

МИНОБРНАУКИ РОССИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«ВОРОНЕЖСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»
(ФГБОУ ВО «ВГУ»)

Факультет компьютерных наук

Кафедра программирования и информационных технологий

*Формирование карты рейтинга жилых районов с точки зрения
инфраструктуры*

Курсовая работа

09.03.04 Программная инженерия

Допущен к защите

Зав. кафедрой _____ С.Д. Махортов, д.ф. - м.н., доцент __.__.2021

Обучающийся _____ А.В. Калиткин, 3 курс, д/о

Руководитель _____ Н.К. Самойлов, ст. преподаватель

Воронеж 2021

Содержание

Содержание.....	2
Введение	3
1 Постановка задачи	4
2 Анализ предметной области	6
2.1 Обзор существующих решений	6
2.2 Сценарий использования приложения	10
2.3 Выбор источника данных об инфраструктуре.....	11
2.4 Выбор средств доступа к данным	12
2.5 Выбор технологий и архитектуры для приложения	14
2.6 Структура данных OSM	15
2.7 Использование Overpass API	15
2.8 Выбор данных об инфраструктуре	16
3 Разработка приложения.....	18
3.1 Структура разработанного приложения.....	18
3.2 Выполнение и оптимизация запросов	19
3.3 Расчёт рейтинга.....	22
3.4 Построение тепловой карты рейтинга	23
3.5 Оптимизация алгоритмов.....	25
Заключение	27
Список использованных источников	29

Введение

Выбор подходящего места для покупки или аренды жилья – одна из задач, с которой сталкиваются люди при планировании или уже в процессе переезда на новое место жительства. Как правило, на выбор всегда оказывает влияние большое количество факторов. К ним можно отнести ограниченность предложения на рынке, отсутствие вариантов с удобным расположением и приемлемой ценой, наличие поблизости социальной инфраструктуры и остановок общественного транспорта, уровень преступности в данном районе, экологическую обстановку и многие другие.

В случае, если возможных вариантов всего несколько, каждый из них можно досконально изучить самостоятельно. В том случае, если выбор достаточно велик или будущее место жительства не привязано к конкретному району города, на самостоятельное изучение обстановки придётся потратить во много раз больше времени.

Один из способов упростить анализ – использовать некую метрику для оценки всех возможных вариантов. Эта метрика может быть как субъективной и формироваться на основе данных социологических опросов, так и объективной и основываться на данных из открытых источников. Но в любом случае, единого варианта существовать не может, так как каждый человек ведёт свой образ жизни и имеет разные потребности, по-разному оценивает важность тех или иных критериев.

Критерий, который очень важен проживающим в городе семьям – наличие поблизости инфраструктуры и возможность удобно добираться из дома до неё и до места учёбы или работы всех членов семьи. Таким образом, метрика, вычисляемая на основе этого критерия будет подходить большинству, кто столкнётся с необходимостью оценить качество того или иного места в городе с точки зрения комфортного проживания.

Для удобства восприятия, полученную метрику логично будет представлять в виде тепловой карты, отображаемой поверх карты города.

1 Постановка задачи

Цель работы – разработать веб-приложение, позволяющее пользователям строить тепловую карту качества района для проживания, основанную на данных об инфраструктуре, получаемых из OpenStreetMap.

Конечные пользователи – люди, изучающие другой город с целью выбора наилучшего возможного будущего места жительства с точки зрения инфраструктуры в связи с планируемым переездом и не обладающие специальными техническими знаниями или умениями.

Разрабатываемое приложение должно соответствовать следующим требованиям:

- Работать в браузерах, актуальных на момент разработки
- Использовать для работы данные, получаемые из OpenStreetMap
- Позволять пользователю открыть на экране произвольный участок карты
- Позволять выбрать пользователю один из заранее определённых наборов инфраструктурных объектов, участвующих в расчёте рейтинга
- Рассчитывать и отображать полупрозрачную тепловую карту рейтинга для районов, отображаемых на экране пользователя, по нажатию кнопки
- Отображать численное значение рейтинга в точке на карте по нажатию на эту точку

В рамках анализа предметной области необходимо:

- Изучить методы и источники данных, использованные в общедоступных исследованиях и программных продуктах
- Определить преимущества и недостатки существующих решений
- Выбрать источник данных об инфраструктуре с учётом их условий использования и накладываемых лицензией ограничений, а также средства для доступа к ним

- Изучить структуру данных выбранного источника
- Выбрать средства реализации и архитектуру приложения

Непосредственно реализация приложения состоит из следующих задач:

- Получение приложением данных об инфраструктуре для выбранного на карте участка
- Поиск ближайшего к точке объекта определённой категории
- Расчёт расстояния и затрат времени на перемещение между точками на карте
- Расчёт рейтинга конкретной точки на карте с точки зрения инфраструктуры
- Формирование картинки тепловой карты рейтинга
- Профилирование и оптимизация используемых алгоритмов

2 Анализ предметной области

2.1 Обзор существующих решений

В ходе анализа предметной области было выяснено, что в большинстве случаев рейтинги районов городов составляются по результатам опросов общественного мнения, в которых жителей просят оценить их текущее место проживания в городе по нескольким критериям. Примерно половину из этих критериев составляют вопросы об обстановке в районе в целом с упором на такие аспекты, как чистота, экология, безопасность и работа коммунальных служб; остальные – о различных видах инфраструктуры: магазинах, развлечениях, инфраструктуры для детей, общественном транспорте и т. д.



Рисунок 1 – Рейтинг районов Воронежа по результатам опросов жителей

Такие опросы позволяют оценить в целом район города, получить некоторую цифру, по которой этот район можно сравнить с другими, а также выявлять тенденции к изменению тех или иных показателей, но они имеют недостаточно высокую точность. Для того, чтобы привязка полученного рейтинга к географическому положению жителей была максимально точной, а погрешность из-за субъективности в ответах на опросы была минимальной,

необходимо проводить опрос, имея довольно большую выборку, что определяет большую сложность данной задачи.

Из общего ряда выделяются исследования Яндекса, посвящённые Москве и Санкт-Петербургу с точки зрения доступности инфраструктуры для жизни. Эти исследования были проведены аналитически, используя данные Яндекс.Карт об организациях. В ходе исследования наиболее заселённые районы городов были разделены на квадраты со стороной 300 метров, и затем каждый из квадратов относили к одному из трёх классов – «Недостаточно инфраструктуры», «Удобно для жизни» или «Очень удобно для жизни» на основании доступности для жителей этого квадрата всей необходимой для жизни инфраструктуры в пешей доступности.



