

МЕТОДЫ КЛАСТЕРНОГО АНАЛИЗА ДОРОЖНЫХ СЕТЕЙ ДЛЯ ВЫЯВЛЕНИЯ УЗКИХ МЕСТ И ОПТИМИЗАЦИИ ДВИЖЕНИЯ

DOI: 10.36724/2072-8735-2025-19-7-34-40

Manuscript received 30 April 2025;
Accepted 05 June 2025

Поваляев Никита Денисович,
Санкт-Петербургский государственный университет,
г. Санкт-Петербург, Россия, nikita03565@gmail.com

Исследование выполнено за счет гранта Российского
научного фонда № 24-79-00192 "Цифровой двойник
городского пассажирского транспорта общего пользования",
<https://rscf.ru/project/24-79-00192/>

Крылатов Александр Юрьевич,
Санкт-Петербургский государственный университет,
г. Санкт-Петербург, Россия, a.krylatov@spbu.ru

Шаталова Наталья Викторовна,
Институт проблем транспорта им. Н.С. Соломенко
РАН, г. Санкт-Петербург, Россия, shatillen@mail.ru

Ключевые слова: дорожные сети, кластеризация,
сетевой анализ, обработка пространственных
данных, управление транспортными потоками,
городское планирование

В статье исследуется пространственная организация дорожных сетей с использованием современных методов сетевого анализа и кластеризации. Цель работы - выявление устойчивых функциональных структур транспортной сети и анализ их изменений под влиянием различных уровней загруженности. Рассматривается применение современных алгоритмов сетевого анализа и кластеризации на улично-дорожных сетях крупных городов. Особое внимание уделено анализу динамики кластеров в условиях разной интенсивности дорожного движения. Эксперименты проведены на реальных данных транспортных потоков и включают оценку изменений топологических характеристик сети при варьирующихся уровнях загруженности. Результаты показали, что предложенные подходы позволяют эффективно выделять устойчивые функциональные области сети и фиксировать их трансформации при изменении условий движения. Выявлены закономерности перераспределения потоков в периоды пиковых нагрузок и формирования потенциальных узких мест. Пространственные паттерны кластеров демонстрируют корреляцию с характером загруженности сети и топографией города. Практическое значение работы заключается в возможности использования предложенной методологии для поддержки принятия решений в области транспортного планирования и управления дорожными потоками. Полученные результаты могут быть использованы для оптимизации маршрутизации, повышения устойчивости транспортной инфраструктуры и предотвращения перегрузок на ключевых участках сети. Исследование подчеркивает ценность применения сетевых методов анализа и кластеризации для изучения динамики транспортных систем и формирования стратегий эффективного управления дорожным движением в условиях современного города.

Информация об авторах:

Поваляев Никита Денисович, Санкт-Петербургский государственный университет, аспирант кафедры математической теории экономических решений, г. Санкт-Петербург, Россия. ORCID: 0009-0003-4926-9086

Крылатов Александр Юрьевич, Санкт-Петербургский государственный университет, доктор физико-математических наук, профессор, заведующий кафедрой математической теории экономических решений, г. Санкт-Петербург, Россия. ORCID: 0000-0002-6634-1313

Шаталова Наталья Викторовна, ФГБУН "Институт проблем транспорта им. Н.С. Соломенко РАН", ведущий научный сотрудник лаборатории проблем организации транспортных систем, кандидат технических наук, г. Санкт-Петербург, Россия.

Для цитирования:

Поваляев Н.Д., Крылатов А.Ю., Шаталов Н.В. Методы кластерного анализа дорожных сетей для выявления узких мест и оптимизации движения // Т-Comm: Телекоммуникации и транспорт. 2025. Том 19. №7. С. 34-40.

For citation:

N.D. Povaliaev, A.Yu. Krylatov, N.V. Shatalova, "Methods of cluster analysis of road networks for bottleneck detection and traffic optimization," T-Comm, 2025, vol. 19, no.7, pp. 34-40. (in Russian)

Введение

Понимание структуры и динамики городских дорожных сетей имеет фундаментальное значение для решения задач управления дорожным движением, снижения уровня загруженности дорог и городского планирования. Городские дорожные сети – это сложные системы, в которых время в пути, связность и уровень перегруженности существенно меняются во времени. Исследование этих сетей с точки зрения теории графов и кластерного анализа как части более широкого и активно развивающегося направления исследований по применению машинного обучения к транспортно-логистическим системам [1-3] может дать ценные сведения об их структуре.

Методы обнаружения сообществ, которые выявляют группы тесно связанных узлов в сети, показали значительные перспективы в определении структурных паттернов в различных областях, включая социальные сети и биологические системы [4]. Алгоритм Лейдена [5] считается одним из лучших для обнаружения сообществ благодаря своей надежности и эффективности [6]. Хотя алгоритмы обнаружения сообществ широко изучены в непространственных сетях, их применение к пространственно-ограниченным сетям, таким как дорожные системы, остается относительно малоизученным.

