

РАЗРАБОТКА ИНТЕРАКТИВНОЙ ГЕОИНФОРМАЦИОННОЙ СИСТЕМЫ ДЛЯ АНАЛИЗА РАЗМЕЩЕНИЯ ОБЪЕКТОВ ГОРОДСКОЙ СФЕРЫ УСЛУГ ПО ДАННЫМ OPENSTREETMAP

*Тувалева Ю.В.,
Самсонов Т.Е.*

Резюме – Картографический сервис OpenStreetMap обладает огромным потенциалом в качестве источника открытых данных для картографирования и изучения объектов городской среды. Для анализа размещения объектов сферы услуг были изучены возможности использования методов пространственного и сетевого анализов, а также пространственной статистики, алгоритмы которых легли в основу обработки данных и вычисления показателей в разработанной геоинформационной системе. С помощью СУБД PostgreSQL, языков программирования R и Python загруженные данные, а также методы анализа размещения объектов городской среды были реализованы в интерактивной геоинформационной системе.

Ключевые слова – Городская среды, сфера услуг, пространственная статистика, сетевой анализ.

DEVELOPMENT OF AN INTERACTIVE GEOINFORMATION SYSTEM TO ANALYZE THE LOCATION OF URBAN SERVICE FACILITIES USING OPENSTREETMAP DATA

*Tuvaleva Y.V.,
Samsonov T.E.*

Abstract – OpenStreetMap is a main source of open spatial data and has a huge potential as a source for mapping and analyzing urban environment. The study applies spatial, network analysis and spatial statistics to the analysis location of urban venues. These methods were implemented in the interactive geographic system based on programming languages R and Python. Data from OpenStreetMap was structures with DBMS PostgreSQL.

Keywords – VGI, OpenStreetMap, urban venues, spatial statistics, network analysis

Введение

Город является сложным многокомпонентным социально-экономическим объектом. Важным компонентом городской среды являются объекты сферы услуг, поскольку их функционирование влияет на комфортность проживания горожан. Развитая сфера услуг позволяет жителям городов более активно взаимодействовать между собой и техногенной оболочкой города, тем самым создавая городскую среду, влияющая как на природную, так и на социальную составляющие [1].

В советский период в нашей стране в основе анализа и размещения объектов сферы услуг лежал нормативный подход, который регулировал нормы потребления опре-

деленного вида услуг. Основная задача заключалась в рациональном проектировании территориальной структуры поселения. Соответственно, анализ сферы услуг был связан с географией расселения и проводился на региональном уровне, редко затрагивая внутригородской. В настоящий момент в отечественных исследованиях большее внимание уделяется не только к пространственному распределению объектов сферы услуг, но также применяется поведенческий метод, опирающийся на реальное поведение и предпочтения потребителя. Кроме того, все чаще анализ проводится на уровне города [2].

Помимо масштаба и подхода к исследованию объектов городской сферы услуг, важным направлением является анализ структуры этих объектов и их классификации. В каждой стране существует государственный классификатор объектов сферы услуг, но зачастую, как например в России, он слишком детален для проведения комплексного анализа. Вследствие этого, различные компании, деятельность которых напрямую связана с данными типами объектов, используют свою собственную классификацию по виду предоставляемых товаров и услуг. Различные консалтинговые компании такие, как Knight Franks, JLL, а также геомаркетинговые компании в своих исследованиях и публичных отчетах в основном применяют следующую классификацию объектов по типу предоставляемых услуг и товаров: продукты, общественное питание (кафе и рестораны), связь и банки, одежда и обувь, техника и электроника, товары для дома и мебель, косметика и парфюмерия, детские товары, зоотовары, ювелирные украшения и подарки.

Стоит отметить, что геомаркетинговые и консалтинговые компании особое внимание уделяют анализу объектов уличной торговли (стрит-ритейл), поскольку они напрямую взаимодействуют с жителями городов, создавая устойчивые городские пространства.

При картографировании и анализе объектов сферы услуг в первую очередь применяются статистические методы для оценки взаимосвязи размещения с определяющими факторами. В основном анализ проводится для одного типа объекта и не освещается комплексный подход к изучению. При этом практически не уделяется внимание картографическому аспекту анализа – составляются карты фактического размещения и строят поверхности плотности.

В качестве источников информации об объектах розничной торговли и услуг используются как официальные статистические данные, так и данные краудсорсинговых систем, активно развивающиеся в последние несколько лет.

Так, на основе данных краудсорсинговой картографической платформы OSM или подобных картографических источников (включая платные картографические сервисы, например 2GIS), создаются различные справочные карты и атласы по городской сфере услуг. Например, электронный атлас г. Москвы или г. Милана являющиеся комплексным справочным ресурсом города как об инфраструктуре города, так и об объектах сферы услуг. Однако данные сервисы носят лишь справочный характер и не содержат результатов какой-либо аналитики, связанной с особенностями размещения объектов розничной торговли и услуг.

