета

Исследование посвящено пространственно-временным закономерностям доставки еды из сетевых ресторанов быстрого питания в городе Минске и прогнозированию количества заказов различными методами. Выполнена оценка прогноза, выявлены основные тренды. Установлено, что общее число заказов имеет положительную динамику, а основные закономерности связаны с погодными условиями и выходными днями.

Ключевые слова: сервис доставки; пространственно-временные закономерности; прогнозирование на основе случайного леса; прогнозирование на основе экспоненциального сглаживания; прогнозирование на основе подгонки кривой; профиль Фурье; кластер временных рядов.

Целью исследования являлись выявление и оценка пространственно-временных закономерностей в доставке еды из сетевых ресторанов быстрого питания, а также прогноз динамики количества заказов.

В ходе выполнения данной работы были поставлены следующие задачи:

1. используя данные по количествам заказов ресторанов быстрого питания, построить куб пространственно-временных закономерностей;
2. проанализировать пространственно-временное распределение количества заказов по микрорайонам города Минска;
3. составить прогноз на две недели по созданному кубу методами экспоненциального сглаживания, случайного леса, подгонки кривой;
4. сравнить результаты прогнозирования и определить наилучший для каждого микрорайона.

Сервис доставки – сервис (услуга) доставки еды и иных товаров на дом или в офис покупателя. Самыми известными сервисами доставки еды в Беларуси являются «Яндекс.Еда» и «Деливио», некоторые из крупных сетей имеют свои службы доставки («МакДональдс», «Бургер Кинг», «Доминос» и т. д.).

В нашем исследовании были использованы данные по количеству заказов из ресторанов быстрого питания одного из крупных сервисов доставки еды в г. Минске.

Всего было проанализировано более 300000 заказов в различные микрорайоны города Минска в период с 19 июля 2021 года по 19 декабря 2021 года. Значения количества заказов были спрогнозированы тремя различными методами с

20 декабря 2021 года по 2 января 2022 года. В исследовании анализировались заказы из трёх крупнейших сетевых ресторанов быстрого питания.

Для анализа количества заказов использовалась группа инструментов углубленного анализа пространственно-временных закономерностей геоинформационной системы «ArcGIS Pro». Набор инструментов содержит статистические инструменты для анализа распределения данных и выявления закономерностей в контексте пространства-времени. Выделяются группы инструментов анализа пространственно-временных закономерностей, прогнозирования временных рядов, визуализации данных куба пространства-времени в 2D и в 3D.

На первоначальном этапе был создан куб пространства-времени. Куб агрегирует точечные объекты в бины пространства-времени в файлы данных netCDF. В пределах каждого бина подсчитываются точки и агрегируются указанные атрибуты. Для всех местоположений бинов проводится оценка трендов числа объектов, а также вычисляются суммы значений полей.

В нашем случае данные были агрегированы по микрорайонам города Минска, временной шаг соответствует одной неделе (с понедельника по воскресенье). Для корректного анализа и прогноза некоторые из исходных микрорайонов были исключены в связи с недостаточным количеством данных. Куб был визуализирован в виде 3D-модели (рисунок 1) для наглядной интерпретации [1].

