

Пространственный анализ данных для исследования потенциала территории

Плотникова М.А., СПбГУ, Санкт-Петербург st087861@student.spbu.ru

Аннотация

В данной работе рассматривается задача выбора наилучшего местоположения для открытия новой точки оператора сотовой связи "Билайн" на территории города Москвы. Потенциальные участки получены путем разбиения географической карты на гексагоны и последующего анализа их привлекательности на основе метода Вороного. В исследовании используются данные из открытых источников, таких как Яндекс.Карты и Реформа ЖКХ. Предложены 10 наилучших мест для открытия магазина, результаты представлены на интерактивной карте.

Введение

При выборе открытия наилучшей точки, ритейлером должны учитываться множество факторов. Существует несколько подходов [1], [2] к решению данной задачи. В данной работе стояла цель использовать только данные из открытых источников, поэтому был выбран подход на основе геопространственного анализа. Этот подход находит широкое применение в различных отраслях (например, проект по установке солнечных батарей [3], построению зданий [4] или заправок станций [5] на местности), поскольку он универсален и интерпретируем, что делает его подходящим инструментом для принятия обоснованных решений для бизнеса.

Алгоритм, представленный в этой работе, будет сосредоточен на расчете универсальных метрик с использованием географических и демографических данных.

Конкретно использовались данные о географическом местоположении текущих магазинов и о плотности населения, информация о которых была выгружена из "Яндекс.Карт" и данных проекта "Реформа ЖКХ" соответственно. На основе выгруженных данных была построена метрика, оценивающая доступность магазинов для населения, и предложены наилучшие географические зоны для улучшения покрытия.

В работе используются такие библиотеки python, как geopandas, numpy, osmnx, folium и shapely. Реализованы интерактивные карты для удобного взаимодействия с данными.

Выгрузка данных

Для анализа использовались следующие источники данных:

Open Street Map [7] представляет собой некоммерческий веб-картографический проект по созданию силами сообщества участников – пользователей интернета подробной свободной и бесплатной географической карты мира.

С помощью библиотеки OSMnx [8] были выгружены административные границы города Москвы в виде объекта Polygon библиотеки shapely.

Яндекс.Карты [9] - поисково-информационная картографическая служба Яндекса. Сервис предоставляет фиксированное число бесплатных поисковых запросов через API к основной системе поиска.

Использовался тип запроса "Поиск по организациям" для выгрузки данных о магазинах операторов сотовой связи. Выгруженные данные:

- название магазина;
- адрес;
- координаты точки на карте мира;

Реформа ЖКХ [10] - автоматизированная информационная система, созданная по заказу Фонда содействия реформированию ЖКХ, ориентированная на мониторинг региональных программ переселения граждан из аварийного жилья, программ капитального ремонта многоквартирных домов. Выгруженные данные:

- площадь жилых помещений, площадь нежилых помещений в м²;
- адрес здания;
- координаты здания.

Расчет проживающих в здании производился по формуле:

$$\left(\frac{\text{area} \times 0.35}{33} + \frac{\text{area} \times 0.65}{18} \right) \quad (1)$$

Где area - это площадь жилого дома. Формула была составлена, исходя из норм для проживающих в однокомнатных и квартирах с большим числом комнат соответственно - 33м² для проживающих в однокомнатной квартире, и 18м² - для семей.

Разбиение территории на полигоны

После загрузки границ Москвы, было построено покрытие гексагонами библиотеки `h3`. 1

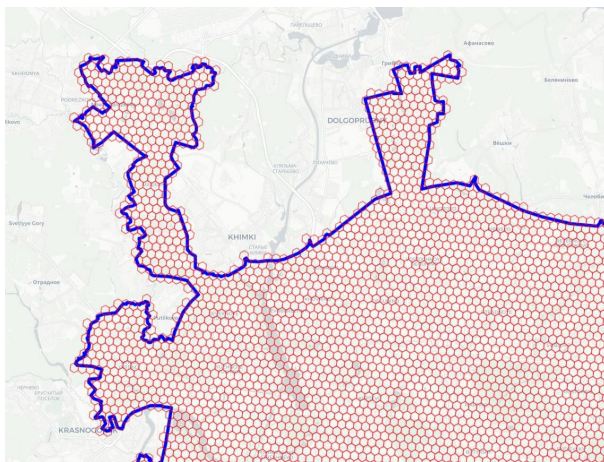


Рис. 1: Построение покрытия города Москвы гексагонами

Гексагон, обладающий наиболее подходящими характеристиками, в будущем будет выбран в качестве решения поставленной задачи.