Улично-дорожные сети (УДС) обычно представлены в виде взвешенных графов, где узлы соответствуют перекресткам, а ребра – сегментам дорог с весами, обозначающими время в пути, расстояние или другие показатели. Эти веса могут быть динамичными и зависеть от таких факторов, как интенсивность движения и пропускная способность дорог.

В предыдущих исследованиях структурные свойства дорожных сетей анализировались с помощью мер центральности, таких как степень посредничества или степень близости. При анализе графов УДС с точки зрения обнаружения сообществ, обнаруженные сообщества будут образовывать так называемые территориальные единицы высокой связности, рассмотрение которых может дать новую ценную информацию. К тому же большинство существующих подходов ориентированы на статичные показатели, которые не учитывают влияние дорожного движения и пробок.

Взаимодействие между перегруженностью и связностью сети имеет важное значение для понимания того, как функционируют городские системы в различных условиях движения. Пробки не только увеличивают время в пути, но и изменяют топологию дорожных сетей, изолируя определенные регионы или создавая узкие места.

В данном исследовании предлагается использовать алгоритм кластеризации Лейдена для анализа городских дорожных сетей, используя время в пути в качестве веса ребра. Исследование направлено на изучение того, как меняется структура сообществ в дорожных сетях при различных сценариях перегрузки — от свободного движения до сильной перегруженности сети.

Обзор литературы

Множество сетей, изучаемых в науке, включая социальные, компьютерные и биологические, естественным образом делятся на сообщества. Задача выявления и описания этой структуры сообществ остаётся одной из центральных проблем в изучении систем, организованных в виде сетей.

Ранние подходы к обнаружению сообществ в сетях фокусировались на выявлении структуры путём анализа связей между вершинами. Была предложена категория алгоритмов для кластеризации сетей на основе итеративного вычисления мер центральности и удаления ребер [4]. Несмотря на успех в идентификации сообществ, алгоритм сталкивался с серьёзными вычислительными ограничениями, которые были впоследствии решены сначала путем оптимизаций исходного алгоритма [7], а затем путем рассмотрения задачи кластеризации как задачи оптимизации, направленной на максимизацию модулярности с помощью матричных вычислений, что заметно улучшило производительность и качество результатов по сравнению с предыдущими подходами.

Путь от первых алгоритмов до современных решений, таких как алгоритмы Лувейна [8] и Лейдена [5], стал заметным шагом вперёд в изучении сетевых структур. Алгоритм Лувейна, впервые предложенный для оптимизации модулярности, стал одним из самых популярных методов благодаря своей вычислительной эффективности и способности работать с большими сетями. Однако ограничения, связанные с возможностью формирования вырожденных сообществ, стимулировали дальнейшие разработки, что привело к появлению алгоритма Лейдена. Этот алгоритм обеспечил более высокое качество разбиений благодаря дополнительным этапам уточнения кластеров и гарантии их внутренней связности, делая его особенно подходящим для анализа сложных транспортных сетей.

В последние годы различные методы анализа данных применяются для улучшения широкого набора аспектов жизни города – от оптимизации решения задач логистики [9, 10] до повышения качества жизни [11] и безопасности [12] горожан. В частности, методы кластеризации активно применяются для анализа дорожных сетей [13], что позволяет выявлять функциональные регионы, анализировать изменения транспортных потоков и улучшать принятие решений. Применение алгоритмов обнаружения сообществ открыло новые возможности для понимания структуры и динамики транспортных сетей. Например, были проведены исследования использования алгоритмов обнаружения сообществ для топологического представления дорожной сети на основе наблюдаемых маршрутов поездок [14, 15].

Это позволило разделить город на функциональные регионы, основанные исключительно на их использовании, а не на географических или административных границах. При этом структуры сообществ и социально-экономические характеристики не всегда совпадают, что подчеркивает уникальность графовых подходов в анализе транспортных систем. Помимо этого, подобные исследования позволяют выявлять паттерны в распределении узких мест в транспортных системах во времени и пространстве, что может быть использовано для предсказания и управления пиковыми нагрузками [16, 17].

Методы кластеризации также применялись к динамическим дорожным сетям [18], однако в данной работе были использованы случайно сгенерированные данные о загруженности дорог. Это ограничение подчеркивает важность использования реальных, изменяющихся во времени данных для лучшего понимания поведения дорожных сетей в различных условиях движения.

В последние годы значительно возрос интерес к применению пространственных алгоритмов кластеризации. Так, разработаны усовершенствованные пространственно-ориентированные версии алгоритмов Лувейна и Лейдена [19, 20], которые обеспечивают более точное выделение функциональных регионов, таких как зоны обслуживания пациентов медицинскими организациями.