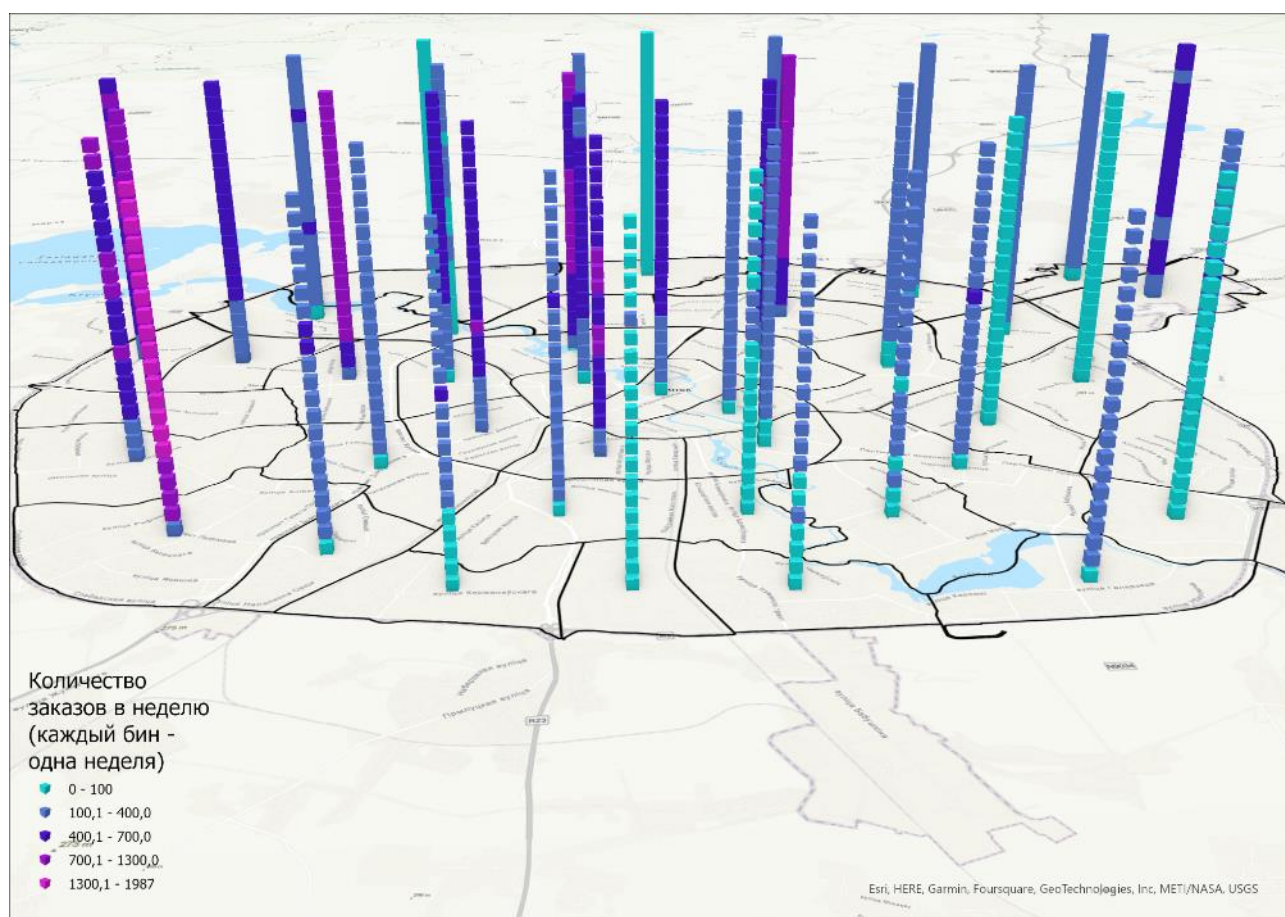


Рис. 1. Визуализация куба пространственно-временных закономерностей

Следующим этапом исследования стало прогнозирование количества заказов на 2 недели и сравнение результатов прогноза.

Прогноз осуществлялся тремя методами:

прогнозирование на основе случайного леса предсказывает значения в каждом местоположении куба пространства-времени при помощи адаптированного алгоритма произвольного леса – метода контролируемого машинного обучения Лео Бреймана и Адели Катлер. Регрессионная модель леса обучается с использованием временных окон в каждом местоположении куба пространства-времени;

прогнозирование на основе экспоненциального сглаживания определяет значения каждого местоположения куба пространства-времени с помощью метода экспоненциального сглаживания Хольта-Винтерса путем разложения временного ряда в каждом кубе местоположения на сезонные компоненты и компоненты тренда;

прогнозирование на основе подгонки кривой вписывает параметрическую кривую в каждое место входного куба пространства-времени и прогнозирует временной ряд, экстраполируя эту кривую на будущие шаги [5].

