

Санкт-Петербургский государственный университет

Математическое обеспечение и администрирование информационных сетей
Кафедра информационно-аналитических систем

Гирин Алексей Романович

Анализ и визуализация транспортных потоков Санкт-Петербургской агломерации

Бакалаврская работа

Научный руководитель:
к.ф.-м.н., доцент Графеева Н.Г.

Рецензент:
Зам. начальника отдела организации перевозок
СПб ГКУ «Организатор перевозок» Махмутов В.Р.

Санкт-Петербург
2021

SAINT-PETERSBURG STATE UNIVERSITY

Software and Administration of Information Systems
Analytical Informational Systems Department

Girin Alexey

Analysis and visualization of traffic flows in the Saint Petersburg agglomeration

Bachelor's Thesis

Scientific supervisor:
PhD, associate professor Grafeeva N.G.

Reviewer:
Deputy head of transportation management department
Makhmutov V.R.

Saint-Petersburg
2021

Оглавление

Введение	4
1. Цели и задачи работы	6
2. Требования к системе визуализации	7
2.1. Пользователи системы	7
2.2. Общие требования к системе	7
2.3. Требования, предъявляемые к совместимости	8
3. Обзор литературы и существующих решений	9
3.1. Подходы к визуализации транспортных данных	9
3.1.1. Пространственная аналитика и модель транспортной сети	10
3.1.2. Аналитика, ориентированная на время	14
3.1.3. Совмещение подходов	17
3.2. Приложения, ориентированные на отображение транспортных данных	18
4. Обзор средств визуализации картографической информации	22
5. Реализация системы	25
5.1. Исходные данные	25
5.2. Программные средства, используемые при разработке системы	28
5.3. Серверная часть и преобразование исходных данных	28
5.4. Клиентская часть и визуализация	30
Заключение	37
Список литературы	38

Введение

Постоянный рост городов естественным образом усложняет процесс перемещения жителей, что особенно актуально для крупных городских агломераций. Естественным решением этой проблемы является развитие системы общественного транспорта, которая значительно расширяет и упрощает возможности перемещения как внутри отдельных районов, так и в пределах городской агломерации в целом, поскольку ориентируется на массовое использование всеми социальными группами населения, а также является экономически и экологически выгодной.

Эффективное планирование системы общественного транспорта и организация пассажирских перевозок является сложной и многоплановой задачей, для решения которой необходимо применение комплексного анализа уже существующих маршрутов. Однако большой объем задействованных данных создаёт дополнительные трудности — в крупных городах информация о транспортной сети может содержать десятки, а порой и сотни миллионов записей, интерпретация которых без дополнительных инструментов и средств диверсифицированного и в то же время наглядного доступа к этим данным крайне затруднительна.

Одним из весьма эффективных и одновременно удобных инструментов, позволяющих получить быстрый доступ к различным типам данных, характеризующих транспортную сеть, является интерактивный графический интерфейс. Системы визуализации данных о городском общественном транспорте, предоставляющие интерфейс такого вида, можно разделить на два типа: ориентированные на использование пассажирами и предназначенные для профессионального анализа и организации транспортной сети [1]. Основной задачей систем первого типа является эффективный поиск оптимального маршрута следования от начального до конечного пункта транспортной корреспонденции путем максимально простого и интуитивно понятного отображения соответствующих элементов маршрутной сети. Однако наибольший приклад-

ной интерес представляют системы второго типа, имеющих значительно более сложную структуру и принципиально иное назначение. Базовая функциональность таких систем должна обеспечивать возможность быстрого и удобного автоматизированного доступа к определенным наборам агрегированных данных о городском пассажирском транспорте, необходимых для анализа существующей маршрутной сети и ее дальнейшей оптимизации.

Сеть городского пассажирского транспорта Санкт-Петербурга, рассматриваемая в данной работе, является крупнейшей в Северо-Западном регионе и второй в Российской Федерации. Ежедневно в городе совершается в общей сложности около 5,7 – 6,1 млн. поездок, значительное число которых приходится на наземный транспорт (в марте 2021 года действовало 43 трамвайных, 48 троллейбусных, 368 автобусных социальных и 212 автобусных коммерческих маршрутов). Важность общественного транспорта Санкт-Петербурга, как одной из ключевых составляющих городской социальной инфраструктуры, определяет необходимость его постоянного анализа и оптимизации, для чего, как уже было отмечено, требуется применение инструментов, упрощающих и ускоряющих работу с данными, используемыми для анализа маршрутной сети. В настоящий момент в СПб ГКУ "Организатор перевозок" (далее – Учреждение) при работе с вышеуказанными данными такого рода средства применяются недостаточно, что делает весьма актуальным вопрос о создании инструмента визуализации данных о характеристиках транспортной системы города.

1. Цели и задачи работы

Целью данной работы является создание приложения, позволяющего визуализировать доступ к массивам данных, используемым для анализа сети городского пассажирского транспорта Санкт-Петербургской агломерации.

Для достижения поставленной цели были определены следующие задачи:

- формирование требований к разрабатываемой системе в соответствии с запросами аналитиков сети;
- изучение релевантных методов представления данных анализа транспортной сети и обзор существующих систем, в которых такие методы используются;
- поиск и изучение средств визуализации картографической информации и соответствующих программных решений;
- реализация системы автоматизированного доступа к данным путем визуализации запросов.