Учитывая степень развития технологий веб-картографирования и разработки различных приложения, а также появления большого объема открытых данных, возникает необходимость создания общедоступной системы с понятным интерактивным интерфейсом, позволяющим любому пользователю создавать необходимые тематические карты. Кроме того, такая система позволила реализовать комплексный подход к анализу, то есть изучение взаимосвязей между двумя и более объектами. Подобная система анализа размещения объектов городской сферы услуг может быть востребована в геомаркетинговых исследованиях, а также планирования развития городов.

Теоретико-методологическая концепция разрабатываемой системы

В основе разработки геоинформационной системы лежит трехуровневая концепция Джилса [4]. На низком уровне анализа внимание уделяется отдельным объектам и их характеристикам. Средний уровень анализа позволяет выделять в пространстве и анализировать группировки объектов (как одного так и разных типов). Наконец, высокий уровень анализа соответствует задаче оценке пространственного распределения всей совокупности объектов. Сочетание уровня анализа и характеристики объектов позволяет определять способ картографирования явления. В рамках данной работы для анализа размещения объектов концепция Джилса была расширена, обогатив ее понятиями контекста и фокуса анализа. Под контекстом анализа понимается агрегирование или перенос информации с объектов типа А на объекты типа С при том, что геометрия объектов может различаться. Для анализируемых объектов розничной торговли и услуг можно выделить 5 различных контекстов:

- Нуль-контекст – фактически отсутствие контекста, то есть каждый объект рассматривается индивидуально, без учета его взаимодействия с другими объектами и положения относительно них. Соответствует низкому уровню анализа;
- Точечный контекст – рассмотрение объектов городской среды по отношению к точечным объектам того же или другого типа. Возникает на среднем уровне анализа;
- Линейный контекст – рассмотрение объектов городской среды относительно линейных объектов, например, относительно дорожной сети или границ города. Возникает на среднем уровне анализа.;
- Площадной контекст – изучение объектов городской среды в пределах площадных объектов, например, анализ с учетом административно-территориального деления города или на основе построенной регулярной сетки. Возникает на среднем уровне анализа;
- Сплошной контекст – изучение объектов в сплошном пространстве как множества с рассеянной локализацией и пространственной вариацией характеристик. Возникает на высоком уровне анализа.

Средний уровень анализа наиболее разнообразен в плане контекстов и позволяет анализировать различные группировки объектов, причем как с себе подобными объектами, так и с объектами другого типа. Сплошной контекст может реализовываться или в картах плотности и потенциала поля распределения объектов (непрерывный сплошной), или в картах районирования и зонирования территории по особенностям размещения объектов (дискретный сплошной).

Фокус анализа определяется множеством типов анализируемых объектов (1, 2 или более), а также характером отношения (или противопоставления) этих типов друг другу. Как правило, даже при наличии большого количества переменных в один момент времени совместно анализируется не более двух из них – например, при расчете корреляции или сопоставлении изображений. Поэтому, ограничив разнообразие типов (переменных) двумя классами А и В, можно определить четыре различных фокуса анализа:

- А – анализируется размещение объектов типа А;
- В – анализируется размещение объектов типа В;
- А+В – анализируется совместное размещение объектов А и В, то есть объекты обоих типов объединяются в одно множество – например, для анализа общей плотности распределения;
- А–В – анализируются сходства и различия в размещении объектов А и В, то есть объекты типов А и В останутся внутри своих множеств и будет произведено сравнение их распределений на предмет сходств и различий по отношению к объекту типа С.

Введем также понятие инверсии контекста — это прием анализа, при котором объекты и контекст меняются местами, что позволяет обогатить анализируемые объекты характеристиками контекста, используемыми при дальнейшем анализе. В случае прямого контекста характеристики будут вычисляться и картографироваться по контекстным объектам. Например, при анализе размещения аптек в контексте улиц каждый сегмент улицы обогащается информацией о количестве или плотности размещения аптек, что затем может быть визуализировано способом линейных знаков. Однако если мы инвертируем контекст, то сможем для каждой аптеки вычислить такую характеристику как центральность ее местоположения, которая определяется для сегмента улично-дорожной сети, на котором она расположена. В этом случае показатель будет картографирован способом значков. Таким образом для прямого контекста способ изображения локализуется по контекстным объектам, для обратного контекста — по объектам картографирования. При этом фокус анализа не меняется — исследователя по прежнему интересует размещение объектов типа А. Но в одном случае оно рассмотрено с точки зрения плотности размещения по улицам, а в другом — с точки зрения центральности этого размещения — характеристики, индуцированной контекстом анализа.

На основе изученной литературы по исследованию объектов сферы услуг, а также используемой концепции контекст-фокус анализа были выбраны наиболее оптимальные методы пространственного и сетевого анализа, которые реализованы в разработанной системе.