Следующим этапом стало сопоставление результатов прогнозирования и выявление наиболее объективных. Сравнение и оценка прогнозов выполнялась с помощью инструмента “Оценка прогнозов по местоположению” группы инструментов “Прогнозирование временных рядов”. Для оценки точности прогнозов данный инструмент определяет для каждого местоположения метод прогнозирования, который дает наименьшее значение средней квадратической ошибки (RMSE) [2]. В результате была составлена картосхема, показывающая наиболее точный метод прогноза в каждом микрорайоне Минска (рисунок 2).

Для анализа полученных результатов была проведена кластеризация данных.

Кластеризация временных рядов разделяет временные ряды, хранящиеся в кубе пространства-времени, на основе их сходства. Временные ряды могут быть кластеризованы на основе трех критериев: одинаковых значений времени, одновременной тенденции к увеличению и уменьшению и сходных повторяющихся закономерностей.

В нашем случае кластеризация проводилась методом профиля Фурье (рисунок 3).

В профиле Фурье временные ряды схожи, если имеют одинаковые гладкие периодические закономерности значений во времени. Эти периоды иногда называют циклами или сезонами, и они представляют собой длительность закономерности, которая затем повторяется в новом периоде. Метод позволяет увидеть периодически повторяющиеся еженедельные закономерности в общих продажах с периодом, начинающимся в понедельник и заканчивающимся в воскресенье [4].

Наиболее объективным методом прогноза в большинстве микрорайонов оказался метод экспоненциального сглаживания. Большая часть микрорайонов с этим методом прогнозирования оказалась во втором кластере. Исходя из рисунков 2 и 3, видно, что метод наиболее эффективен в районах со средним

количеством заказов и в районах, где отсутствует явно выраженная динамика (количество заказов минимально изменяется на всей кривой).