Рисунок 2 – Проведённое Яндексом исследование Москвы с точки зрения доступности инфраструктуры для жизни

В данном исследовании квартал считался удобным, если не более чем за 10 минут из его центра можно дойти до двух любых продуктовых магазинов, общеобразовательной школы, детского сада и аптеки, а так же не более, чем за 15 минут – до объектов инфраструктуры тех видов, которые не предполагают частого посещения среднестатистических жителей, такие, как поликлиники, хозяйственные магазины и парикмахерские.

Очень удобными считались те кварталы, для которых выполнены требования удобного квартала, а также не более чем за 15 минут можно было дойти до ближайшего бара, предприятия общественного питания, кинотеатра и станции метро или МЦК. Кварталы, которые не выполнили требования удобных для жизни, были отнесены к классу «Недостаточно инфраструктуры».

Аналогичные карты были составлены для развлекательных организаций, или учитывая все виды инфраструктуры вместе. Затем полученные по квадратам данные были сгруппированы по районам, к которым они относятся, и по полученным результатам были сформированы интерактивная карта и таблица с рейтингами и особенностями районов – относительной долей квадратов каждого из классов, среднее время пешком до заведений и организаций разных видов и комментарии исследователей.

Стоит так же упомянуть сервис Яндекс.Недвижимость, который кроме отображения на карте доступного для продажи или аренды жилья даёт так же даёт возможность посмотреть их собственный рейтинг инфраструктуры и транспортной доступности в виде тепловой карты.

Для расчёта индекса транспортной доступности в данных городах были выделены по 50 наиболее популярных точек, таких как крупные офисы и торговые центры, вокзалы и аэропорты, университеты и места отдыха и рассчитано среднее время в пути из каждого жилого дома до этих точек. Цвета на карте показывают примерное время в пути – зелёные зоны показывают места, для которых полученное суммарное время вышло минимальным.

Цвета́ми тёмно-красных тонов помечены районы, добираться из которых до выбранных точек дольше всего.

Карта инфраструктуры показывает, насколько плотно расположены в районе сервисы для жизни и развлечения. В ней учитываются 9 типов организаций: покупки на каждый день (супермаркеты и т.п.), не ежедневные покупки (книги, мебель, бытовая техника и т.п.), еда (кафе, рестораны), здоровье, красота, бытовые услуги, спорт, развлечения и культура. Для расчёта город был поделён на квадраты 250x250 метров, и каждому из них присвоена оценка от 1 до 8. Районы с единицей окрашены в тёмно-красный цвет и имеют не более одного из полезных сервисов. Районы с максимальным рейтингом – это зоны, в которых работает более 66 таких сервисов, и окрашены они в тёмно-зелёный цвет.

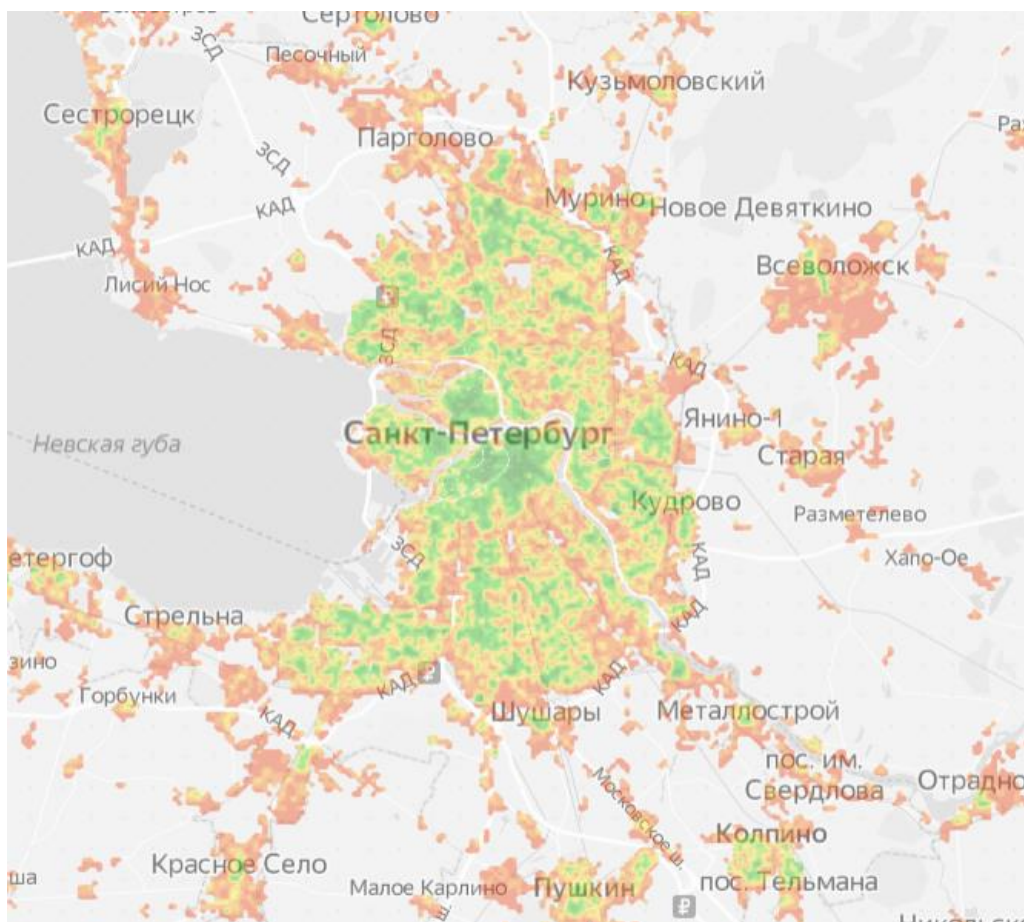


Рисунок 3 – Тепловая карта инфраструктуры для жизни и развлечений на примере Санкт-Петербурга

К сожалению, как исследования Яндексa, так и тепловая карта, доступная в сервисе Яндекс.Недвижимость, обладают несколькими недостатками. Самый существенный – это ограниченность их покрытия только столичными городами России при том, что данный сервис, как и Яндекс.Карты, доступен по всей стране и даже в некоторых соседних странах. Второй по значимости недостаток - невозможность выбрать пользовательский сценарий использования, тем самым ограничив набор инфраструктурных объектов, участвующих при формировании рейтинга.