2. Требования к системе визуализации

В данном разделе приводится перечень основных требований к системе визуализации данных о характеристиках транспортной сети города, определенный в ходе консультаций с сотрудниками Учреждения.

2.1. Пользователи системы

Конечными пользователями разрабатываемой системы выступают работники Учреждения.

2.2. Общие требования к системе

Система визуализации данных о характеристиках транспортной сети города должна обеспечивать следующие возможности:

- Автоматизированный доступ к данным, реализуемый с помощью иерархического меню с параллельным отображением картографической информации о транспортной сети:
 - визуальная интерактивная модель следующих объектов (с возможностью выбора той или иной части объекта и с последующим детализированным представлением соответствующей этому выбору информации о характеристиках):
 - * отдельный маршрут;
 - * транспортная сеть для конкретного вида транспорта;
 - сведения о следующих характеристиках (как посуточные, так и усредненные за месяц):
 - * скорость прохождения транспортными средствами одного маршрута соответствующих этому маршруту межостановочных перегонов с разбивкой по различным временным промежуткам;
 - * время прохождения транспортными средствами одного маршрута соответствующих этому маршруту межостановочных перегонов;

- вочных перегонов с разбивкой по различным временным промежуткам;
- * наполнение салонов транспортных средств одного маршрута на соответствующих этому маршруту межостановочных перегонах с разбивкой по часовым временным промежуткам;
 - * пассажирооборот на маршруте для каждого соответствующего этому маршруту остановочного пункта с разбивкой по часовым временным промежуткам;
 - * пассажирооборот каждого остановочного пункта с обобщением по виду транспорта и разбивкой по часовым временным промежуткам;
- Картографическое интерактивное представление межостановочной матрицы корреспонденций с числовым сопровождением отображаемой информации.

2.3. Требования, предъявляемые к совместимости

Система визуализации данных о характеристиках транспортной сети должна обеспечивать доступ посредством web-интерфейса к базе данных анализа транспортной сети, которая хранится на стороне сервера. Запросы к базе должны обрабатываться стандартными браузерами Google Chrome, Opera и Mozilla Firefox.

3. Обзор литературы и существующих решений

3.1. Подходы к визуализации транспортных данных

В литературе описано большое количество методов отображения данных анализа систем общественного транспорта. В целом их можно разделить на три условные категории: ориентированные на пространственное отображение, представляющие динамику изменения различных параметров сети и совмещенные. Выбор того или иного метода основывается на конкретных характеристиках транспортной системы, которые требуется визуализировать, и на необходимом для поставленных целей уровне детализации представления. Как правило, данные о движении средств общественного транспорта и о пассажирских поездках являются многопараметрическими и описываются очень большими объемами, что определяет ряд общих принципов выбора методов их отображения, позволяющих сделать его максимально наглядным. Здесь можно выделить: агрегацию (для избежания избыточного числа объектов на карте), фильтрацию (позволяет пользователю работать только с теми данными, которые необходимы ему в текущий момент), абстракцию (отображение наиболее значимых данных) и анимацию (для представления динамики изменения определенных характеристик объектов в зависимости от меняющихся внешних параметров) [1]. Применение этих правил обуславливает необходимость использования интерактивных методов визуализации, которые должны не просто обеспечивать пользователя статичной информацией с рядом фиксированных показателей, но и позволять ему напрямую взаимодействовать с отображаемыми объектами (например, визуально выбирать те или иные их части или изменять масштаб для получения более или менее подробного описания). Далее приводится обзор различных подходов представления данных о характеристиках транспортной сети.

3.1.1. Пространственная аналитика и модель транспортной сети

Структура сети общественного транспорта задается множеством остановочных пунктов и перегонов, которые их связывают. Остановочный пункт является точечным объектом, определяемым набором географических координат, что делает возможным его представление на карте с помощью метки. Для отображения перегона подойдет полилиния, начальной и конечной точками которой являются начальный и конечный остановочные пункты данного перегона. Использование меток и полилиний в качестве базовых объектов отображения картографической информации лучше всего подходит для визуализации на картографической подложке как моделей отдельных элементов транспортной системы (например, маршрутов), так и ее агрегированной структуры [1].

Свойства, присваиваемые моделирующим объектам, могут быть полезны при демонстрации данных о характеристиках транспортной сети. Для каждой метки и полилинии можно установить свой цвет и размер в соответствии с величиной выбранной характеристики, что позволяет наглядно, быстро и просто представлять общую картину либо для всей сети, либо для отдельных ее элементов (рис. 1) [2].