Используемые python библиотеки: `numpy` [11], `pandas` [12], `geopandas` [13], `h3` [14].

Целью при обработке географических данных являлась оценка качества доступности торговых точек для населения.

Был применен метод Вороного [6] для разбиения карты в зависимости от удаленности до точки. Полученное разбиение было пересечено с полигонами. Для каждого полигона полученного разбиения была посчитана дистанция до ближайшего магазина, рассчитанная, как максимальное расстояние из полигона до точки (метрика Хаусдорфа). Используемые python библиотеки: `geovoronoi` [15], `shapely` [16], `folium` [17].

Оценка доступности торговых точек для населения

Полученные полигоны были пересечены с координатами населенных домов. Итоговая метрика ρ была подсчитана по формуле

$$\rho = \frac{p}{d}$$

где p - это плотность населения на участке, d - максимальное расстояние до ближайшего магазина.

В случае, когда гексагон пересекался с разбиением Вороного и имел несколько наиболее близких к нему точек оператора "Билайн", метрика подсчитывалась отдельно для каждого участка, и затем усреднялась для всего гексагона. Результаты были сформированы в виде файла csv и выведены на интерактивную карту 2.

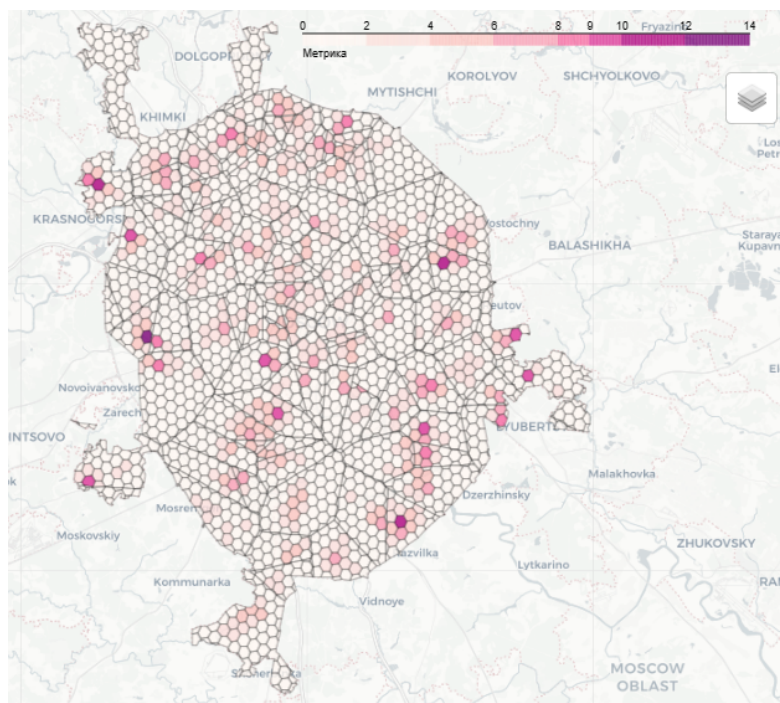


Рис. 2: Гексагоны окрашены в зависимости от значения построенной метрики.

Улучшение покрытия

В качестве наиболее подходящих мест для открытия новых точек были выбраны 10 значений с наибольшим значением метрики.

В центрах данных гексагонов были созданы новые точки, которые в дальнейшем участвовали в повторном разбиении 3.