Первая группа методов включает пространственный анализ и оверлей с подсчетом количества объектов по регулярной сетки или по району [3], которая также позволяет строить некоторые статистические функции оценки типа распределения, например, индекс Моришита, по диаграмме индекса можно определить, имеют ли исследуемые объекты тенденцию к группировке или размещаются независимо друг от друга.

Следующая группа методов для оценки типа распределения включает эмпирические функции распределения расстояний. Их существует несколько типов, но стоит отметить, что они все позволяют сравнить фактическое распределение со случайным и сделать вывод о том, как распределены эти точечные объекты в пространстве: имеют ли они какую-либо тенденцию к группировке, отталкиваются или размещаются независимо друг от друга [5]. На основе расстояний до ближайшего объекта строятся матрицы расстояний, позволяющие выяснить, насколько близко расположены объекты одного типа к объектам того же типа, чем к объектам другого типа и наоборот. На основе сравнения расстояний до ближайшего объекта того же типа и расстояний до ближайшего объекта другого типа строятся матрицы кластеризации, которые позволяют выявить объекты, которые в среднем размещаются ближе не к подобному объекту, а к другому типу, что дает возможность предположить, что между их размещением имеется взаимосвязь. Ячейки такой матрицы показываются тип кластеризации — внешняя или внутренняя. Матрица не является симметричной и по ней можно проследить некоторые закономерности в размещении объектов (рис.1).

Для оценки плотности распределения в пределах исследуемой территории на высоком уровне анализа применяется метод ядерного сглаживания. В работе используется метод поиска оптимального размера окна, который позволяет масштабировать его в зависимости от фактического размещения объектов, что позволяет учесть пространственную неоднородность размещения множества объектов.

Следующая группа используемых методов относится к сетевому анализу. Структура города определяется рисунком улично-дорожной сети, что также влияет на размещение объектов сферы услуг, поэтому при изучении их распределения в городе важно использовать показатели, описывающие структуру и параметры элементов дорожной сети [7].

Среди этих метрик выделяется важная группа показателей центральности – срединность и промежуточность. Срединность позволяет оценить насколько узел расположен близко ко всем другим узлам через кратчайшее расстояние в сети. Промежуточность позволяет выделить узлы улично-дорожной сети, которые связывают кратчайшим путем между собой большое число узлов в анализируемой сети [6].