В целом применение такого примитива отображения на карте, как метка, удобно для моделирования любого точечного объекта, будь он статичным (как остановочный пункт) или динамично изменяющим свое местоположение (например, транспортное средство в движении). Однако если число одновременно представляемых объектов превышает некий предел, усложняется восприятие отображаемой информации (общая картина становится трудночитаемой). В этом случае эффективным методом является кластеризация меток (например, кластеры отображаются на маленьких зумах, а сами объекты только при достаточном приближении), их объединение в более общие структуры (например, множество точечных координат, получаемых от какого-либо движущего-

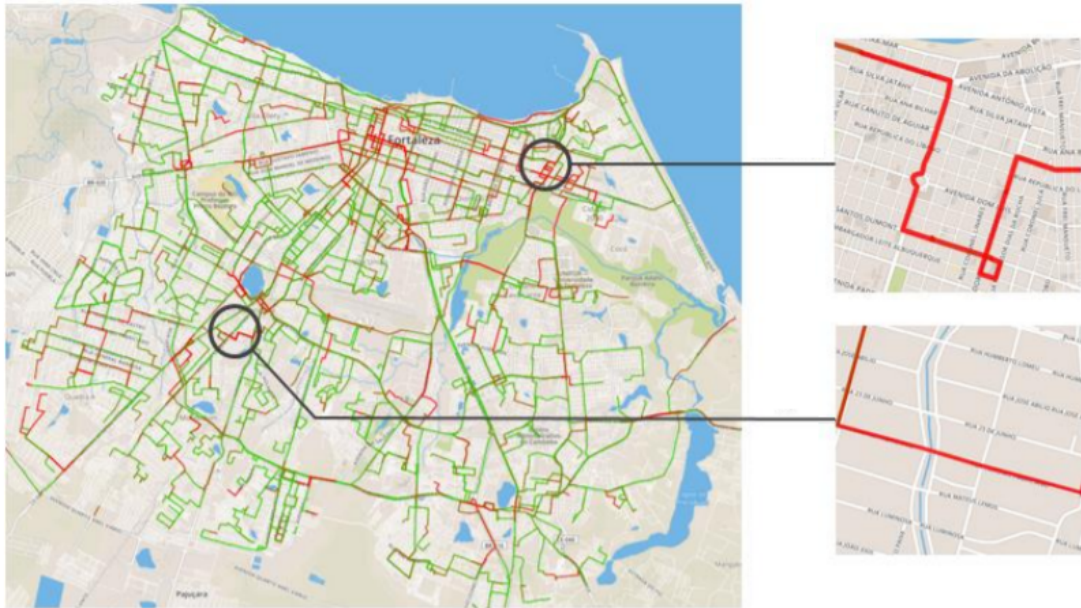


Рис. 1: Демонстрация автобусной сети города с использованием полилиний и выделением красным цветом проблемных участков (г. Форталеза) [3]

гося объекта, задают траекторию его пути и могут быть объединены в полилинию), либо выполнение оценки плотности их расположения и использование такого средства обобщения, как тепловая карта, для демонстрации интенсивности на которой применяется цветовая палитра (рис. 2) [4]. Использование полилиний не ограничивается отображением перегонов транспортной сети, этот метод полезен при визуализации любых траекторий движения, а также набора каких-либо объектов с идентичными характеристиками, что в дальнейшем описывается на примере методов отображения передвижения пассажиров.

Общим инструментом для визуализации направления и интенсивности движения является карта потоков (рис. 3) [2]. Природа такого движения может быть различной (например, перемещение торговых товаров или материалов и ресурсов в какой-либо среде), но в контексте рассматриваемой темы интерес представляет отображение пассажирских потоков. Структура этой карты определяется множеством точечных объектов (узлов), задающих начальные и конечные пункты движения, а также линиями, которые соединяют их, если имеется соот-

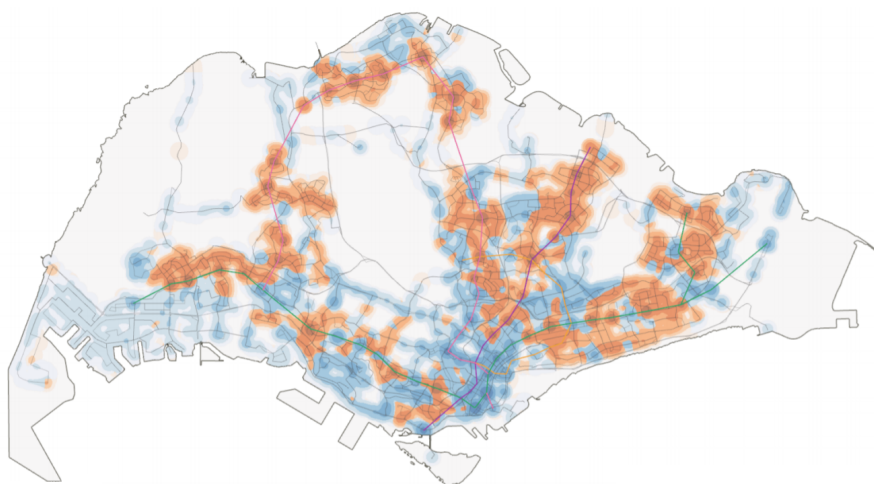


Рис. 2: Тепловая карта, показывающая плотность начальных (красный цвет) и конечных (синий цвет) граничных точек для поездок пассажиров в утреннее время (г. Сингапур) [2]

ветствующий поток. Интенсивность показывается толщиной отображаемых линий, а направление потоков – добавлением стрелок или уменьшением толщины с приближением к конечному пункту, в данном случае возможно представление узлов в древовидной иерархической структуре с последующим слиянием нескольких соседних потоков, исходящих из одного пункта [5], а также использование метода спирального дерева для ухода от пересечений между потоками [6]. Стоит отметить, что применение данного подхода сопряжено с некоторыми трудностями – одновременное отображение большого числа потоков приведет к затруднению восприятия отдельных элементов изображения, кроме того, возможно поглощение меньших потоков большими при наложении изображений.