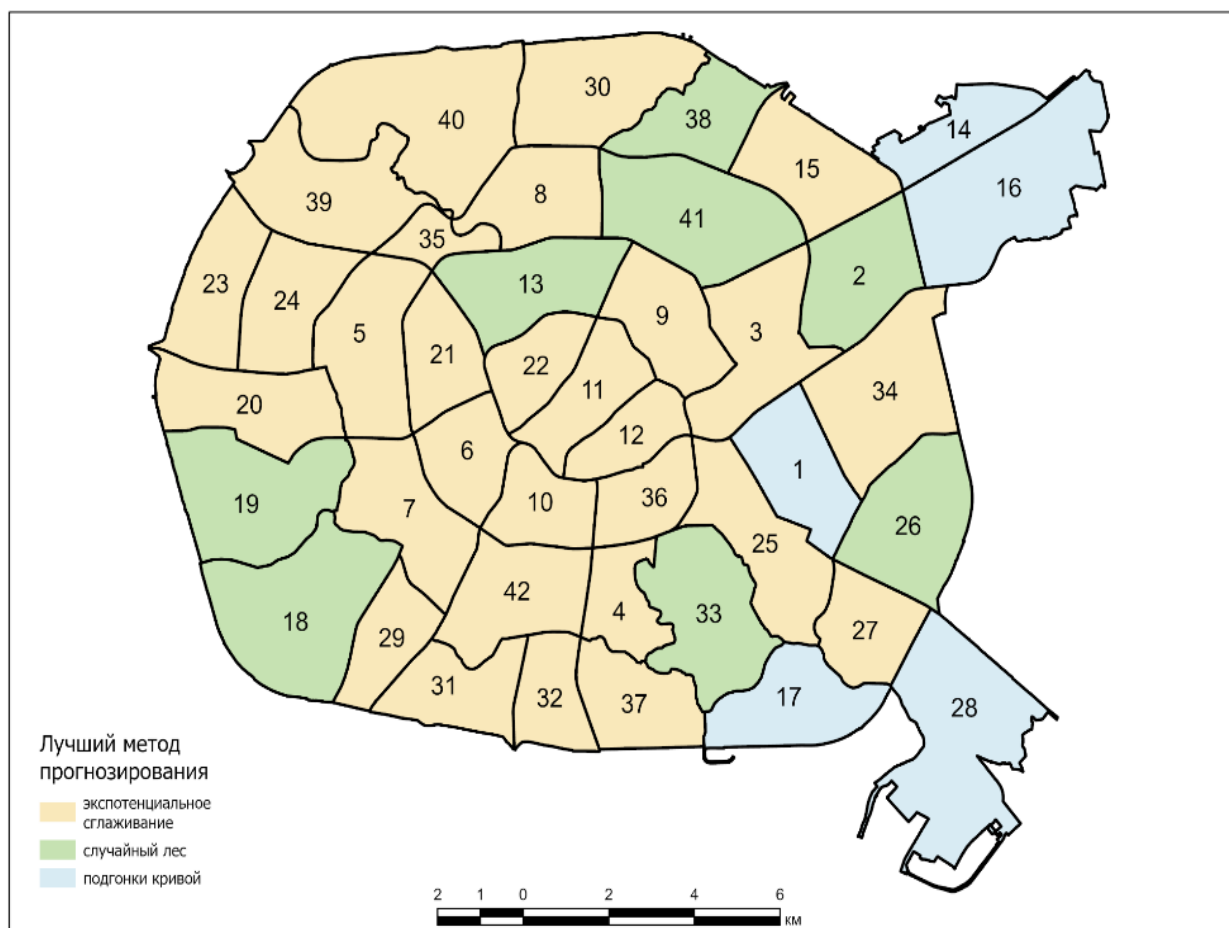


Рис. 2. Наиболее объективные методы прогноза (соответствие районов номерам:

1 – Тракторный, 2 – Староборисовский Тракт, 3 – Макаенка – Столетова, 4 – Камвольный, 5 – Прищыцкого – Сердича, 6 – Проспекта Дзержинского, 7 – Гурского, 8 – Старовиленский, 9 – Площади Я. Коласа, 10 – Жуковского, 11 – Центр-2, 12 – Центр-3, 13 – Парк Победы, 14 – Уручье – Копище, 15 – Зеленый Луг – Восток, 16 – Уручье – Восточный, 17 – Чижевка, 18 – Юго-Запад – Малиновка, 19 – Сухарево, 20 – Запад – Красный Бор, 21 – Кальварийская, 22 – Центр-1, 23 – Каменная Горка, 24 – Масюковщина – Кунцевщина, 25 – Ванеева – Васнецова, 26 – Ангарская, 27 – Автозавод, 28 – Ельница – СЭЗ, 29 – Дружба, 30 – Новинки, 31 – Курасовщина, 32 – Серова, 33 – Серебрянка, 34 – Дrajня – Степянка, 35 – Веснянка, 36 – Коммунарка, 37 – Лошица, 38 – Зеленый Луг, 39 – Лебяжий, 40 – Зацень, 41 – Богдановича – Толбухина, 42 – Аэропорт).

Для районов с наименьшим количеством заказов (1-й кластер) более эффективен метод подгонки кривой. Тренд здесь еще более умеренный, чем во втором кластере, количество заказов в среднем в два раза меньше.

Единственным районом, который оказался в 3-ем кластере, является район “Юго-Запад – Малиновка”. В данном микрорайоне самый неоднородный тренд, а лучшим методом прогноза является случайный лес. Этот микрорайон выделяется аномально высоким числом заказов. Более того, если подробно анализировать кривую, то можно обнаружить, что количество заказов уменьшается во время введения дистанционного обучения и отъезда на каникулы иногородних

студентов, которые в большом количестве проживают в данном микрорайоне (студенческая деревня находится именно здесь). Четкий рост заметен между 8 и 10 неделями (с 1 по 15 сентября) – это период приезда студентов после летних каникул. На 12 и 13 неделе (с 11 по 23 октября) по графику ввели дистанционное обучение, и также заметен чёткий рост числа заказов. Явно падение числа заказов выявляется ближе к 20 неделе, то есть с новогодними праздниками и каникулами.