2.2 Сценарий использования приложения

Рассматривая опыт уже существующих решений, для разрабатываемого приложения был сформирован следующий типовой сценарий использования:

- Открыть в браузере страницу приложения
- Перейти на интерактивной карте на интересующий регион
- Выбрать наиболее подходящий сценарий использования
- Нажать на кнопку отрисовки результата
- Получить на экране тепловую карту рейтинга

Так же у пользователей должна быть возможность при необходимости посмотреть численное значение рейтинга в конкретной точке, нажав на эту точку на карте.

На рисунке 4 представлена диаграмма прецедентов, составленная с учётом сформированного выше сценария использования.

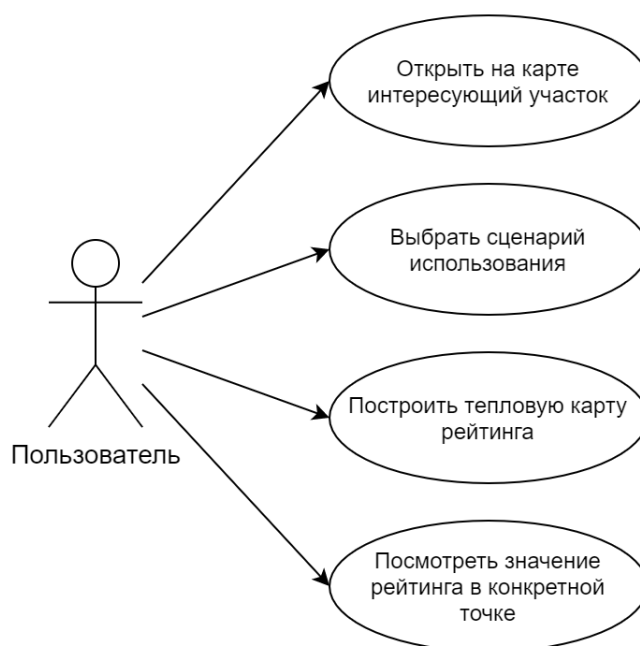


Рисунок 4 – Диаграмма прецедентов

2.3 Выбор источника данных об инфраструктуре

На данный момент есть несколько сервисов, предоставляющих доступ к картам. Большой популярностью в России пользуются Яндекс.Карты и 2GIS, на западе – Google Maps. Однако, это всё закрытые проекты, которые требуют регистрации и соблюдения некоторых ограничений даже в случае некоммерческого использования и не подходят требованиям проекта.

Лицензия Яндекс.Карт не допускает использования бесплатного API в приложениях с закрытым исходным кодом, в коммерческих целях или в проектах с ограниченным доступом. Так же не допускается изменение и сохранения полученных данных при бесплатном использовании. Тарифы коммерческой версии рассчитаны на их использование с большим количеством запросов в сутки, в случае необходимости обработки получаемых данных стоимость дополнительно увеличивается на 500 000 рублей в год.

Подобные ограничения так же накладываются лицензионным соглашением API 2GIS. Как в случае Яндекса, так и в случае 2GIS, нужно предварительно получать разрешение на использование данных их сервиса и

придерживаться данных ограничений, которые не позволяют решить поставленную задачу в полной мере.

В качестве источника данных об инфраструктуре был выбран проект OpenStreetMap (далее OSM), предоставляющий большое количество разнообразных географических данных, которые могут быть использованы совершенно свободно. Данные OpenStreetMap доступны для последующего использования под лицензией для открытых данных Open Database License (далее ODbL), построенной на принципе share-alike.

Лицензия ODbL не накладывает ограничения на способ использования данных, допускает их использование в приложениях с закрытым исходным кодом, в коммерческих целях или в приложениях с ограниченным доступом, а так же не накладывает ограничения на лицензию, под которой должны распространяться работающие с этими данными приложения и их исходные коды. Единственные выдвигаемые лицензией требования – необходимость упомянуть сообщество OSM как источник данных и делать свои изменения доступными под той же лицензией при дополнении базы геоданных. Лицензия не заставляет распространять или делать доступными публично данные, полученные на основе данных из OSM. Однако в случае их распространения, на них должны распространяться на тех же условиях, что и данные OpenStreetMap, в том числе и другими лицами.

Наполнение карты (количество обозначенных объектов, актуальность, количество метаданных), а так же её качество (соответствие метаданных принятым стандартам) может варьироваться от города к городу, но в большинстве городов России и почти всех городах Европы достаточно хорошее, чтобы рассчитывать свой собственный рейтинг на основе этих данных.

2.4 Выбор средств доступа к данным

Данные, предоставляемые OpenStreetMap, можно получить несколькими способами, каждый из которых имеет свои ограничения и недостатки в каждом конкретном сценарии его использования.

При необходимости использования приложения оффлайн или обработки запросов от большого числа пользователей целесообразно развернуть свой собственный сервер базы данных, на котором будет находиться копия данных OSM. Для первоначального развёртывания необходимо загрузить все данных планеты или региона полностью, в дальнейшем для поддержания актуальности данных необходимо раз в неделю загружать с серверов OSM файл изменений и применять его к своей копии данных.

OpenStreetMap предоставляет два API – первый, editing API, используется только приложениями для редактирования карт, так как позволяет извлекать самые актуальные данные и вносить изменения на сервер. Он не позволяет извлекать данные с больших территорий и довольно ограничен. Второй, Overpass API, позволяет писать в декларативном стиле довольно сложные запросы к базе и извлекать с их помощью данные сразу на большой территории.

Сервисы, предоставляющие данные API, могут быть развёрнуты локально для использования с собственным сервисом базы данных. Одним из преимуществ использования своей копии OSM – возможность отправлять запросы на выборку данных или производить поиск маршрутов, обращаясь напрямую к серверу базы данных с использованием расширения PostGIS для СУБД PostgreSQL. Однако, объём данных OSM в несжатом виде со всеми индексами достаточно велик, и для развёртывания своей копии OSM требуется сервер или рабочая станция с диском достаточной производительности и вместимости. По этой причине на данном этапе решено было использовать общедоступные API, а собственный сервер развернуть на следующих этапах выполнения работы.

По заверению их администраторов общедоступного сервера Overpass API, они достаточно производительны, чтобы их можно было использовать, не беспокоясь, что может быть оказано влияние на работу других приложений,

если приложение совершает в сутки не более 10 000 запросов и выгружает в сутки не более 5 ГБ данных.

Для тестирования запросов к Overpass API во время выполнения работы использовался сервис Overpass Turbo, позволяющий писать и выполнять запросы, а также отображать результат запроса в текстовом виде и на карте прямо в окне браузера.