Методология

Подготовка данных

Улично-дорожная сеть (УДС) представляется в виде взвешенного направленного графа, где вершины соответствуют некоторым точкам на дороге, а дуги – сегментам дорог, соединяющим эти точки. В качестве весов в графе рассматривается время движения по сегменту дороги, обозначенному соответствующей дугой. В качестве основного источника данных об УДС используется OpenStreetMap (OSM) [21], а получение и непосредственная обработка данных выполняются с помощью библиотеки OSMnx [22].

Данные о загруженности дорог получены с помощью картографических сервисов [23], таких как Google Maps и Яндекс.Карты. Каждый цвет сегмента дороги на карте соответствует определенному уровню загруженности: от зеленого цвета (слабая загруженность) до темно-красного (очень высокая загруженность). Пример представлен на рисунке 1 Рис. .



Рис. 1. Пример графа

На основе уровня загруженности можно оценить плотность потока, а с помощью потока – пересчитать скорость движения [24] по формуле:

$$v = v_f \exp\left(-\frac{1}{2}\left(\frac{k}{k_m}\right)^2\right)$$

где v – скорость движения, v_f – скорость свободного движения, k – текущая плотность потока, k_m – критическая плотность потока.

Большинство объектов, описывающих дороги в OSM имеют атрибут максимальной скорости либо в виде числа, либо с помощью переменной вида RU:urban. Для тех дорог, где нет данного атрибута, ему будет присвоено значение по умолчанию на основе обязательного атрибута highway, показывающего тип дороги. Например, для значения highway «secondary» – 60 км/ч, а для highway «residential» – 20 км/ч. Таким образом гарантируется, что скорости и время движения везде корректны и близки к истинным.

Кластеризация

В исследовании используется алгоритм кластеризации Лейдена. Для оценки применения алгоритма к дорожным сетям были проведены эксперименты по кластеризации при четырех различных условиях движения:

- Свободный поток: используется время свободного движения.
- Низкая загруженность: используется время в пути при уровне загруженности, типичном для раннего утра или позднего вечера.
- Умеренная загруженность: используется время в пути при уровне загруженности, типичном для середины дня.
- Сильная загруженность: используется время движения в условиях сильной загруженности для моделирования сценария движения в час пик.

Основные детали эксперимента

Алгоритм максимизирует метрику, которая представлена формулой:

$$Q = \sum_{ij} \left(A_{ij} - \gamma \frac{k_i^{out} k_j^{in}}{m} \right) \delta(\sigma_i, \sigma_j)$$

где A – матрица смежности, k_i^{in} и k_i^{out} обозначают соответственно входящую и исходящую степени вершины i , а A_{ij} обозначает дугу от вершины i до вершины j , m – суммарное количество дуг (или сумма весов в случае взвешенного графа), σ_i обозначает сообщество вершины i , а $\delta(\sigma_i, \sigma_j) = 1$, если $\sigma_i = \sigma_j$ и 0 в противном случае. В данной формуле может варьироваться параметр γ , принимая разные значения больше нуля.

Так как в алгоритме Лейдена без использования фиксированных начальных меток кластеров будут выбираться случайные, результаты запусков будут существенно различаться между собой. Поэтому кластеризация выполняется несколько раз при каждом условии трафика и значении параметра γ , а результаты усредняются.

Дальнейший анализ проводится для такого значения параметра γ , при котором получается наибольшее значение модулярности.

Для усреднения результатов кластеризации делаем следующее. Матрица совпадений строится по всем запускам с фиксированными условиями дорожной ситуации, чтобы определить, как часто пары узлов попадают в один и тот же кластер. Стоит отметить, что полученная матрица будет разреженной. Эта матрица нормируется на общее количество запусков и выравнивается по порогу 70%, вычисленному эвристически, для получения консенсусной кластеризации, отражающей стабильные структуры сообществ.

Далее матрица совпадений представляется в виде графа, где узлы представляют отдельные элементы, а ребра указывают на частое совпадение выше заданного порога. В этом графе необходимо найти связанные компоненты. Каждый связанный компонент соответствует консенсусному кластеру, поскольку представляет собой группу узлов, которые часто объединяются в кластеры в ходе нескольких запусков.

Такой подход гарантирует, что консенсусные кластеры определяются стабильно и устойчивы к изменчивости результатов индивидуальной кластеризации.

Для оценки устойчивости результатов кластеризации в различных условиях трафика полученные в результате вычисления консенсуса метки кластеров используются в качестве начальных меток и производится серия запусков алгоритма для последующего усреднения.

Анализ результатов

Визуализация кластеров

Результаты кластеризации для каждого запуска алгоритма были визуализированы. Примеры представлены на рисунке 2. Было отмечено, что есть кластеры, на которые дорожная ситуация не влияет, но также есть кластеры очень чувствительные к изменениям в весах ребер. Из наиболее часто наблюдаемых изменений в структуре кластеров является их разделение на несколько новых кластеров по ребрам с высокой степенью загруженности.