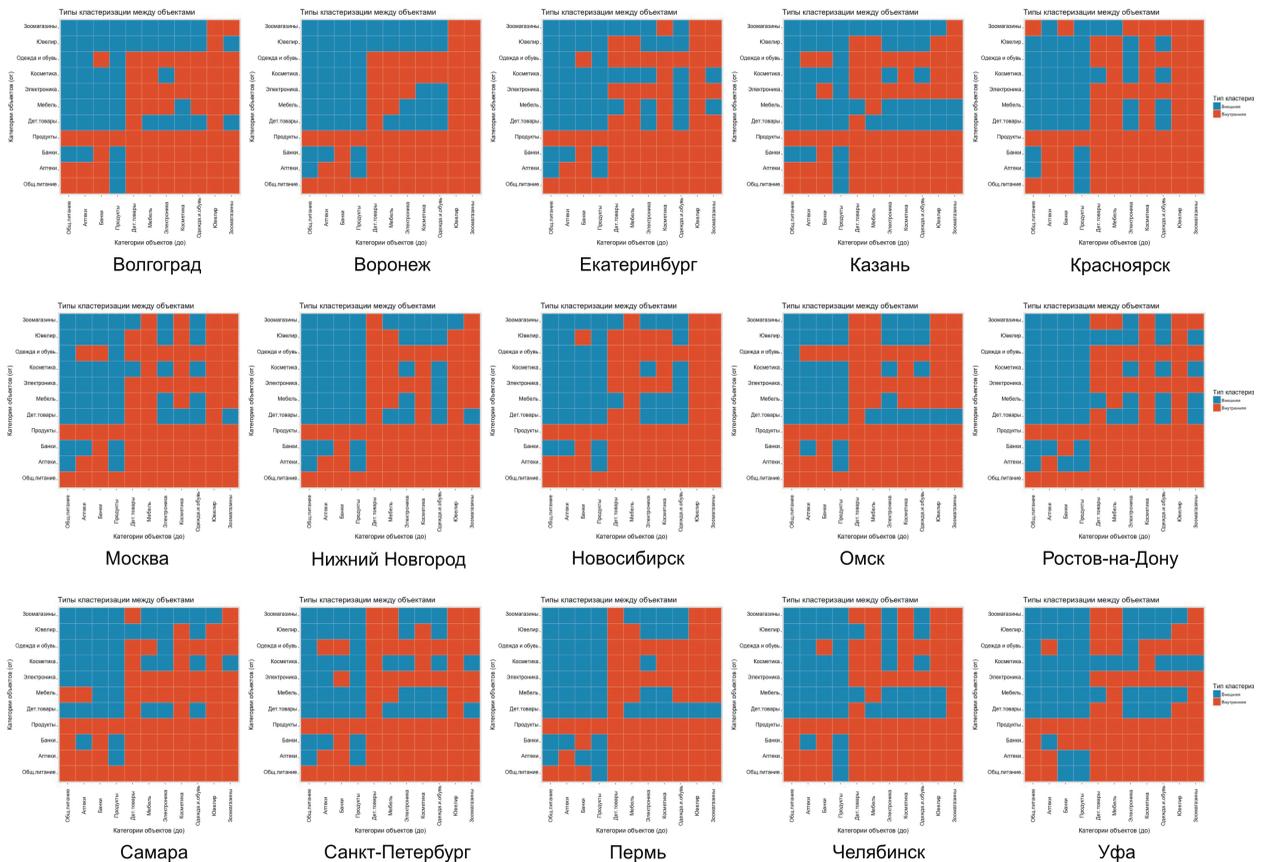


Рис. 1. Матрицы кластеризации

Для отображения характеристик, вычисленных с помощью статистических функции и других методов количественного анализа на различных уровнях и контекстах анализа могут составляться инвентаризационные, аналитические и синтетические карты.

Визуализация на нуль-контексте позволяет в целом оценить размещение точек, покрытие набором данных исследуемой территории. Для этих задач составляются инвентаризационные карты для отображения фактического положения объектов. Используется значковый способ изображения.

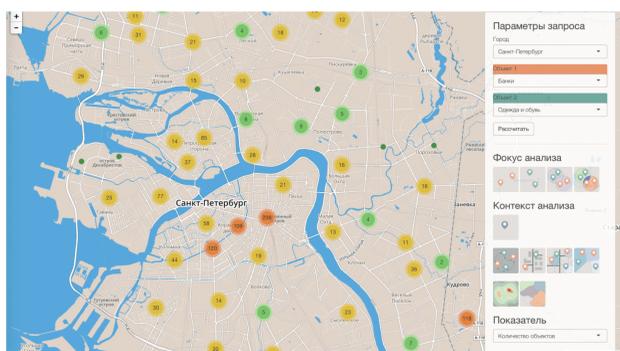
При точечном контексте анализа необходимо определить неравномерность в размещении и выделить кластеры. Для этого также составляются инвентаризационные карты. Для отображения на таких картах используется способ значков с изменением их цвета или размера.

Картографическое изображение на линейном и площадном контекстах позволяет проанализировать неравномерность размещения точечных объектов относительно дорожной сети, в пределах муниципального района или регулярной сетки. Для данного контекста создаются аналитические карты плотности размещения. В зависимости от контекста анализа используются разные способы изображения. Так, для линейного контекста плотность размещения объектов будет отображаться с помощью линейных знаков, цвет которых будет соответствовать количеству объектов на единицу длины дорожной сети. Для площадного контекста применяются картограммы (при использова-

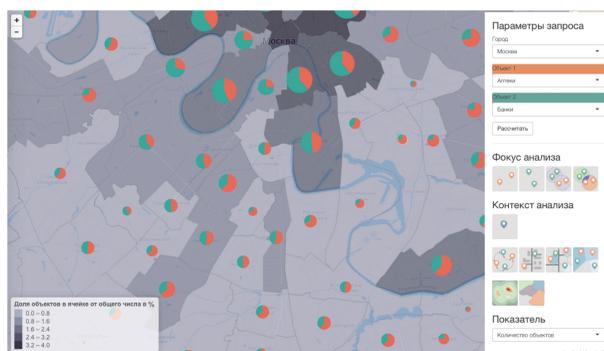
нии абсолютных значений) и картодиаграммы (при отображении количества объектов на единицу какого-либо показателя или доли объектов в данном в районе от их общего числа в пределах исследуемой территории). В этом случае может использоваться моно-тонная цветовая шкала, интенсивность которой будет возрастать с увеличением числа объектов.

Для сплошного контекста анализа составляются аналитические карты плотности распределения точечных объектов, необходимые для оценки размещения объектов в пределах города. Для этого строятся поверхности плотности распределения. При создании таких карт используются изолинии с послышной окраской, отражающие плотность размещения (рис.2).