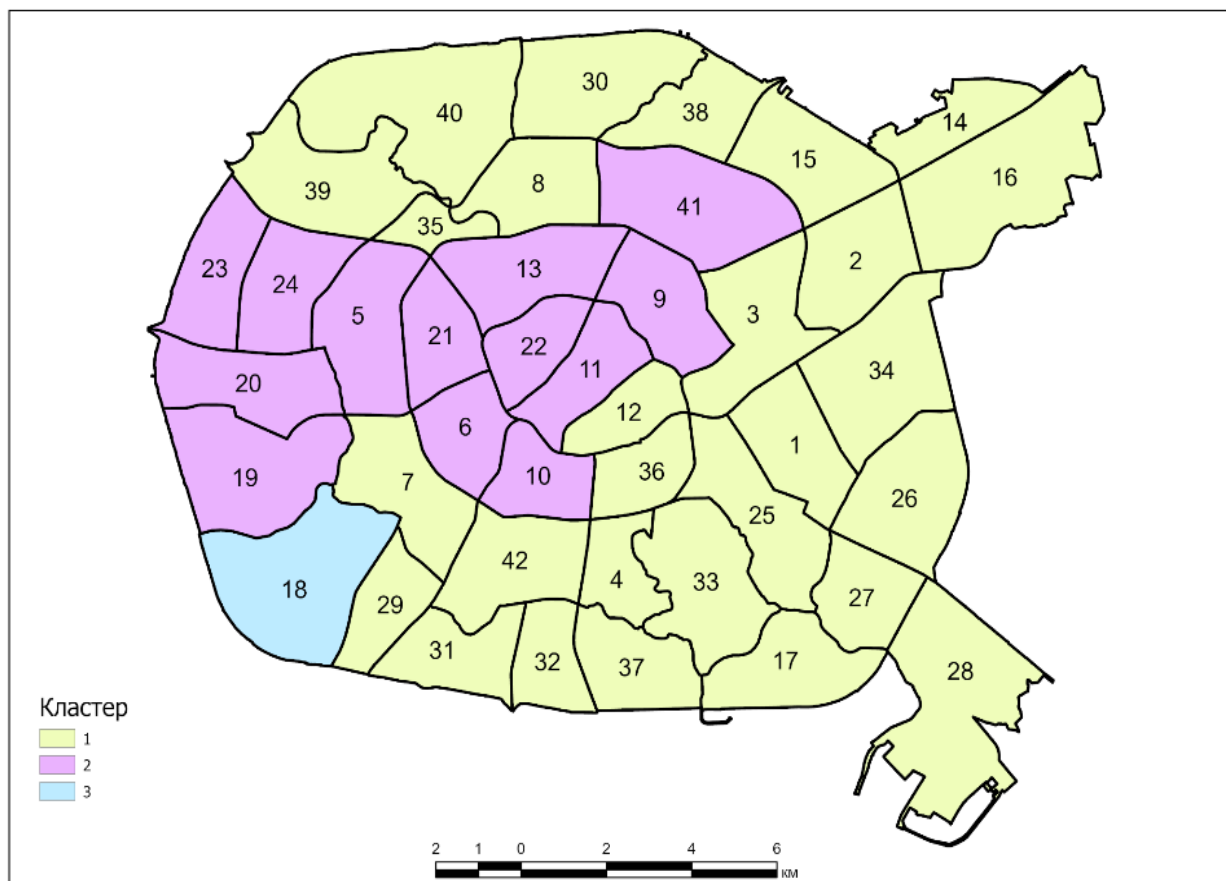


Рис. 3. Результат кластерного анализа

(соответствие районов номерам: 1 – Тракторный, 2 – Староборисовский Тракт, 3 – Макаенка – Столетова, 4 – Камвольный, 5 – Притыцкого – Сердича, 6 – Проспекта Дзержинского, 7 – Гурского, 8 – Старовиленский, 9 – Площади Я. Коласа, 10 – Жуковского, 11 – Центр-2, 12 – Центр-3, 13 – Парк Победы, 14 – Уручье – Копище, 15 – Зеленый Луг – Восток, 16 – Уручье – Восточный, 17 – Чижевка, 18 – Юго-Запад – Малиновка, 19 – Сухарево, 20 – Запад – Красный Бор, 21 – Кальварийская, 22 – Центр-1, 23 – Каменная Горка, 24 – Масюковщина – Кунцевщина, 25 – Ванеева – Васнецова, 26 – Ангарская, 27 – Автозавод, 28 – Ельница – СЭЗ, 29 – Дружба, 30 – Новинки, 31 – Курасовщина, 32 – Серова, 33 – Серебрянка, 34 – Дrajня – Степянка, 35 – Веснянка, 36 – Коммунарка, 37 – Лошица, 38 – Зеленый Луг, 39 – Лебяжий, 40 – Зацень, 41 – Богдановича – Толбухина, 42 – Аэропорт).