Рис. 2. Визуализация кластеризации с использованием времени свободного движения (слева) и высокой загруженности (справа)

Для визуализации результатов кластеризации дорожного графа используется подход, основанный на преобразовании дуг, входящих в соответствующие кластеры, в геометрические области. На первом этапе каждая дорога (дуга графа) расширяется до буферной зоны фиксированного радиуса, что позволяет объединить близлежащие дороги в единую геометрическую область. Для этого создаются буферные зоны заданного размера, а затем зоны, относящиеся к одному кластеру, объединяются в общую фигуру. Чтобы снизить сложность геометрии и упростить её отображение, проводится упрощение контуров.

После объединения буферных зон может появиться ситуация, когда внутри фигур остаются внутренние пустоты.

Чтобы этого избежать, внутренние пустоты удаляются, остается только внешний контур каждой фигуры.

После обработки геометрий и удаления внутренних пустот создаются итоговые полигоны, представляющие границы кластеров. Эти полигоны преобразуются обратно в географическую систему координат для дальнейшего анализа и визуализации на карте.

Этот подход позволяет эффективно представить результаты кластеризации на карте, визуализируя не только принадлежность дорог к разным кластерам, но и их пространственную структуру. Полученные фигуры дают наглядное представление о границах кластеров и позволяют анализировать изменения в дорожной сети под различными условиями намного лучше, чем фигуры, полученные путем построения выпуклой оболочки или диаграмм Вороного по вершинам графа.

Оценки результатов кластеризации

Зависимость модулярности и количества кластеров от параметра γ представлена на рисунке 3. Отчетливо видно, что модулярность принимает максимальные значения при определенных значениях γ . Такие значения можно считать оптимальными для кластеризации.

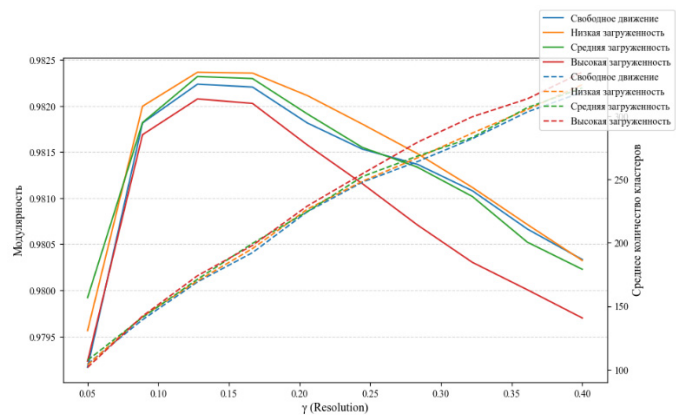


Рис. 3. График зависимости модулярности и количества кластеров от γ

Было проведено исследование стабильности результатов кластеризации при зафиксированном оптимальном значении параметра γ , равном 0.127. Сравнялся результат конкретного запуска с полученным консенсусом для запусков на одних и тех же входных данных при условии случайного распределения начальных кластеров. Для этого была использована метрика Adjusted Rand Score, результаты представлены на рисунке 4. Из результатов видно, что конкретные запуски с одними и теми же значениями параметров близки к консенсусу, полученному при серии запусков, но заметно различаются при разных значениях параметров.

На рисунке 5 показано аналогичное распределение, но при использовании консенсусных кластеров в качестве начального состояния. На данном графике видно, что при использовании времени свободного движения кластеры почти идентичны, а при увеличении загруженности возрастает разброс значений метрики вследствие роста неустойчивости кластеризации.