Можно выделить еще один метод, описанный в статье по анализу использования сотовой сети [7], который подойдет для визуализации величины пассажирооборота или сведений из матрицы межстаночных или межрайонных корреспонденций. Авторы этой работы предложили отображать на карте объем входящих и исходящих звонков из различных точек города с помощью кругов. Каждый круг сопоставляется какой-либо из этих точек, его цвет принимает одно из двух значений

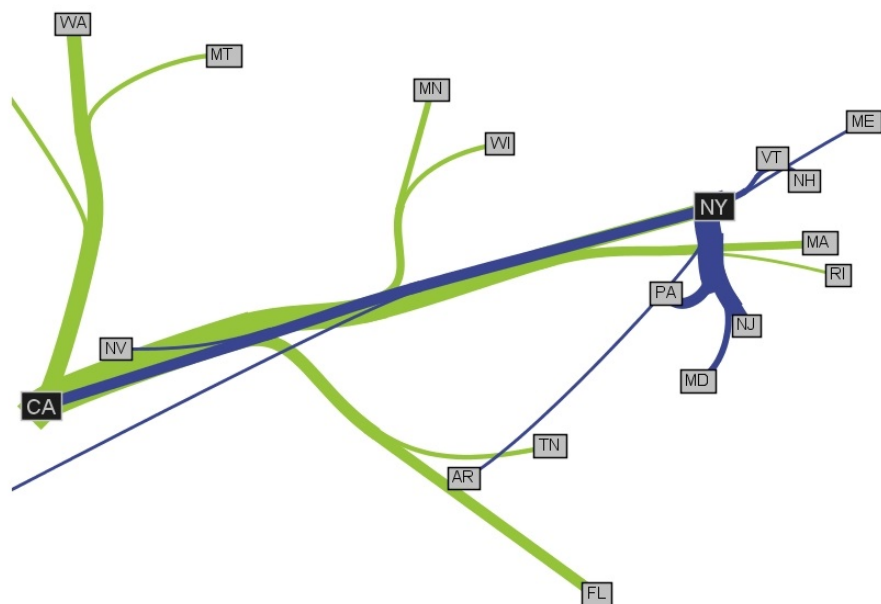


Рис. 3: Карта потоков древовидной структуры, визуализирующая передвижение населения между некоторыми штатами в США [5]

(в зависимости от типа звонков – входящие или исходящие), а радиус определяется их числом (рис. 4). Такие круги можно использовать для отображения величины пассажирооборота остановочных пунктов с установлением зависимости радиуса круга от числа вошедших или вышедших пассажиров.

Как было отмечено выше, для визуального представления модели транспортной сети подойдет использование полилиний и меток (другими словами, графа из меток в качестве вершин и полилиний в качестве ребер). Подобное представление с помощью графов может быть удобно для отображения на карте данных из матрицы корреспонденций – метки соотносятся с остановочными пунктами или районами города (узлами), которые соединяются линиями, если между соответствующими узлами имеются пассажирские поездки. Избежать избыточности при визуализации информации помогает выделение отдельных линий в общей связке или фильтрация элементов графа (рис. 5) [2].

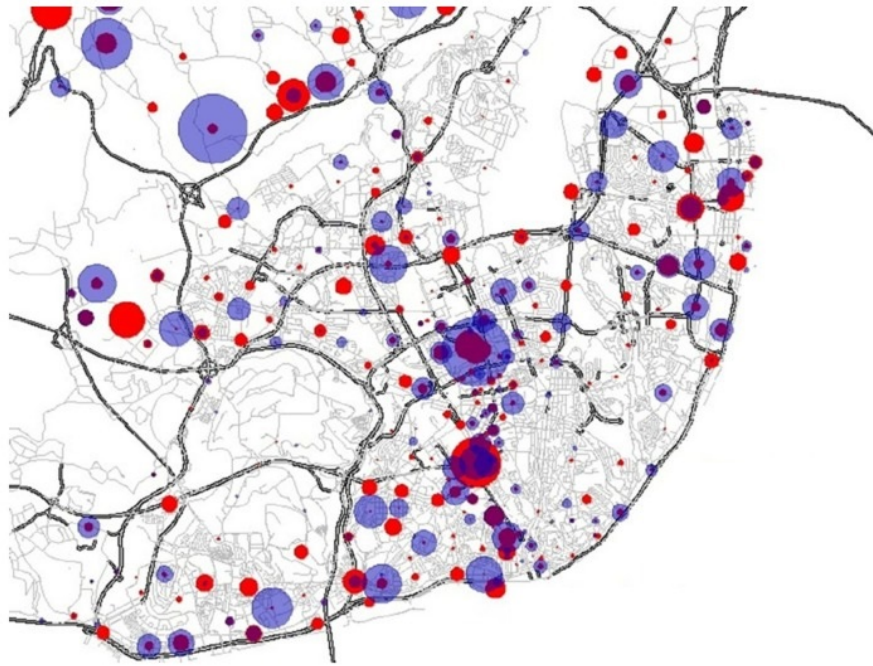


Рис. 4: Использование круговых объектов для визуализации объема входящих (красный цвет) и исходящих (синий цвет) звонков [7]