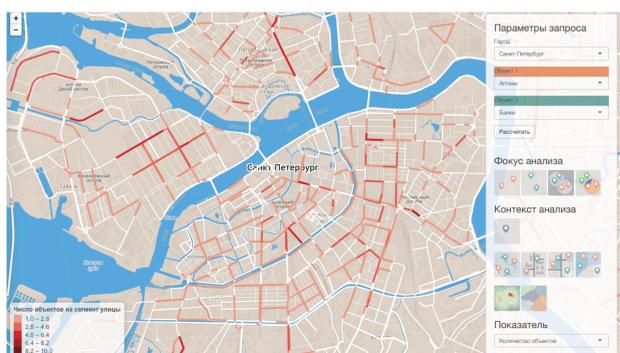
Точечный контекст



Площадный контекст



Линейный контекст



Сплошной контекст

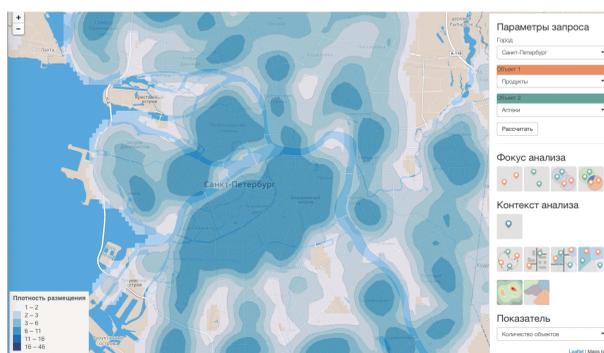


Рис. 2. Картографическое представление контекстов анализа

Архитектура системы

При разработке системы на основе предложенной концепции анализа был определен ряд требований, которым должна отвечать система. Одними из важных требований являются:

- Разработка системы в виде веб-приложения, доступного через Интернет;
- Картографическая компонента должна быть свободно управляемой, то есть наличие возможности совмещения тематических слоев, полученных по результатам анализа;
- Открытость, а именно использование открытых источников данных, содержащих информацию об объектах розничной торговли и услуг и применение в разработке открытых ПО, обладающих широким инструментарием.

В качестве используемых открытых программных средств (согласно одному из требований к разрабатываемой системе) использовались язык программирования R,

а также различные дополнительные модули, значительно расширяющие его функционал, а именно, модули Shiny и Overpass и JavaScript библиотека Leaflet.

В качестве источника информации об объектах городской сферы услуг использовались данные открытого картографического сервиса OpenStreetMap.

Архитектура системы состоит из модулей загрузки и подготовки данных, серверной обработки и взаимодействия с пользователем.

Модуль загрузки и подготовки данных является начальной ступенью во всей архитектуре, поскольку осуществляет выгрузку и предварительную обработку данных OSM. Загрузка данных из БД OSM производилась с помощью языка запросов Overpass API. С помощью PostgreSQL осуществлялось хранение и управление загруженными данными.

Модуль серверной обработки включает два важных блока обработки и вычисления. Первый связан с непосредственными вычислениями статистических функций, а также показателей для картографирования. Для этого применяются различные библиотеки языка R, например, rgeos, spatstat и maptools. Второй блок отвечает за визуализацию полученных расчетов с применением таких библиотек, как leaflet и plotly.

Модуль взаимодействия с пользователем, предоставляющий доступ к параметрам управления аналитическими и картографическими функциями системы, в частности к визуализации данных и построению карт. Основным функционалом этого модуля создается при помощи библиотеки shiny.

Подобная архитектура системы (рис.3) позволяет в автоматизированном режиме производить загрузку данных, а также построение карт после того, как пользователь выберет город и типы объектов для анализа, то есть определит параметры для вычисления (данные, контекст и фокус анализа). Изначально по умолчанию отображается карта с фактическим размещением двух типов объектов в городе. Далее при включении и отключении тематических слоев регулируется их отображение, то есть они не удаляются полностью, что оптимизирует работу системы, избегая дополнительную создание карт и построения графиков.