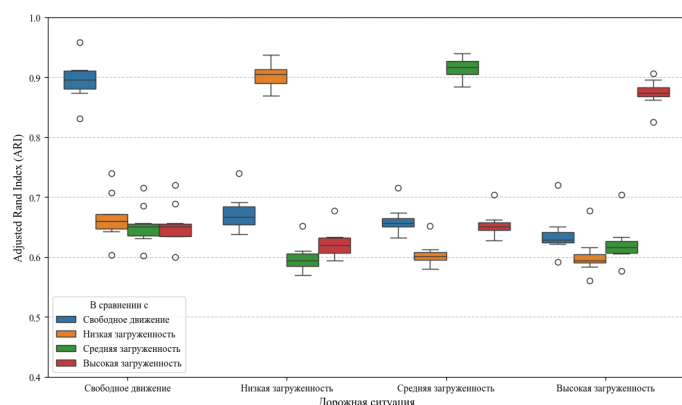


Рис. 4. Распределение ARI

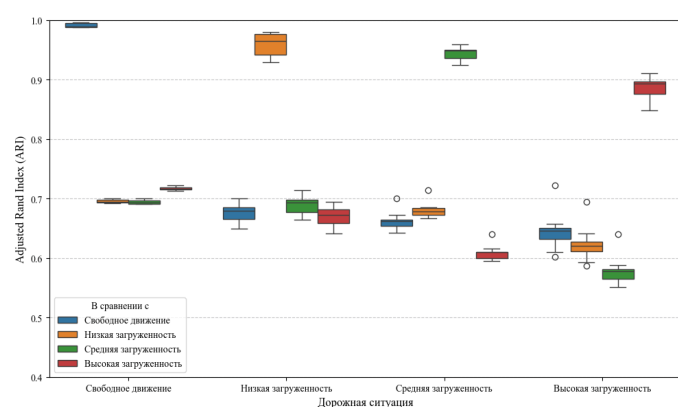


Рис. 5. Распределение ARI при использовании начального состояния

Некоторые узлы, назовем их спорными, могут принадлежать разным кластерам на нескольких запусках с одними и теми же значениями параметров. Спорные узлы, которые могут быть связаны с несколькими кластерами, становятся все более нестабильными в условиях перегруженности. По мере роста загруженности разница в весах (например, увеличение времени в пути) может сделать ранее стабильные границы неоднозначными, что приведет к появлению небольших фрагментов или новых кластеров. Консенсусный подход к кластеризации объединяет несколько запусков, чтобы получить стабильное разделение, но при высокой транспортной загруженности результаты кластеризации могут сильно отличаться. Этот эффект является отражением чувствительности сети к изменениям дорожных условий.

Ограничения и допущения

Хотя данное исследование дает ценные сведения о применении алгоритма кластеризации Лейдена в дорожных сетях при различных условиях движения, важно признать несколько ограничений. Используются несколько статических состояний загруженности дорог, которые могут не полностью отражать динамику и изменчивость во времени реальных условий движения.

Для эффективной работы с получением большого количества результатов кластеризации или вычисления консенсуса необходимо использовать параллельные вычисления. Также

вычисление матриц совпадений требует использования специальных подходов по работе с разреженными матрицами, таких как представление матриц в Йельском формате, и использования операций над разреженными матрицами.

Выводы

В данном исследовании рассматривается применение алгоритма Лейдена для кластеризации дорожных сетей в условиях переменного трафика с целью получения значимых наблюдений для улучшения понимания процессов и принятия решений в задачах городской мобильности, городского планирования и управления движением. С использованием методов кластеризации на основе консенсуса, были успешно зафиксированы стабильные структуры сообществ, а также проанализировано, как эти структуры изменяются в условиях растущей загруженности дорог.

Усредненные результаты кластеризации с использованием матриц совместного совпадения и выявлением связанных компонент для получения консенсусных кластеров оказалось эффективным для определения устойчивых кластеров. Примечательно, что в условиях свободного движения результаты кластеризации были весьма стабильными. Однако увеличение загруженности дорог приводило к росту изменчивости результатов, особенно в спорных областях вблизи границ кластеров.

Исследование показало, что с ростом загруженности дорог кластеры, как правило, распадаются, а спорные узлы вблизи основных узких мест с большей вероятностью будут образовывать границу между кластерами. Это явление показывает, как дорожные сети становятся все более разобщенными в условиях высокой загруженности, определенные регионы изолируются, и нарушаются устоявшиеся закономерности в маршрутах поездок.

Модулярность и среднее количество кластеров были проанализированы для различных разрешений и условий движения. Результаты показали, что модулярность в целом снижается с ростом загруженности дорог, что свидетельствует о снижении сплоченности кластеров. Аналогично, более высокий уровень загруженности соответствует более фрагментированным кластерам, что отражается в увеличении среднего количества кластеров.

С практической точки зрения выявление узких мест и зон, склонных к фрагментации в условиях перегруженности, может послужить основой для принятия целенаправленных мер, таких как расширение дорог, добавление альтернативных маршрутов общественного транспорта или изменение стратегий управления движением. Кроме того, выявление районов, подверженных риску изоляции в условиях сильных заторов, открывает возможности для расширения связей между районами и повышения устойчивости городских транспортных систем.

Исследование проведено с использованием упрощенной и статичной модели дорожной сети, которая может не полностью отражать сложности реального мира. Включение общественного транспорта, использование данных о трафике в реальном времени и усовершенствование методов кластеризации [25, 26] – перспективные направления будущих исследований.