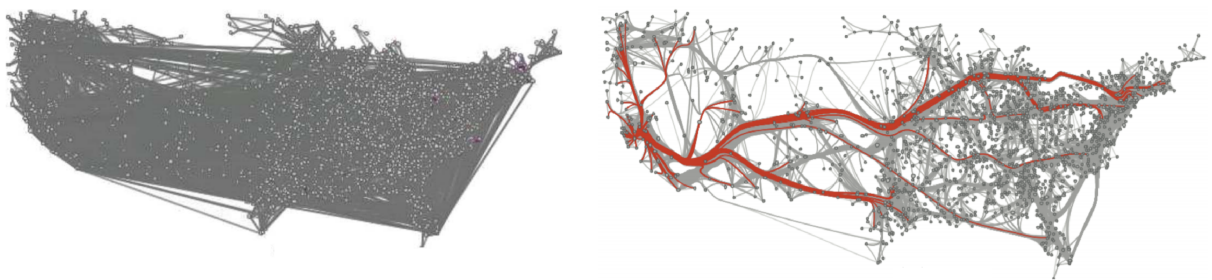


Рис. 5: Представление с помощью графов передвижения населения в США внутри страны (слева используется линейное отображение ребер, справа применяется алгоритм их объединения) [8]

3.1.2. Аналитика, ориентированная на время

Характеристики транспортной сети изменяются в зависимости от времени, это определяет необходимость в методах, позволяющих визуализировать динамику этих изменений. Такая зависимость может быть линейной и периодической. В первом случае для отображения изменения значений характеристики подойдет использование гистограмм и графиков. Для более общего и наглядного представления динамики при линейной зависимости можно применять различные подходы анимации – моделирование движения объектов на карте, использование ползунка

на временной шкале, при изменении положения которого будет меняться, например, насыщенность цветовыделения графа транспортной сети.

В работе [9] был предложен подход к визуализации загруженности (в зависимости от времени суток) объектов, представляемых на карте точечными узлами, с помощью диаграммы роз. В исходном варианте метод используется для демонстрации интенсивности дорожных заторов в определенных точках города. Каждой такой точке сопоставляется круг, разбитый на сегменты, которые соответствуют интервалам времени суток. Цветовое задание для одного сегмента определяется количеством возникших пробок в данной части города за соответствующий промежуток времени, а радиус сегмента их продолжительностью (рис. 6). Представляется возможным применение такого подхода для отображения величины пассажирооборотов остановочных пунктов города – узлы сопоставляются остановочным пунктам, а радиус или цвет какого-либо сегмента задается числом вошедших или вышедших пассажиров в соответствующее время суток.

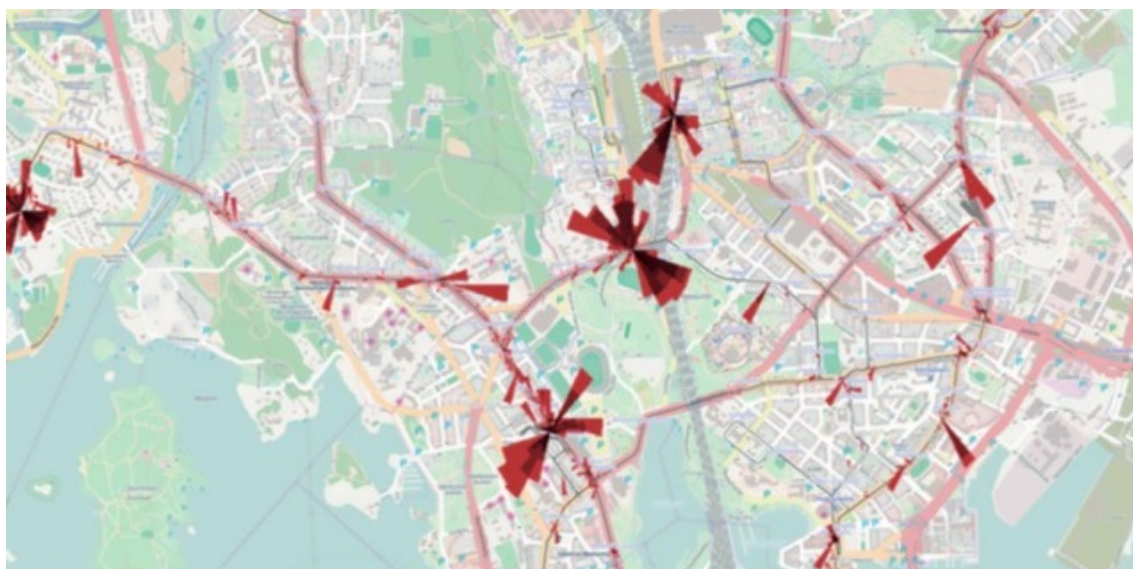


Рис. 6: Использование диаграммы роз для анализа загруженности дорог в г. Хельсинки [9]

Похожий подход применяется еще в одной системе визуализации дорожного трафика [10]. Аналогично построению диаграммы роз вы-

полняется разбиение кольца на сегменты, каждый из которых соответствует определенному времени суток, однако для демонстрации интенсивности какой-либо характеристики (в исходной работе такой характеристикой является степень загруженности дорог в определенном месте города) используется только выделение сегментов различными цветами без изменения их радиусов. Такая кольцевая структура удобна для представления, если есть периодическая зависимость характеристики от времени, так как она позволяет делать несколько уровней вложения и отображать информацию, например, сразу за неделю (рис. 7а). Данный метод носит название "отпечаток пальца", его также можно использовать для визуализации пассажирооборота остановочных пунктов.