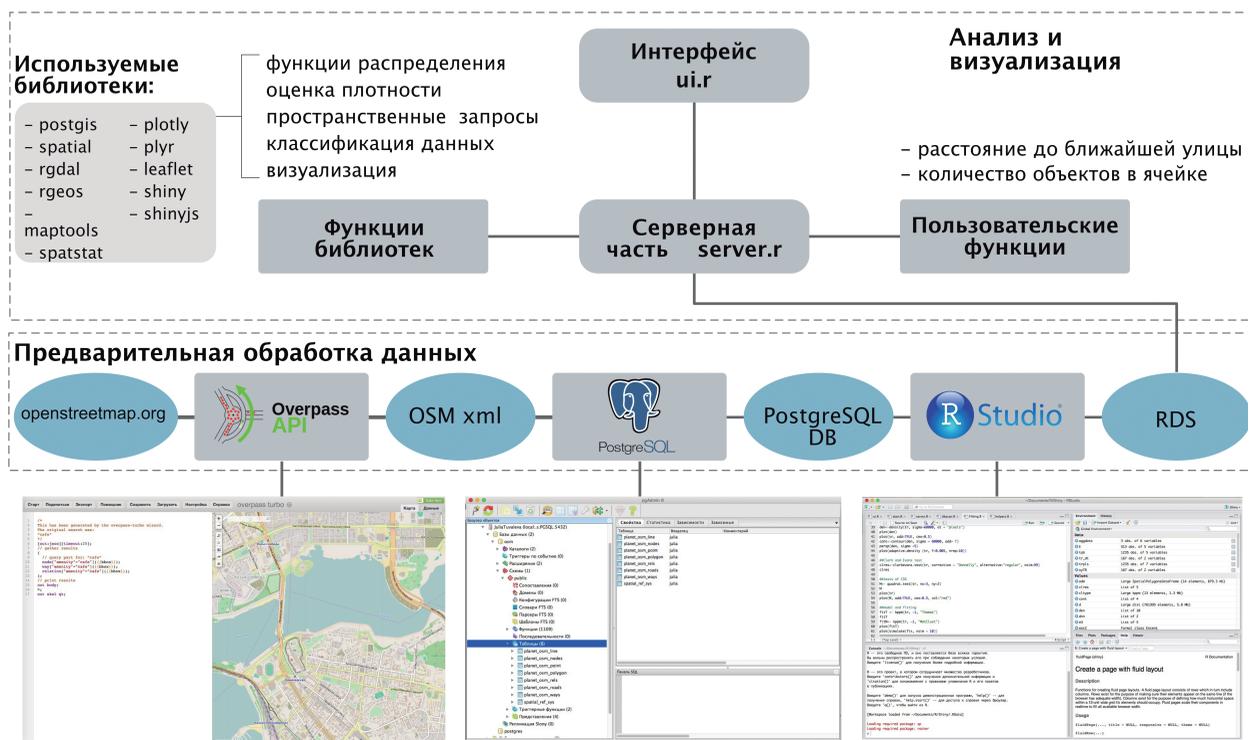
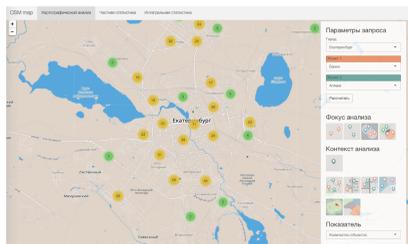


Рис. 3. Архитектура системы

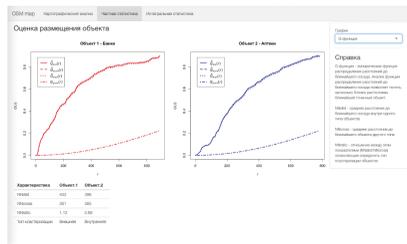
Интерфейс системы включает три раздела – картографический анализ, частная статистика и интегральная статистика (рис.4).

В основной части картографического раздела размещается интерактивная карты, а справа от нее расположена панель управления, позволяющая выбрать города, объекты для анализа, а также контекст и фокус.

Карты



Частная статистика



Интегральная статистика



Рис. 4. Интерфейс разделов системы

Раздел частной статистики позволяет определить тип пространственного распределения на основе функция распределения расстояния, тип кластеризации, а также их отличие для двух видов услуг.

Раздел интегральной статистики дает обобщенную оценку в размещении всех видов услуг и отображает матрицу кластеризации всех объектов в выбранном городе.

Результаты

Разработанная система была использована для анализа объектов розничной торговли и услуг на 15 городах-миллионерах России. Наиболее интересные взаимосвязи были получены при анализ матриц кластеризации, а также показателей центральности.

Для многих городов, например, для Москвы, Нижнего Новгорода, Санкт-Петербурга, Екатеринбурга характерны устойчивые высокие значения срединности, ими обладают такие типы услуг, как банки, заведения общественного питания, магазины одежды и обуви, а также ювелирные магазины. Устойчивые низкие значения характерны для зоомагазинов. Также выделяются категории услуг, значение срединности для которых меняются по городам. Например, в Санкт-Петербурге и Екатеринбурге, магазины детских товаров имеют низкие значения срединности, а в Новосибирске и Омске – высокие (рис. 5).

Устойчивые высокие значения показателя промежуточности характерны для магазинов электроники и общественного питания. Также есть объекты, значения промежуточности для которых меняются по городам. Так, в Казани, Ростове-на-Дону и Екатеринбурге мебельные ювелирные магазины имеют высокие значения показателя, тогда как в Красноярске, Самаре, Уфе и Омске – низкие значения. В некоторых городах, например, в Москве, Челябинске и Перми низкие значения промежуточности имеют аптеки и продуктовые магазины (рис. 6).

В анализируемых городах преобладает односторонняя связь между типами объектов. Однако есть города с наибольшей (Екатеринбург, Пермь, Нижний Новгород) и с наименьшей (Москва, Санкт-Петербург) долей односторонних связей, что может быть связано как с особенностями пространственной структурой города, так и наличием других важных объектов городской инфраструктуры.

Несмотря на то, что разработанная система отвечает поставленным требованиям и выполняет поставленные перед ней задачи, она требует дальнейших доработок в расширении алгоритмов для анализа объектов на всех контекстах анализа. Также же-

лательное дальнейшее усовершенствование модуля взаимодействия с пользователем, увеличение числа параметров, управляемых с помощью интерфейса. Дополнительно стоит проработать методику комплексного анализа городов, например более детальный анализ связей между объектами городской сферы услуг, а также их влияния на размещение других типов объектов.