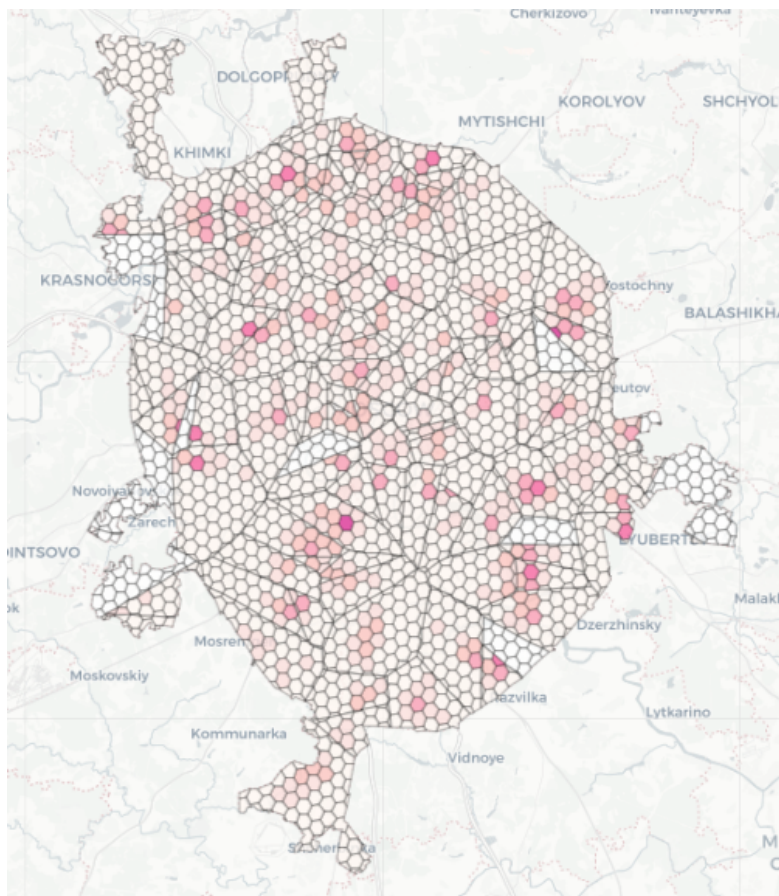


Рис. 3: Карта нового разбиения, построенная после добавления 10 дополнительных точек. Гексагоны окрашены в зависимости от значения построенной метрики.

На карте наглядно продемонстрировано улучшение качества покрытия после добавления 10 новых точек.

Приложение

Код, разработанный в рамках данного исследования, а также интерактивные карты доступны по ссылке:

https://github.com/Ima0203/GEO_searching

Заключение

В ходе работы были исследованы существующие подходы в решении подобных задач, реализована загрузка и обработка данных из открытых источников с использованием API технологии и библиотек языка python. Был проведен отбор наиболее качественных из них и подходящих для решения задачи.

На основе алгоритма Вороного и разбиения территории на гексагоны, была разработана метрика доступности магазинов для населения. На основе ее показателей, была реализована интерактивная карта города Москвы с указанием наиболее подходящих местоположений для открытия новой точки. Решение было представлено с помощью интерактивных карт folium.

Список литературы

- [1] Grin, Petr. *Utilization of machine learning algorithms to support retail chain store location decisions*. University of Northern Iowa, 2020.
- [2] Ting, Choo-Yee, Mang Yu Jie. *Location Profiling for Retail-Site Recommendation Using Machine Learning Approach*, 2020.
- [3] Al Garni, H. Z., Awasthi, A. *Solar PV Power Plant Site Selection Using a GIS-AHP Based Approach with Application in Saudi Arabia*. Applied Energy, 2017.
- [4] Kumar, S., Bansal, V. *A GIS-based Methodology for Safe Site Selection of a Building in a Hilly Region*. Frontiers of Architectural Research, 2016.
- [5] Messaoudi, D., Settou, N., Negrou, B., Rahmouni, S., Settou, B., Mayou, I. *Site Selection Methodology for the Wind-powered Hydrogen Refueling Station Based on AHP-GIS in Adrar, Algeria*. Energy Procedia, 2019.
- [6] *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams*. John Wiley & Sons, 2009.
- [7] «Open Street Map». URL: <https://openstreetmap.org/>
- [8] OSMnx. URL: <https://github.com/gboeing/osmnx>
- [9] «Yandex API». URL: <https://yandex.ru/maps-api/>
- [10] «Открытые данные ЖКХ». URL: <https://dom.mingkh.ru/>
- [11] NumPy. URL: <https://numpy.org/>
- [12] Pandas. URL: <https://pandas.pydata.org/>
- [13] GeoPandas. URL: <https://geopandas.org/>
- [14] H3. URL: <https://github.com/uber/h3-py>
- [15] Geovoronoi. URL: <https://github.com/WZBSocialScienceCenter/geovoronoi>
- [16] Shapely. URL: <https://github.com/Toblerity/Shapely>
- [17] Folium. URL: <https://github.com/python-visualization/folium>