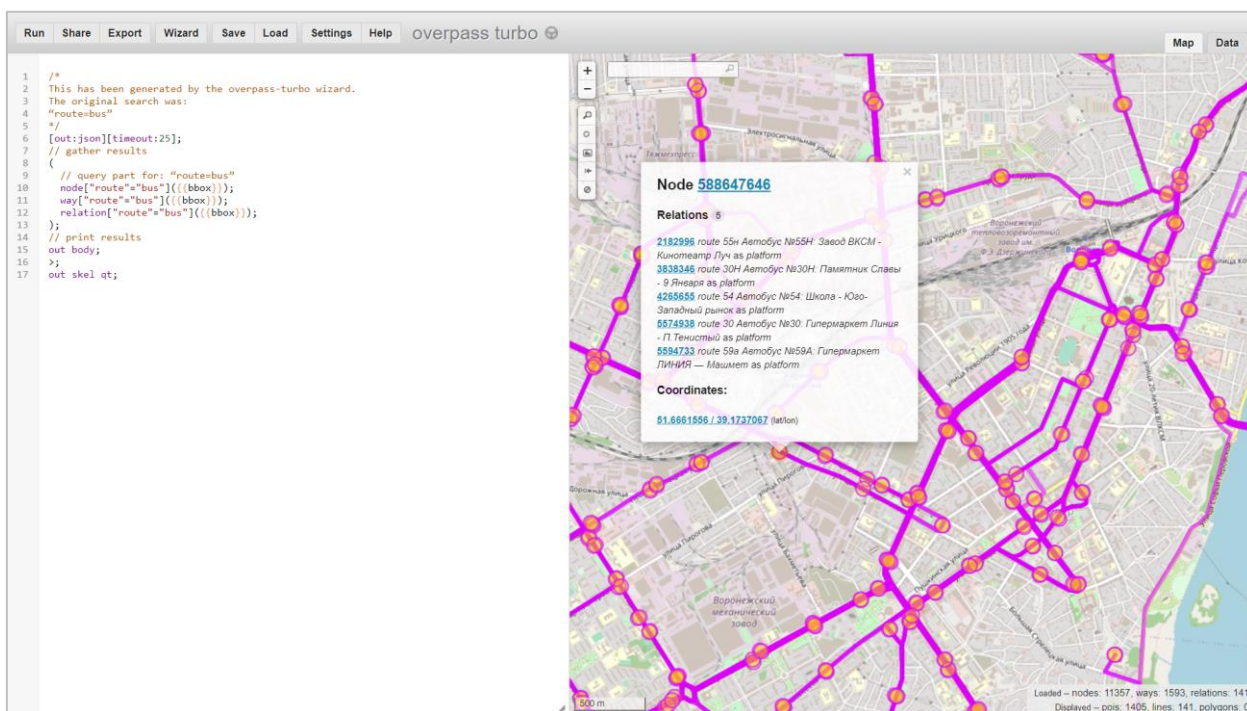


Рисунок 5 – Интерфейс Overpass Turbo с примером запроса к API

Стоит заметить, что несмотря на свои возможности, этот Overpass API позволяет только извлекать объекты, но не проводить с ними вычисления на стороне сервера-интерпретатора запросов. По этой причине невозможно использовать Overpass Turbo как решение поставленной задачи без его доработки.

2.5 Выбор технологий и архитектуры для приложения

Так как приложение не должно требовать от пользователей для своей работы его установки, логично реализовать его в виде одностороннего веб-приложения, которое может распространяться как набор статических файлов с любого веб-сервера. Для создания клиентского приложения было решено использовать фреймворк Vue и язык программирования TypeScript, для

создания пользовательского интерфейса – библиотеку Vuetify, для отображения интерактивной карты – библиотеку Leaflet.

Выбор языка обусловлен наличием в нём строгой типизации и возможности писать код в привычном объектно-ориентированном стиле, что существенно ускоряет разработку и уменьшает вероятность ошибок. На этапе сборки проекта код на TypeScript компилируется в JavaScript-код, который сможет выполняться в браузере, но в котором уже отсутствует информация о типах. Совместимость TypeScript с JavaScript позволяет использовать JavaScript-библиотеки из TypeScript кода, и наоборот.

2.6 Структура данных OSM

В OSM используется только три вида базовых элементов, которыми можно описать любую сущность, требующую представления на карте. Эти элементы – это точка (node), путь (way) и отношение (relation). Точка – элемент, содержащий широту и долготу. Путь – это ломаная линия, соединяющая точки в заданном порядке. Ломаная линия может быть замкнутой (полигон) или не замкнутой. Их часто используют для улиц и других объектов, имеющих геометрию. Отношения задают принадлежность точек, путей или полигонов к определённой группе.

У каждого элемента могут быть заданы атрибуты (теги), которые представляют собой пары ключ=значение. По ним можно понять, что перед нами объект какого-то определённого типа либо извлечь метainформацию, например, о названии и режиме работы заведения.

2.7 Использование Overpass API

Запросы к Overpass API позволяют получать с сервера данные, удовлетворяющим условиям запроса. Однако, Overpass API обрабатывает каждый возможный вид данных отдельно. Итого, например, чтобы выбрать все элементы с тегом amenity=school нам будет необходимо воспользоваться оператором объединения, поместив внутрь результаты выборки точек, путей и отношений с этим тегом.

На рисунке 5 показан пример запроса к Overpass API, в результате которого возвращаются все автобусные маршруты, которые попали в отображаемую на интерактивной карте в правой половине окна зону.

2.8 Выбор данных об инфраструктуре

Большинство объектов инфраструктуры имеют тег `amenities`. Объекты, являющиеся магазинами, принято помечать тегом `shop`. Объекты, относящиеся к системе здравоохранения, обычно помечают тегом `healthcare` (например, `healthcare=hospital` или `healthcare=pharmacy`), но также часто встречаются больницы, поликлиники и другие объекты здравоохранения, помеченные только как `amenity`, без тега `healthcare`.

Так же этими тегами помечены большое количество объектов, такие как скамейки, мусорные баки и велосипедные парковки, которым стоит быть обозначенными на картах, но которые явно не стоит учитывать при расчёте рейтинга района с точки зрения инфраструктуры.