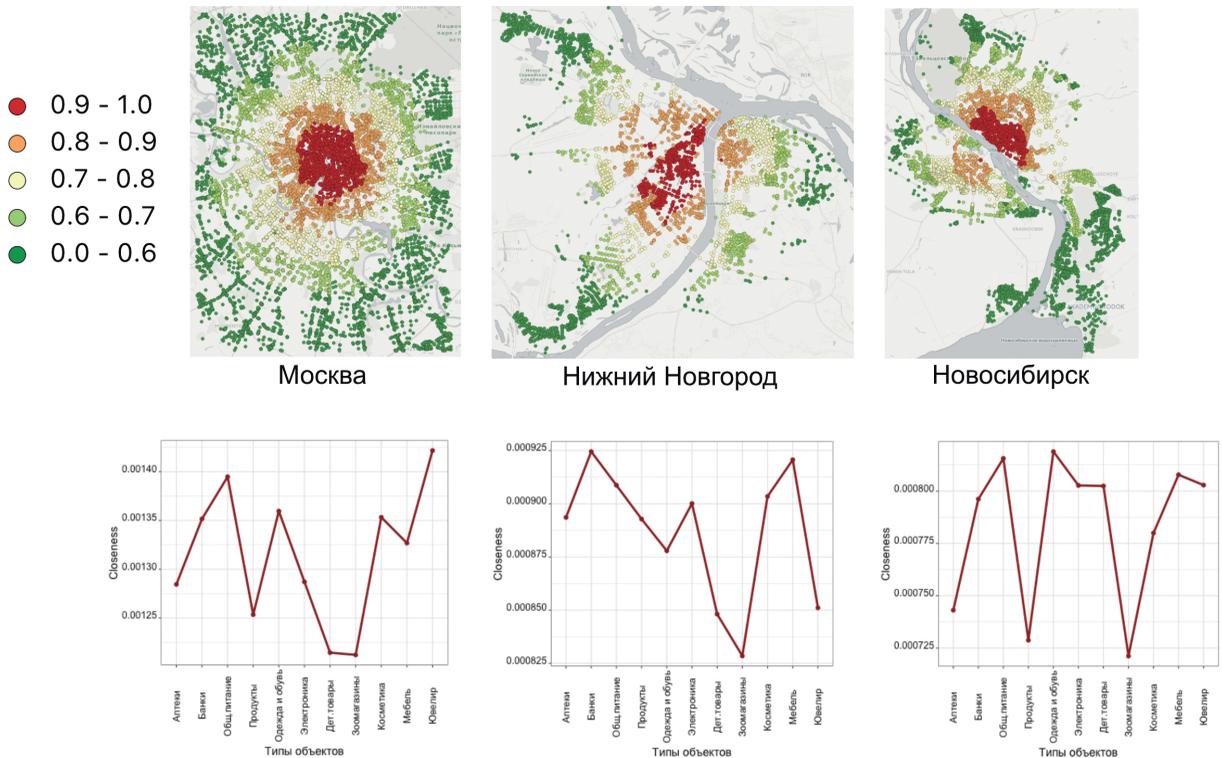


Рис. 5. Срединность узлов улично-дорожной сети и ее влияние на размещение объектов