Литература

1. Шаталова Н.В., Бородина О.В., Король М.М. Применение методов машинного обучения в управлении транспортно-логистическими системами Институт проблем транспорта им. Н.С. Соломенко Российской академии наук, 199178, Санкт-Петербург, 12-я линия В.О., д. 13, Санкт-Петербург, 2024. С. 258-260.
2. Волк А.Г. и др. Интегрированная Система Управления Трафиком Для Умного Города На Основе Машинного Обучения: Перспективы Для Эффективного Городского Транспорта // Политематический Сетевой Электронный Научный Журнал Кубанского Государственного Аграрного Университета. 2024. № 199. С. 229-239.
3. Akopov A.S., Beklaryan L.A. Traffic Improvement in Manhattan Road Networks With the Use of Parallel Hybrid Biobjective Genetic Algorithm // IEEE Access. 2024. (12). С. 19532-19552.
4. Newman M.E.J., Girvan M. Finding and evaluating community structure in networks // Physical Review E. 2004. № 2 (69). С. 026113.
5. Traag V.A., Waltman L., Van Eck N.J. From Louvain to Leiden: guaranteeing well-connected communities // Scientific Reports. 2019. № 1 (9). С. 5233.
6. Hairol Anuar S.H. и др. Comparison between Louvain and Leiden Algorithm for Network Structure: A Review // Journal of Physics: Conference Series. 2021. № 1 (2129). С. 012028.
7. Newman M.E.J. Fast algorithm for detecting community structure in networks // Physical Review E. 2004. № 6 (69). С. 066133.
8. Blondel V.D. и др. Fast unfolding of communities in large networks // Journal of Statistical Mechanics: Theory and Experiment. 2008. № 10 (2008). С. P10008.
9. Гатиятуллин М.Х., Николаев И.А. Оптимизация организации дорожного движения на УДС города // Техника И Технология Транспорта. 2023. № 3 (30).
10. Аламир Хайдер Сагбан Хуссейн Х.с.х., Заргарян (Финаева) Е.В., Заргарян Ю.А. Интеллектуальная система контроля заторов на дорогах с использованием контролируемого алгоритма машинного обучения на базе адаптивного IOTN // Известия Юфу. Технические Науки. 2023. № 2 (232). С. 175–186.
11. Boeing G. и др. Using open data and open-source software to develop spatial indicators of urban design and transport features for achieving healthy and sustainable cities // The Lancet Global Health. 2022. № 6 (10). С. e907-e918.
12. Golze J., Feuerhake U., Sester M. Spatial Analysis of External Influences on Traffic Accidents Using Open Data // The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. 2022. (XLIII-B4-2022). С. 125-132.
13. Naizabayeva L. и др. Identification of an algorithm for the analysis and study of urban road network trajectories // Eastern-European Journal of Enterprise Technologies. 2024. № 3 (128) (2). С. 14-27.
14. Manley E. Identifying functional urban regions within traffic flow // Regional Studies, Regional Science. 2014. № 1 (1). С. 40-42.
15. De Leo V. и др. Community core detection in transportation networks // Physical Review E. 2013. № 4 (88). С. 042810.
16. Yang S. и др. Analysis of traffic state variation patterns for urban road network based on spectral clustering // Advances in Mechanical Engineering. 2017. № 9 (9). С. 168781401772379.
17. Zhao P. и др. A network distance and graph-partitioning-based clustering method for improving the accuracy of urban hotspot detection // Geocarto International. 2019. № 3 (34). С. 293-315.
18. Nejad M.M., Mashayekhy L., Chinnam R.B. Effects of traffic network dynamics on hierarchical community-based representations of large road networks Anchorage, AK, USA: IEEE, 2012. С. 1900-1905.
19. Wang C., Wang F., Onega T. Network optimization approach to delineating health care service areas: Spatially constrained Louvain and Leiden algorithms // Transactions in GIS. 2021. № 2 (25). С. 1065-1081.
20. Wang C., Wang F. GIS-Automated Delineation of Hospital Service Areas in Florida: From Dartmouth Method to Network Community Detection Methods // Annals of GIS. 2022. № 2 (28). С. 93-109.
21. Herfort B. и др. A spatio-temporal analysis investigating completeness and inequalities of global urban building data in OpenStreetMap // Nature Communications. 2023. № 1 (14). С. 3985.
22. Boeing G. OSMnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks // Computers, Environment and Urban Systems. 2017. (65). С. 126-139.
23. Hossain I., Nower N. Traffic Data Collection and Visualization Tool for Knowledge Discovery Using Google Maps: // International Journal of Software Innovation. 2022. № 1 (10). С. 1-12.
24. Gaddam H.K., Rao K.R. Speed-density functional relationship for heterogeneous traffic data: a statistical and theoretical investigation // Journal of Modern Transportation. 2019. № 1 (27). С. 61-74.
25. Sahu S. GVE-Louvain: Fast Louvain Algorithm for Community Detection in Shared Memory Setting. 2024.
26. Sahu S., Kothapalli K., Banerjee D.S. Fast Leiden Algorithm for Community Detection in Shared Memory Setting Gotland Sweden: ACM, 2024. С. 11-20.