В таблице ниже приведены несколько групп объектов, и сценарии использования приложения, при которых эти категории используются при расчёте рейтинга:

Категория	Атрибуты	Сценарии использования
Общественный транспорт	<code>highway=bus_stop</code>	Все
Магазины и рынки	<code>highway=bus_stop</code> <code>shop=convenience</code> <code>shop=supermarket</code> <code>amenity=marketplace</code>	Все
Аптеки	<code>amenity=pharmacy</code> <code>healthcare=pharmacy</code>	Все
Кафе и быстрое питание	<code>amenity=fast_food</code> <code>amenity=cafe</code>	Любой студент
Детские сады	<code>amenity=kindergarten</code>	Семьянин

Школы	amenity=school	Семьянин
Больницы и поликлиники	amenity=school amenity=clinic amenity=hospital amenity=doctor amenity=doctors healthcare=clinic healthcare=hospital healthcare=doctor healthcare=doctors healthcare=centre	Все
Парикмахерские	shop=hairstylist	Семьянин
Кино	shop=hairstylist	Безответственный студент
Пабы и бары	amenity=pub amenity=bar	Безответственный студент
Ночные клубы	amenity=nightclub	Безответственный студент

Так как довольно часто тег может быть установлен не на точке, а на полигоне или отношении, в данном важно извлекать не информацию о геометрии, а о центре объекта.

3 Разработка приложения

3.1 Структура разработанного приложения

Разрабатываемое приложение состоит из нескольких классов и модулей, реализующих логику приложения:

- Класс Home – класс, реализующий пользовательский интерфейс и обеспечивающий взаимодействие приложения с пользователем.
- Модуль Leaflet – используется для отображения на экране пользователя интерактивной карты и для дальнейшего вывода поверх неё тепловой карты рейтинга.
- Класс OverlayDrawer – используется для отрисовки тепловой карты, используя функции расчёта рейтинга, определённые в MapService.
- Модуль Haversine – используется для расчёта расстояний между точками на карте по их координатам.
- Класс MapService – содержит логику поиска ближайших объектов инфраструктуры и расчёта рейтинга.
- Модуль K-d tree – структура данных, представляющая собой многомерное дерево, используемое для оптимизации поиска ближайших объектов на больших выборках данных
- Модуль Overpass – используется для запроса данных из OSM посредством Overpass API

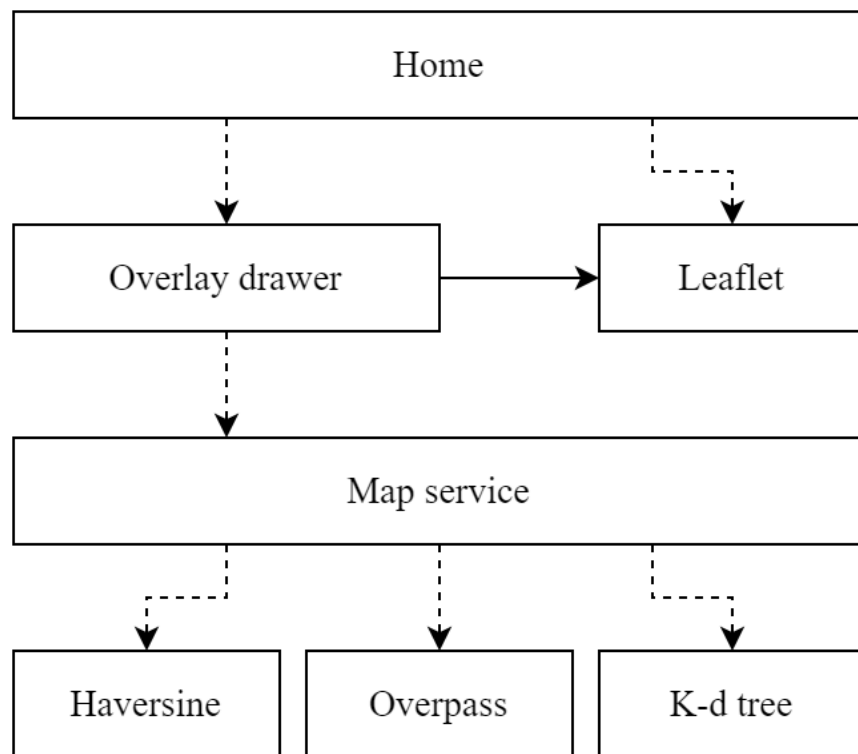


Рисунок 6 – Диаграмма классов разрабатываемого приложения

3.2 Выполнение и оптимизация запросов

Самый простой способ запросить через Overpass API данные об объектах, удовлетворяющих хотя бы одному из условий – воспользоваться оператором объединения, передав по одному набору данных для каждого отдельного условия. Однако, интерпретатор Overpass не выполняет оптимизацию таких запросов перед исполнением, в результате чего для извлечения требуется в несколько раз больше времени, чем если бы запрос использовал всего одно условие, использовавшее бы для фильтрации данных регулярные выражения.

Overpass API поддерживает два варианта фильтрации атрибутов с помощью регулярных выражений. Первый – атрибут имеет фиксированный ключ, но значение проверяется с помощью регулярного выражения. Второй – и ключ, и значения проверяются с помощью двух отдельных заданных им регулярных выражений.

Используя первое свойство, можно оптимизировать запрос на выборку данных, у которых значение одного из атрибутов может принимать несколько значений. Ниже приведён пример, показывающий суть оптимизации:

```
[out:json][bbox:{{bbox}}];  
(  node[amenity=school];  
    node[amenity=kindergarten];  
    node[shop=hairstresser];);  
out;
```

Листинг 1 – Использование оператора объединения, 3 отдельных условия

```
[out:json][bbox:{{bbox}}];  
(  node[amenity~"^(school|kindergarten)$"];  
    node[shop=hairstresser];);  
out;
```

Листинг 2 – Использование оператора объединения вместе с регулярным выражением для группировки значений, 2 условия

Учитывая, что множества возможных значений для атрибутов с ключами amenity, shop и healthcare практически не пересекаются, можно дополнительно ускорить запрос, отказавшись от использования оператора объединения и используя всего один большой запрос с двумя регулярными выражениями. В этом случае сервер вернёт некоторое небольшое количество лишней информации, но запрос всё равно будет выполнен заметно быстрее. Пример такого запроса показан ниже:

```
[out:json][timeout:25][bbox:{{bbox}}];
node[~"^(amenity|shop)$"~"^(school|kindergarten|
hairstresser)$"];
out;
```

Листинг 3 – Использование всего одного условия с регулярными выражениями

Использование этой особенности интерпретатора запросов позволило ускорить обработку запроса на сервере более чем в семь раз по сравнению с запросом, выбирающим данные без использования регулярных выражений и использующим оператор объединения.

На графиках ниже показано среднее время выполнения и объём возвращённых в ответ данных нескольких вариантов запросов, извлекающих данные об инфраструктуре в европейском городе средней величины.