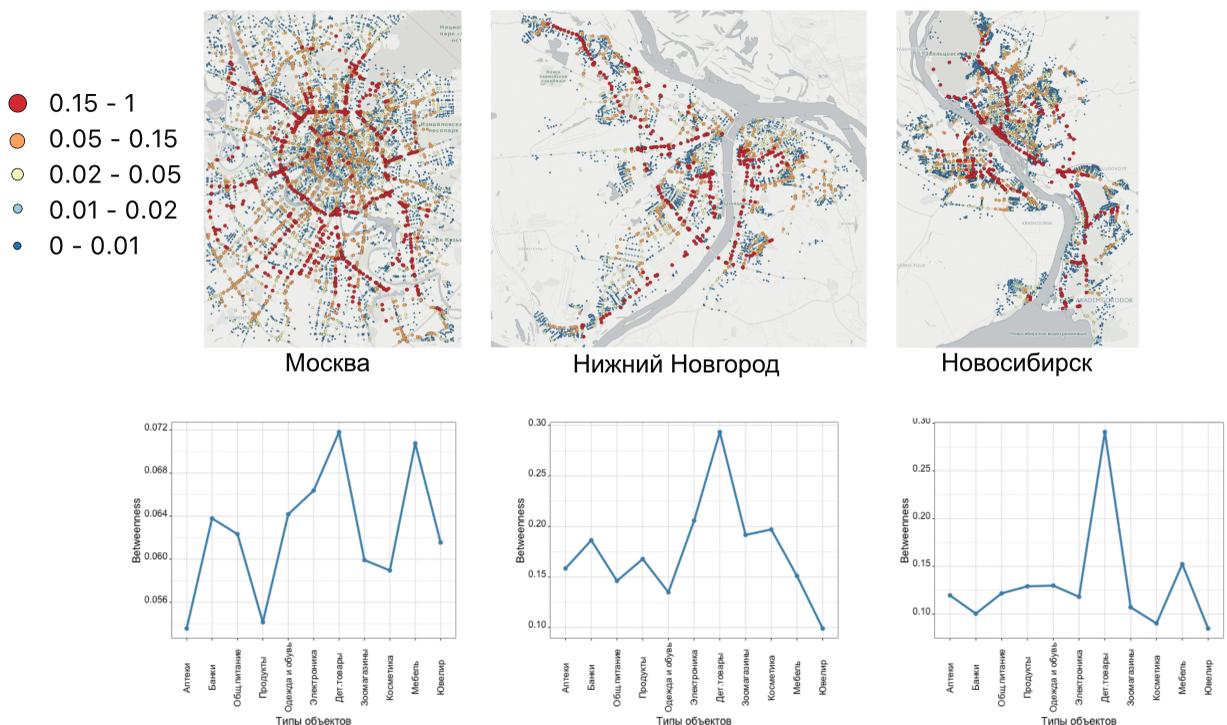


Рис. 6. Промежуточность узлов улично-дорожной сети и ее влияние на размещение объектов

Список литературы

- [1] *Глазычев В. Л., Егоров М. М., Ильина Т. В.* Городская среда. Технология развития Настольная книга М.: Изд-во «Ладья». — 1995. — 240 с.
- [2] *Имангалин А. Ф., Ушакова Л. А.* Комплексная оценка расположения объектов аптечной сети с использованием геоинформационных технологий (на примере г.Уфы) // //Материалы Международной конференции ИнтерКарто/ИнтерГИС 17: Устойчивое развитие территории: теория ГИС и практический опыт. (Белокуриха, Денпасар, 14–19 декабря 2011 г.). — г. Барнаул, 2011. — С. 259–264.
- [3] *Лурье И. К.* Геоинформационное картографирование. — 2 изд. — Москва : КДУ, 2010. — 424 с.
- [4] Geels G. H. Een model voor de keuze van vlaksymbolen I // Kartografisch Tijdschrift.— 1987. — Vol. 13, no. 4. — P. 22–27.
- [5] *Baddeley A., Rubak E. and Turner R.* Spatial Point Patterns: Methodology and Applications with R. Chapman and Hall/CRC Press, 2015.
- [6] *Newman M. E. J.* Networks An introduction. — Oxford University Press, 2010. — 1042 p. — ISBN: 978–0–19–920665–0.
- [7] Street centrality and densities of retail and services in Bologna, Italy / Sergio Porta, Vito Latora, Fahui Wang et al. // Environment and Planning B: Planning and Design. — 2009. — Vol. 36, no. 3. — P. 450–465.

References

- [1] *Glazychev V. L., Egorov M. M., Il'ina T. V.* Gorodskayasreda. Tekhnologiya razvitiya Nastol'naya kniga M.: Izd-vo «Lad'ya». — 1995. — 240 pp.
- [2] *Imangalin A. F., Ushakova L. A.* Kompleksnaya ocenka raspolozheniya ob"ektov aptechnoj seti s ispol'zovaniem geoinformacionnyh tekhnologii (na primere g.Ufy) // //Materialy Mezhdunarodnoj konferencii InterKarto/InterGIS 17: Ustojchivoe razvitie territorii: teoriya GIS i prakticheskij opyt. (Belokuriha, Denpasar, 14–19 dekabrya 2011 g.). — g. Barnaul, 2011. — p. 259–264.
- [3] *Lur'e I. K.* Geoinformacionnoekartografirovanie. — 2 izd. — Moskva : KDU, 2010. — 424 pp.
- [4] Geels G. H. Een model voor de keuze van vlaksymbolen I // Kartografisch Tijdschrift.— 1987. — Vol. 13, no. 4. — P. 22–27.
- [5] *Baddeley A., Rubak E. and Turner R.* Spatial Point Patterns: Methodology and Applications with R. Chapman and Hall/CRC Press, 2015.
- [6] *Newman M. E. J.* Networks An introduction. — Oxford University Press, 2010. — 1042 p. — ISBN: 978-0–19–920665–0.
- [7] Street centrality and densities of retail and services in Bologna, Italy / Sergio Porta, Vito Latora, Fahui Wang et al. // Environment and Planning B: Planning and Design. — 2009. — Vol. 36, no. 3. — P. 450–465.

Авторы

Тувалева Ю. В. — магистр географии по специальности картография и геоинформатика.

Самсонов Т. Е. — ведущий научный сотрудник кафедры картографии и геоинформатики, кандидат географических наук.

Authors

Tuvaleva Y. V. — Master of Geography in the specialty of cartography and geoinformatics.

Samsonov T. E. — Leading Scientist of the Department of Cartography and Geoinformatics, Candidate of Geographical Sciences.