Если вложенных колец в "отпечатке пальца" много, радиус каждого следующего кольца по направлению к центру будет уменьшаться, что в конечном итоге приведет к свертыванию диаграммы. Для решения этой проблемы можно раскрыть "отпечаток" в матрицу тепловой карты (рис. 7б), тогда число отображаемых слоев может быть сколь угодно большим без потери в наглядности представления [11].

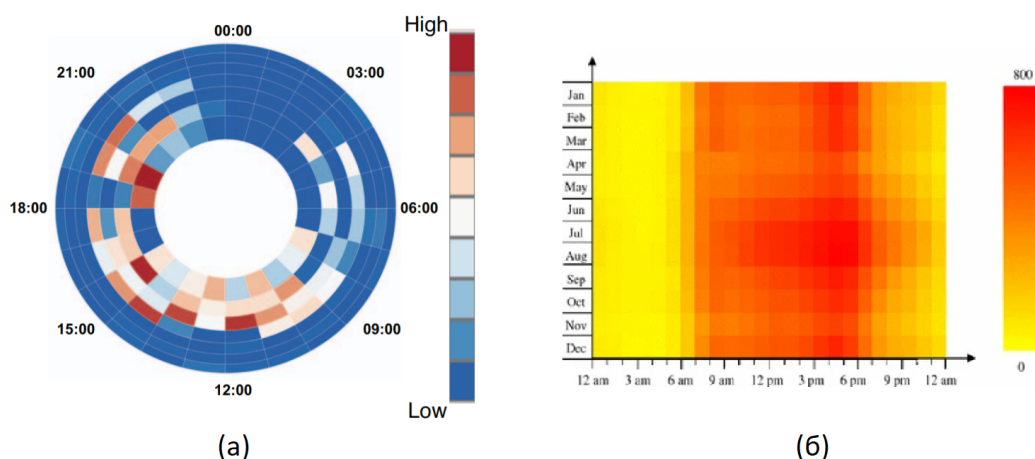


Рис. 7: Демонстрации загруженности движения на участке дороги в зависимости от времени суток с использованием (а) диаграммы "отпечаток пальца" (слои соответствуют дням недели) [10], (б) матрицы тепловой карты (слои соответствуют месяцам) [11]

Стоит указать еще один метод, который позволяет визуализировать транспортную доступность городских объектов, а именно карту изохрон – на географической карте выделяются зоны достижимости из начальной точки за определенное в параметрах время (рис. 8) [12].



Рис. 8: Карта изохрон, демонстрирующая достижимость городских объектов из начальной точки за 30 минут (темно-синий цвет) и 60 минут (светло-синий цвет) (г. Сингапур) [2]

3.1.3. Совмещение подходов

Как уже было отмечено, интерактивность инструмента особенно важна при отображении данных о характеристиках транспортной системы. Она может быть получена путем совмещения нескольких подходов к визуализации – для разного уровня детализации (который может изменяться, например, при выборе пользователем какой-либо части отображаемого элемента или смене уровня масштабирования карты) будет использоваться тот или иной метод, представляющий соответствующие сведения о системе [13]. Пример применения интерактивного подхода с семантическим масштабированием приведен на рис. 9.

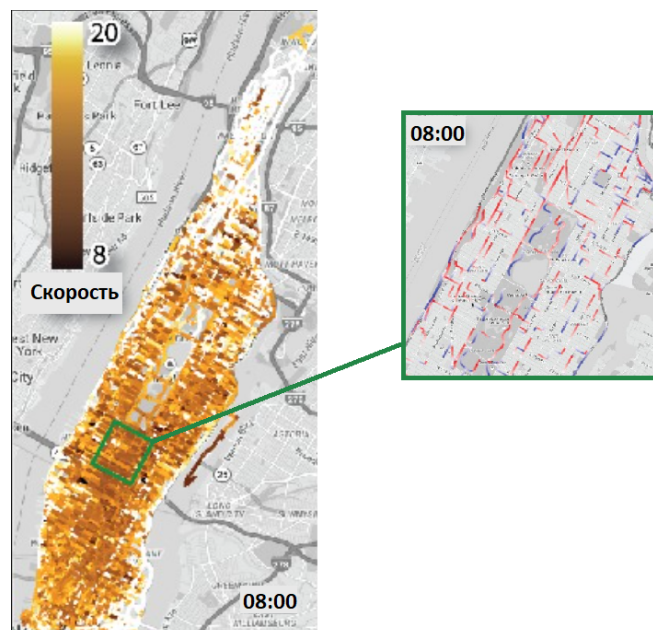


Рис. 9: Совмещение методов визуализации, для отображения скорости движения на маленьких зумах используется тепловая карта, а при приближении – цветное выделение сегментов дороги [13]

3.2. Приложения, ориентированные на отображение транспортных данных

Весомая часть существующих сервисов по визуализации данных о городском общественном транспорте ориентирована на использование пассажирами. В качестве примера такого сервиса можно отметить популярное приложение от компании "Яндекс" Яндекс.Карты¹, которое позволяет следить за перемещением средств общественного транспорта в режиме реального времени, строить оптимальные маршруты проезда и прогнозировать время ожидания. Исходными данными для работы приложения являются геометрии траекторий движения транспортных средств и сведения о координатах отдельных транспортных средств в определенный момент времени, а в качестве алгоритма расчета времени прибытия и сглаживания движения на карте выступает собственный прогнозатор Яндекса. Ввиду того, что это приложение предназначено для использования пассажирами, а все отображаемые сведения о движении транспорта являются неточными и имеют непрофессионально