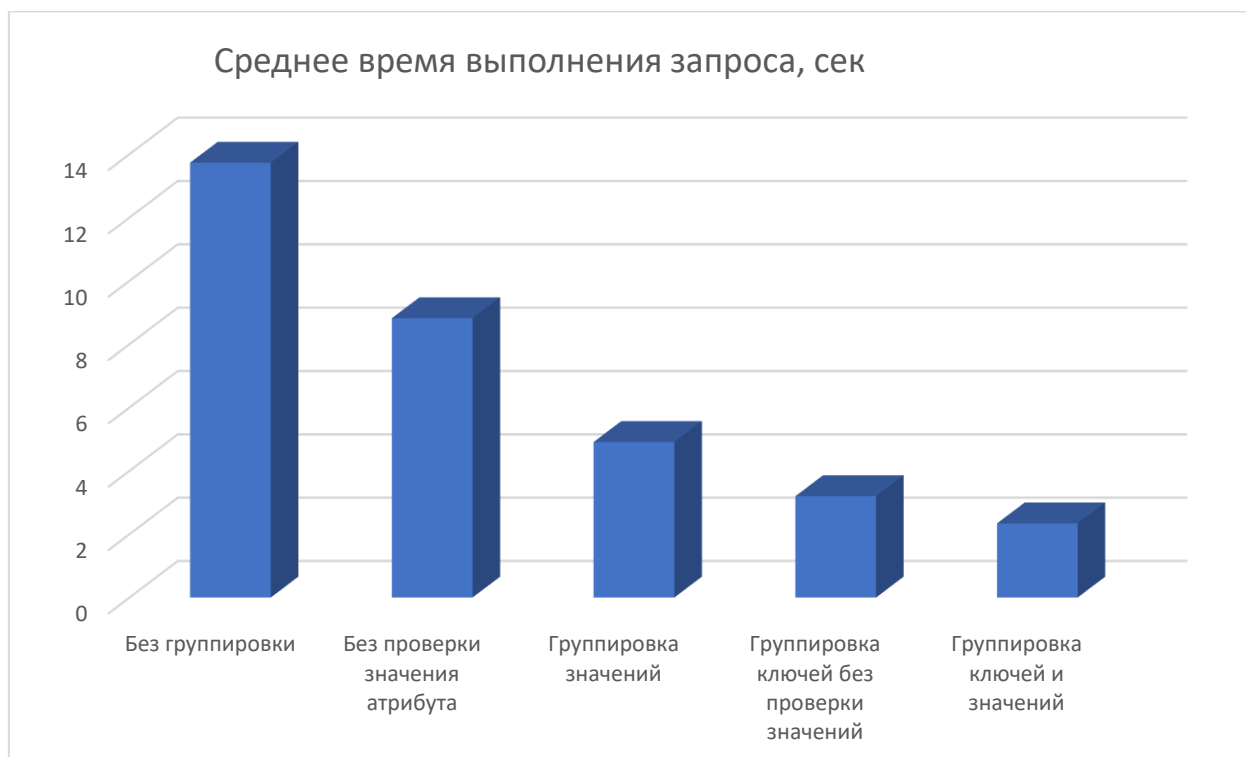


Рисунок 7 – Среднее время выполнения сформулированных разными способами запросов



Рисунок 8 – Объем передаваемых данных в результате выполнения запросов

3.3 Расчёт рейтинга

Пусть максимальный допустимый рейтинг для категории будет равен 5 и выставляется только в том случае, если до ближайшего объекта заданной категории можно дойти пешком по прямой не более чем за две с половиной минуты, считая, что средняя скорость пешехода равна 5 км/ч.

Рейтинг должен убывать достаточно быстро, чтобы на карте была видна разница между максимальным и почти максимальным рейтингами и области без инфраструктуры не были затронуты высоким рейтингом из соседних областей.

Под требования такой функции будет подходить ограниченная сверху гипербола, в числителе которой будет коэффициент, равный $2,5 * 5$, а в знаменателе – время. Максимальное значение функции рейтинга – 5.

Время пешком от одной точки до другой по прямой легко находится из расстояния между точками, а расстояние между точками может быть вычислено по формуле гаверсинусов, зная координаты точек.

Общий рейтинг в точке можно представить в виде среднего арифметического рейтингов, рассчитанных для каждой из применимых к рассматриваемому случаю категорий.

Расчёт расстояния между точками на карте по их координатам осуществляется по следующей формуле:

$$a(A, B) = \sin^2\left(\frac{lat_A - lat_B}{2}\right) + \sin^2\left(\frac{lon_A - lon_B}{2}\right) * \cos(lat_A) * \cos(lat_B),$$

$$dist(A, B) = R * 2 * \operatorname{atan2}(\sqrt{a}, \sqrt{1-a}),$$

где lat_A , lon_A , и lat_B , lon_B – широта и долгота первой и второй точек в радианах соответственно, $R = 6371000$ – радиус Земли в метрах, $\operatorname{atan2}(x, y)$ – функция угла между радиус-вектором точки с координатами x, y и положительным направлением оси абсцисс в радианах.

Время, необходимое для перемещения из точки A в точку B на карте определяется по формуле

$$time(A, B) = \frac{dist(A, B)}{V_{\text{пеш}}},$$

где $V_{\text{пеш}} = 5$ км/ч – средняя скорость перемещения пешехода по городу.

Рейтинг произвольной точки на карте рассчитывается по формуле

$$rating(A) = \frac{1}{n} \sum_{i=1}^n \max(R_{\max}, \frac{T_{\min} R_{\max}}{time(A, B_{i \text{ ближ}})}),$$

где $R_{\max} = 5$ – максимальный возможный рейтинг; $T_{\min} = 2,5$ – минимальное время в пути, соответствующее максимальному рейтингу; $B_{i \text{ ближ}}$ – ближайшая точка, относящаяся к i -той группе объектов инфраструктуры, участвующих в расчёте рейтинга.

3.4 Построение тепловой карты рейтинга

В случае, если мы хотим видеть на карте сразу много рейтингов одновременно, лучшим решением будет обозначить их цветом. Выбранная в моём случае цветовая гамма состоит из прозрачного и цветов радуги от синего

до красного, причём прозрачный соответствует минимальному (нулевому) рейтингу в точке, а красный – максимальному.

Градиент, который будет использоваться при отрисовке карты рейтинга, просчитывается заранее и сохраняется в памяти в виде картинки шириной 1 и высотой 256 пикселей. Каждое возможное значение рассчитанного рейтинга будет соответствовать одному из цветов градиента.