METHODS OF CLUSTER ANALYSIS OF ROAD NETWORKS FOR BOTTLENECK DETECTION AND TRAFFIC OPTIMIZATION

Nikita D. Povaliaev, Saint Petersburg State University, Saint Petersburg, Russia, nikita03565@gmail.com
Alexander Yu. Krylatov, Saint Petersburg State University, Saint Petersburg, Russia, a.krylatov@spbu.ru
Natalia V. Shatalova, Solomenko Institute of Transport Problems of the Russian academy of sciences, St. Petersburg, Russia, shatillen@mail.ru

Abstract

The article studies the spatial organization of road networks using modern methods of network analysis and clustering. The goal of the work is to identify stable functional structures of the transport network and analyze their changes under the influence of different levels of congestion. The application of modern algorithms of network analysis and clustering on street and road networks of large cities is proposed. Special attention is paid to the analysis of cluster dynamics in conditions of different traffic intensity. Experiments are carried out on real data of traffic flows and include estimation of changes in topological characteristics of the network under varying levels of congestion. The results show that the proposed approaches allow to effectively identify stable functional areas of the network and capture their transformations under changing traffic conditions. Regularities of flow redistribution during peak load periods and formation of potential bottlenecks were revealed. The spatial patterns of clusters show correlation with the nature of network congestion and the

topography of the city. The practical significance of the work lies in the possibility of using the proposed methodology for decision support in the field of transportation planning and traffic flow management. The results obtained can be used to optimize routing, improve the sustainability of transport infrastructure and prevent congestion on key sections of the network. The study highlights the value of applying network analysis and clustering methods to study the dynamics of transport systems and form strategies for effective traffic management in the conditions of a modern city.

Keywords: road networks, clustering, network analysis, spatial data analysis, traffic flow control, urban planning