На основе кластерного анализа составлен график динамики заказов по кластерам (рисунок 4), отражающий основные тренды [3].

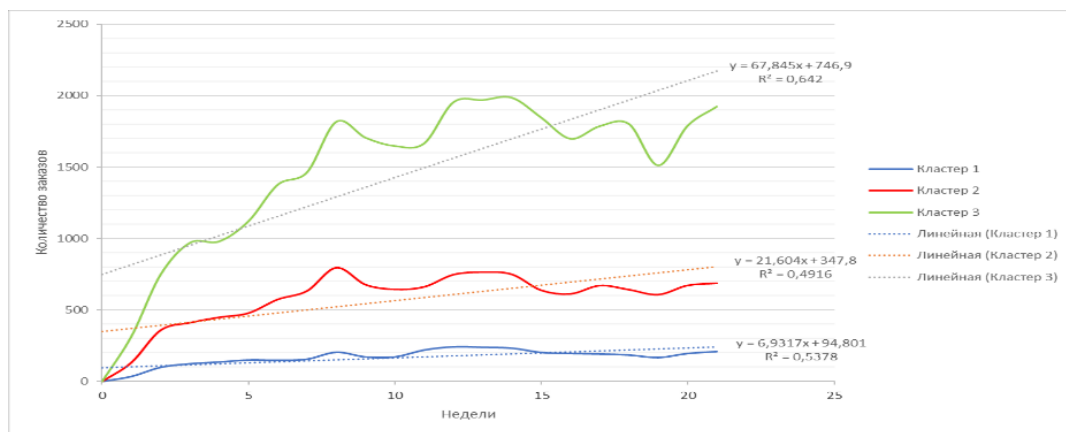


Рис. 4. Диаграмма среднего количества заказов по каждому кластеру

В целом по городу Минску просматривается тенденция на рост заказов. Возможно, это связано с тем, что временной срез взят практически с самого начала открытия сервиса доставки. Сервис приобретал популярность, со временем минчане узнавали о его существовании и начали пользоваться им активнее.

БИБЛИОГРАФИЧЕСКИЕ ССЫЛКИ

1. Andrienko, N.: Exploratory spatio-temporal visualization: an analytical review / N. Andrienko, G. Andrienko, P. Gatalsky // Journal of Visual Languages and Computing/ – 2003. –14(6). – P. 503–541.
2. Putrenko, Viktor and Pashynska, Natalia: Application of geoinformation modeling tools for intelligent data analysis of fire hazards. Bulletin of National Technical University “KhPI”, Series New Solutions in Modern Technologies. – 2017. – №7 – P. 156–163
3. Putrenko, V. Data Mining of Network Events With Space-Time Cube Application [Electronic resource] / V. Putrenko, N. Pashynska, S. Nazarenko. – Kyiv, 2018. – Mode of access: https://www.researchgate.net/publication/328082401_Data_Mining_of_Network_Events_with_Space-Time_Cube_Application. – Date of access: 10.11.2022.
4. Афанасьев, В. Н. Анализ временных рядов и прогнозирование [Электронный ресурс] / В.Н. Афанасьев: Министерство науки и высшего образования Российской Федерации, Оренбургский государственный университет. – Оренбург, 2020. – 286 с.
5. Обзор набора инструментов Углубленный анализ пространственно-временных закономерностей [Электронный ресурс] / Официальный сайт ESRI. – Режим доступа: <https://pro.arcgis.com/ru/pro-app/2.9/tool-reference/space-time-pattern-mining/an-overview-of-the-space-time-pattern-mining-toolbox.htm>. – Дата доступа: 10.11.2022.