Отрисовка тепловой карты выполняется в несколько этапов. Сначала рисуется картинка с низким разрешением, используя квадраты со стороной 8 пикселей, а затем детализация увеличивается с помощью перерисовки, используя квадраты со стороной 2 пикселя.

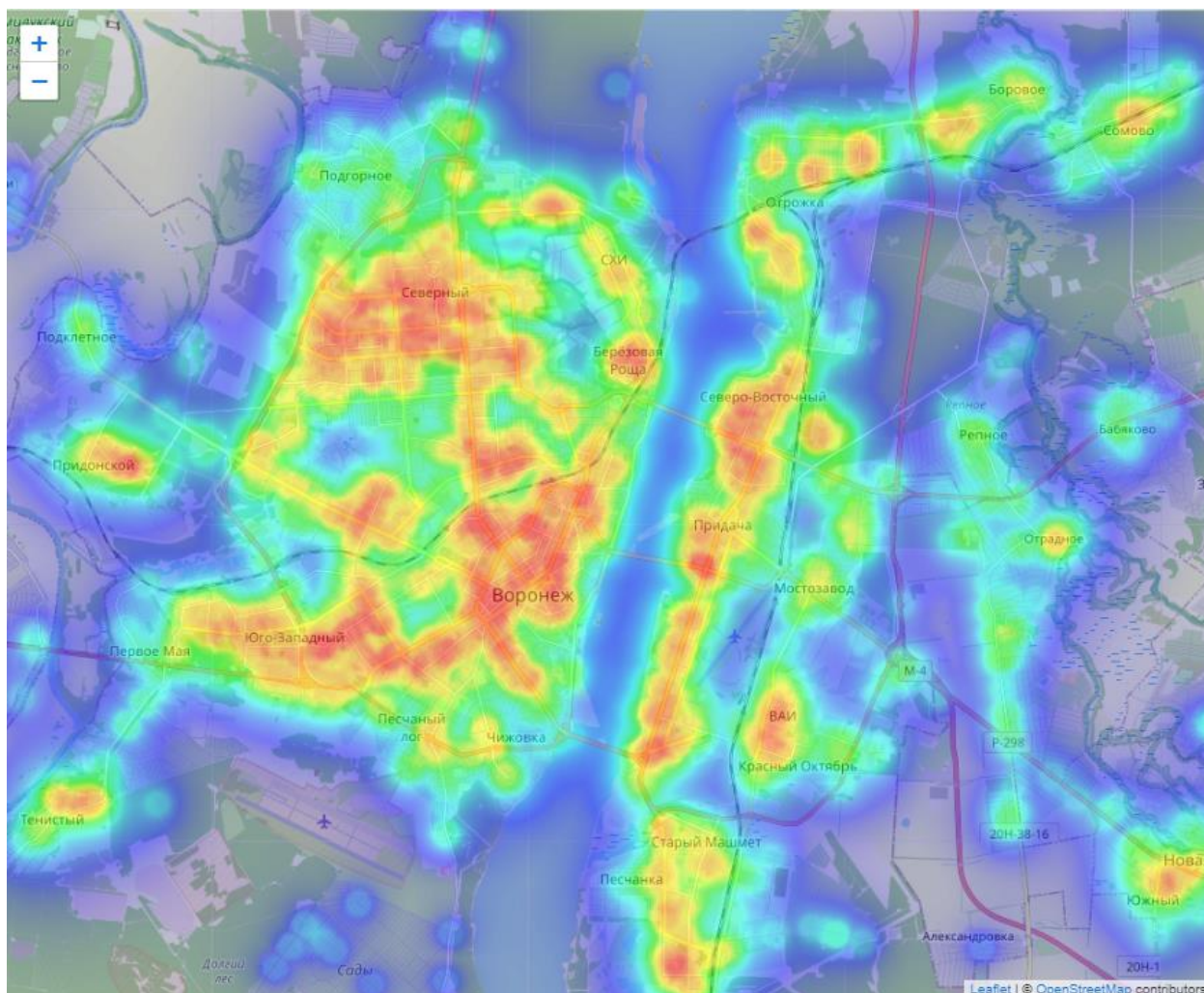


Рисунок 9 – Тепловая карта рейтингов районов с точки зрения доступности инфраструктуры

Сразу после каждого этапа отрисовки полученная картинка прикрепляется поверх отображаемой карты, замещая старую. Этот подход оказался более простым и производительным, чем использование сторонних плагинов.

При формировании тепловой карты работа ведётся не элементом Canvas, предоставляющим инструменты рисования в браузере, а напрямую с массивами байт, представляющих данную картинку в несжатом виде.