¹URL: <https://yandex.ru/map/> (дата обращения: 01.04.2021)

ориентированный характер, невозможно его применение для представления данных о характеристиках транспортной сети (скорости, времени, величины пассажиропотока или наполнения салона), необходимых для отслеживания состояния системы. Прочие клиентоориентированные приложения (например, Карты Google²) по этой же причине не удовлетворяют поставленным в работе требованиям. Кроме того, в соответствии с требованиями общегосударственной политики при выборе программных продуктов или решений на их основе, предпочтение отдаётся решениям на базе бесплатного ПО либо продуктов OpenSource.

Примером приложения для визуализации городской автобусной транспортной сети является система BusViz, используемая управлением наземного транспорта г. Сингапура [14]. Архитектура этой системы состоит из трех уровней. На самом нижнем уровне организовано хранение данных о пассажирских перевозках и движении транспортных средств, получаемых непосредственно с бортовых компьютеров. На втором уровне эти данные анализируются, предоставляется интерфейс для доступа к ним в обобщенном виде. Верхний уровень представления обращается к этому интерфейсу по запросам пользователей и выполняет отображение. Система акцентируется на визуализации данных по отдельным маршрутам, траекторию каждого из них можно представить на карте в интерактивном виде с использованием полилиний и меток, при выборе той или иной метки отображаются подробные сведения о заполненности прибывающих на соответствующий остановочный пункт автобусов (рис. 10). Используется пространственно-временное представление движения транспортных средств одного маршрута для демонстрации их загруженности и скопления в определенных местах (рис. 11).

Еще одним примером сервиса, предназначенного для анализа городских автобусных перевозок, может служить система Smartbus [3], разработанная сотрудниками Аризонского университета и Универси-

²URL: <https://www.google.com/maps/> (дата обращения: 01.04.2021)

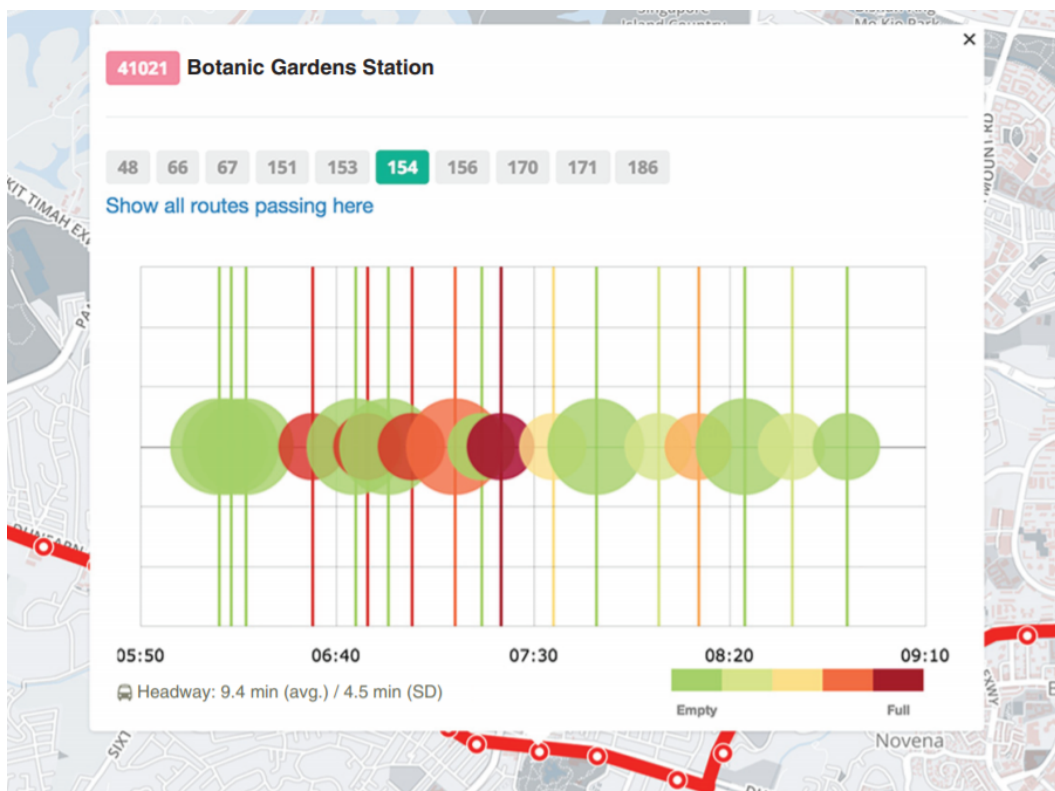


Рис. 10: Демонстрация в BusViz заполненности транспортных средств одного маршрута, прибывающих на конкретный остановочный пункт в определенное время. Цвет соответствует числу пассажиров в салоне, маленькие круги относятся к стандартным автобусам, а большие к двухэтажным [14]

тета Форталезы (рис. 12). Аналогично предыдущей системе Smartbus является web-приложением с трехуровневой архитектурой, каждая ее часть выполняет обработку исходных данных, вычисления или анализ, все уровни системы взаимодействуют с общим модулем визуализации. Данный модуль предоставляет карту, на которую накладывается граф, соответствующий автобусной сети города, и диаграммы, на которых отображается агрегированная информация о состоянии этой сети. Используется выделение цветом элементов графа для выполнения визуального анализа, определения проблемных участков (рис. 1).