References

- [1] N. V. Shatalova, O. V. Borodina, and M. M. Korol, "Application of machine learning methods in the management of transportation and logistics systems", Solomenko Institute of Transport Problems of the Russian academy of sciences, St. Petersburg, 2024, pp. 258-260.
- [2] A. G. Volik, S. A. Yarutin, I. S. Ugryumov, D. A. Shorvoglyan, and V. V. Dovgal, "Integrated Traffic Management System for a Smart City Based on Machine Learning: Prospects for Efficient Urban Transport", *Polythematic Network Electronic Scientific Journal of Kuban State Agrarian University*, no. 199, pp. 229-239, 2024, doi: 10.21515/1990-4665-199-021.
- [3] A. S. Akopov and L. A. Beklaryan, "Traffic Improvement in Manhattan Road Networks With the Use of Parallel Hybrid Biobjective Genetic Algorithm", *IEEE Access*, vol. 12, pp. 19532-19552, 2024, doi: 10.1109/ACCESS.2024.3361399.
- [4] M. E. J. Newman and M. Girvan, "Finding and evaluating community structure in networks", *Phys. Rev. E*, vol. 69, no. 2, p. 026113, Feb. 2004, doi: 10.1103/PhysRevE.69.026113.
- [5] V. A. Traag, L. Waltman, and N. J. Van Eck, "From Louvain to Leiden: guaranteeing well-connected communities", *Sci Rep*, vol. 9, no. 1, p. 5233, Mar. 2019, doi: 10.1038/s41598-019-41695-z.
- [6] S. H. Hairol Anuar et al., "Comparison between Louvain and Leiden Algorithm for Network Structure: A Review", *J. Phys.: Conf. Ser.*, vol. 2129, no. 1, p. 012028, Dec. 2021, doi: 10.1088/1742-6596/2129/1/012028.
- [7] M. E. J. Newman, "Fast algorithm for detecting community structure in networks", *Phys. Rev. E*, vol. 69, no. 6, p. 066133, Jun. 2004, doi: 10.1103/PhysRevE.69.066133.
- [8] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of communities in large networks", *J. Stat. Mech.*, vol. 2008, no. 10, p. P10008, Oct. 2008, doi: 10.1088/1742-5468/2008/10/P10008.
- [9] M. H. Gatiyatullin and I. A. Nikolaev, "Optimization of traffic organization on the city's road network", *Transport Engineering and Technology*, no. 3 (30), 2023.
- [10] Alamir Haider Saghan Hussein H.s.h., E. V. Zargaryan (Finaeva), and Yu. A. Zargaryan, "Intelligent Traffic Congestion Monitoring System Using Supervised Machine Learning Algorithm Based on Adaptive IOTN", *Izvestiya SFedU. Engineering Sciences*, no. 2 (232), pp. 175-186, 2023, doi: 10.18522/2311-3103-2023-2-175-186.
- [11] G. Boeing et al., "Using open data and open-source software to develop spatial indicators of urban design and transport features for achieving healthy and sustainable cities", *The Lancet Global Health*, vol. 10, no. 6, pp. e907-e918, Jun. 2022, doi: 10.1016/S2214-109X(22)00072-9.
- [12] J. Golze, U. Feuerhake, and M. Sester, "Spatial Analysis of External Influences on Traffic Accidents Using Open Data", *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XLIII-B4-2022, pp. 125-132, 2022, doi: 10.5194/isprs-archives-xliii-b4-2022-125-2022.
- [13] L. Naizabayeva, G. Turken, Z. Abdiakhmetova, Temirbekova Zh., and M. Satymbekov, "Identification of an algorithm for the analysis and study of urban road network trajectories", *Eastern-European Journal of Enterprise Technologies*, vol. 2, no. 3 (128), pp. 14-27, 2024, doi: 10.15587/1729-4061.2024.298274.
- [14] E. Manley, "Identifying functional urban regions within traffic flow", *Regional Studies, Regional Science*, vol. 1, no. 1, pp. 40-42, Jan. 2014, doi: 10.1080/21681376.2014.891649.
- [15] V. De Leo, G. Santoboni, F. Cerina, M. Mureddu, L. Secchi, and A. Chessa, "Community core detection in transportation networks", *Phys. Rev. E*, vol. 88, no. 4, p. 042810, Oct. 2013, doi: 10.1103/PhysRevE.88.042810.
- [16] S. Yang, J. Wu, G. Qi, and K. Tian, "Analysis of traffic state variation patterns for urban road network based on spectral clustering", *Advances in Mechanical Engineering*, vol. 9, no. 9, p. 168781401772379, Sep. 2017, doi: 10.1177/1687814017723790.
- [17] P. Zhao, X. Liu, J. Shen, and M. Chen, "A network distance and graph-partitioning-based clustering method for improving the accuracy of urban hotspot detection", *Geocarto International*, vol. 34, no. 3, pp. 293-315, Feb. 2019, doi: 10.1080/10106049.2017.1404140.
- [18] M. M. Nejad, L. Mashayekhy, and R. B. Chinnam, "Effects of traffic network dynamics on hierarchical community-based representations of large road networks", *2012 15th International IEEE Conference on Intelligent Transportation Systems*, Anchorage, AK, USA: IEEE, Sep. 2012, pp. 1900-1905. doi: 10.1109/ITSC.2012.6338845.
- [19] C. Wang, F. Wang, and T. Onega, "Network optimization approach to delineating health care service areas: Spatially constrained Louvain and Leiden algorithms", *Transactions in GIS*, vol. 25, no. 2, pp. 1065-1081, Apr. 2021, doi: 10.1111/tgis.12722.
- [20] C. Wang and F. Wang, "GIS-Automated Delineation of Hospital Service Areas in Florida: From Dartmouth Method to Network Community Detection Methods", *Ann GIS*, vol. 28, no. 2, pp. 93-109, 2022, doi: 10.1080/19475683.2022.2026470.
- [21] B. Herfort, S. Lautenbach, J. Porto De Albuquerque, J. Anderson, and A. Zipf, "A spatio-temporal analysis investigating completeness and inequalities of global urban building data in OpenStreetMap", *Nat Commun*, vol. 14, no. 1, p. 3985, Jul. 2023, doi: 10.1038/s41467-023-39698-6.
- [22] G. Boeing, "OSMnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks", *Computers, Environment and Urban Systems*, vol. 65, pp. 126-139, Sep. 2017, doi: 10.1016/j.compenurbysys.2017.05.004.
- [23] I. Hossain and N. Nower, "Traffic Data Collection and Visualization Tool for Knowledge Discovery Using Google Maps", *International Journal of Software Innovation*, vol. 10, no. 1, pp. 1-12, Jan. 2022, doi: 10.4018/IJSI.293270.
- [24] H. K. Gaddam and K. R. Rao, "Speed-density functional relationship for heterogeneous traffic data: a statistical and theoretical investigation", *J. Mod. Transport.*, vol. 27, no. 1, pp. 61-74, Mar. 2019, doi: 10.1007/s40534-018-0177-7.
- [25] S. Sahu, "GVE-Louvain: Fast Louvain Algorithm for Community Detection in Shared Memory Setting", Aug. 05, 2024, arXiv: arXiv:2312.04876. doi: 10.48550/arXiv.2312.04876.
- [26] S. Sahu, K. Kothapalli, and D. S. Banerjee, "Fast Leiden Algorithm for Community Detection in Shared Memory Setting", *Proceedings of the 53rd International Conference on Parallel Processing*, Gotland Sweden: ACM, Aug. 2024, pp. 11-20. doi: 10.1145/3673038.3673146.

Information about authors:

Nikita D. Povaliaev, Saint Petersburg State University. Post graduate student of the Department of Mathematical Theory of Decision Making in Economy, Saint Petersburg, Russia. ORCID: 0009-0003-4926-9086

Alexander Yu. Krylatov, Saint Petersburg State University. Doctor of Physics and Mathematics, Associate Professor, head of the Department of Mathematical Theory of Decision Making in Economy, Saint Petersburg, Russia. ORCID: 0000-0002-6634-1313

Natalia V. Shatalova, Solomenko Institute of Transport Problems of the Russian academy of sciences, Leading researcher of the laboratory of problems of transport systems organization, PhD (technical sciences), Saint Petersburg, Russia