3.5 Оптимизация алгоритмов

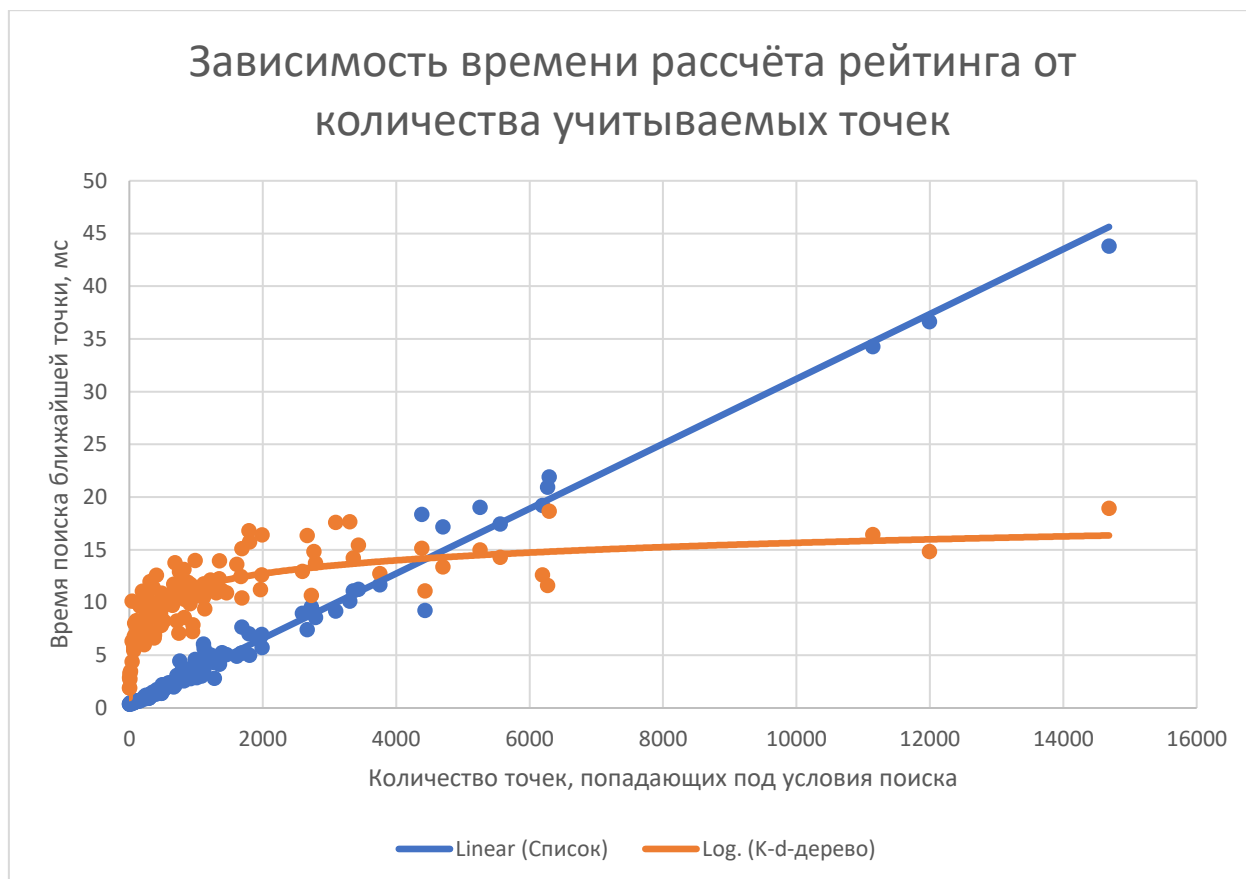
Нахождение точного расстояния между двумя точками по их координатам – довольно затратная операция, которую не следует вызывать для каждой точки для каждой проверки. Поэтому, чтобы уменьшить количество вычислений, соотношение между расстоянием в метрах и расстоянием в градусах определяется один раз в момент перед загрузкой данных. Затем все вычисления проводятся для точек с координатами в полученных «условных» метрах.

Метры «условные» так как текущее соответствие будет верным только для той широты, на которой было первоначально вычислено это соответствие. В то же время, вблизи этого места искажения будут довольно малы, чтобы эти значения можно было всё так же использовать для оценки расстояний.

При сравнении расстояний используется метрика, равная сумме квадратов разностей координат точек (расстояние в квадрате). Квадратный корень вычисляется только при необходимости посчитать точное значение уже непосредственно для расчёта рейтинга.

Другая операция, которая могла занимать продолжительное время в случае работы с большим количеством данных – это поиск ближайшего объекта, принадлежащего заданной категории. На малых объёмах данных линейный поиск справлялся за приемлемое время, но при попытке отрисовать карту рейтинга, например, для всей области этот процесс оказывался слишком медленным.

Чтобы проверить возможность ускорения отрисовки при большом объёме данных, вместе с линейными списками данные параллельно сохранялись в структуру, представляющую собой k-d-дерево. Затем была проведена отрисовка карты на разных масштабах, с разным количеством учитываемых данных. Полученные в ходе замеров производительности данные были усреднены для каждой величины рассматриваемого в момент поиска списка/дерева и использованы для построения графика.



Графики, построенные по точкам, полученным в ходе эксперимента, пересекаются при количестве элементов в списке/дереве немного большем 4000. Это означает, что начиная с этого значения, использование дерева требует меньше времени и ресурсов для поиска ближайшего элемента, чем при линейном поиске в списке.

Заключение

В ходе работы было разработано веб-приложение, позволяющее пользователям строить тепловую карту качества района для проживания, основанную на данных об инфраструктуре, получаемых из OpenStreetMap.

Конечные пользователи разработанного приложения - люди, изучающие другой город с целью выбора наилучшего возможного будущего места жительства с точки зрения инфраструктуры в связи с планируемым переездом и не обладающие специальными техническими знаниями или умениями.

Разработанное приложение соответствует следующим требованиям:

- Работает в браузерах, актуальных на момент разработки
- Использует для работы данные, получаемые из OpenStreetMap
- Позволяет пользователю открыть на экране произвольный участок карты
- Позволяет выбрать пользователю один из заранее определённых наборов инфраструктурных объектов, участвующих в расчёте рейтинга
- Рассчитывает и отображает полупрозрачную тепловую карту рейтинга для районов, отображаемых на экране пользователя, по нажатию кнопки
- Отображает численное значение рейтинга в точке на карте по нажатию на эту точку

В рамках анализа предметной области были выполнены следующие задачи:

- Изучены методы и источники данных, использованные в общедоступных исследованиях и программных продуктах
- Определены преимущества и недостатки существующих решений
- Выбран источник данных об инфраструктуре с учётом их условий использования и накладываемых лицензией ограничений, а также средства для доступа к ним

- Изучена структура данных выбранного источника
- Выбраны средства реализации и архитектура приложения

В ходе разработки приложения были реализованы следующие возможности:

- Получение приложением данных об инфраструктуре для выбранного на карте участка
- Поиск ближайшего к точке объекта определённой категории
- Расчёт расстояния и затрат времени на перемещение между точками на карте
- Расчёт рейтинга конкретной точки на карте с точки зрения инфраструктуры
- Формирование картинки тепловой карты рейтинга
- Профилирование и оптимизация используемых алгоритмов

Список использованных источников

1. Вычисление расстояния и начального азимута между двумя точками на сфере
<https://gis-lab.info/qa/great-circles.html>
2. Исследования Яндекса – Москва для жизни и для развлечений
https://yandex.ru/company/researches/2017/moscow_districts
3. Лучшие и худшие районы Воронежа
https://www.domofond.ru/statya/luchshie_i_hudshie_rayony_voronezha/7557
4. Тепловые карты – Яндекс.Недвижимость. Справка
<https://yandex.ru/support/realty/search/indicators.html>
5. OpenStreetMap Wiki – Elements
<https://wiki.openstreetmap.org/wiki/Elements>
6. OpenStreetMap Wiki – Overpass API
https://wiki.openstreetmap.org/wiki/Overpass_API
7. OpenStreetMap Wiki – Overpass QL
https://wiki.openstreetmap.org/wiki/Overpass_API/Overpass_QL
8. OpenStreetMap TagInfo
<https://taginfo.openstreetmap.org/>
9. The TypeScript Handbook
<https://www.typescriptlang.org/assets/typescript-handbook.pdf>
10. Vue.js
<https://vuejs.org/>
11. Vuetify — A Material Design Framework for Vue.js
<https://vuetifyjs.com/en/>