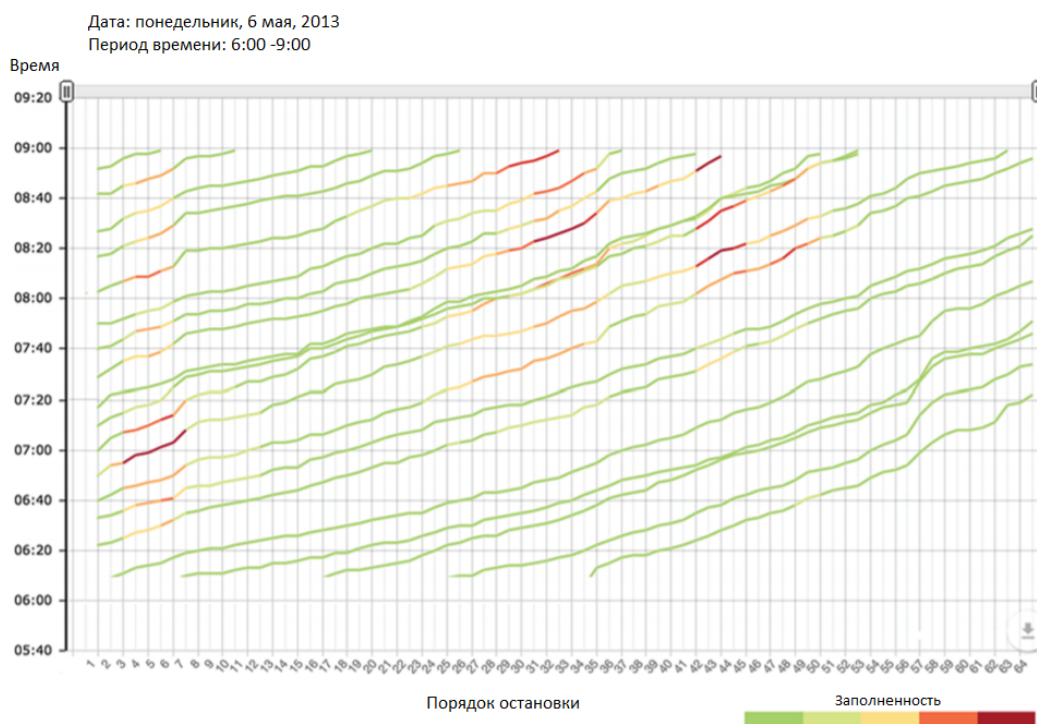


Рис. 11: Временная диаграмма в системе BusViz, линии соответствуют движениям автобусов по одному маршруту (их цвет задается наполнением салона), с помощью перекрывающихся линий можно определять скопления транспортных средств [14]

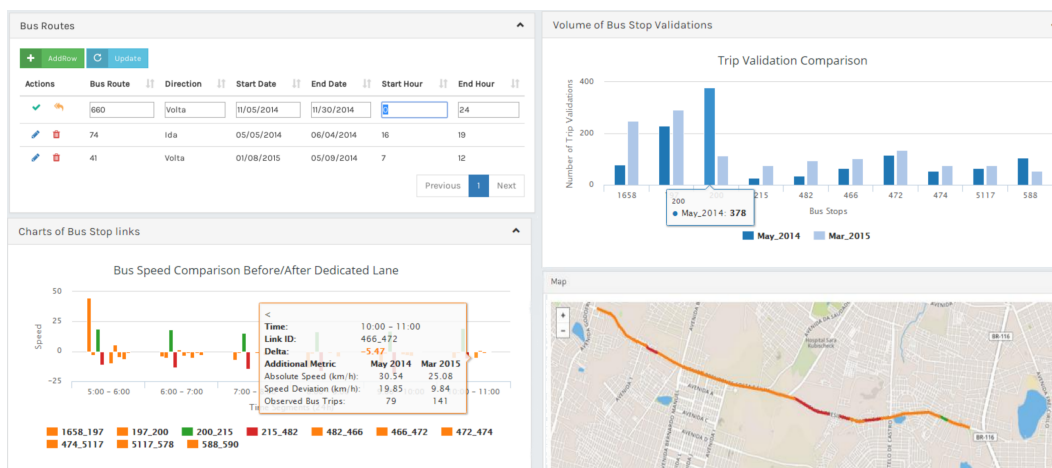


Рис. 12: Интерфейс приложения Smartbus, представлен инструмент выбора автобусного маршрута с отображением подробной информации о нем [3]

4. Обзор средств визуализации картографической информации

Существует большое число инструментов и библиотек для работы с картографическими данными в web-браузере. Они могут различаться, например, функциональными возможностями, условиями использования, набором поддерживаемых протоколов хранения пространственной информации и т.д. Поскольку визуализация информации о транспортной сети в значительной части выполняется на картографической подложке, использование таких средств в разрабатываемой системе является необходимым. Далее приведен их краткий обзор, а также определены критерии выбора.

Исходя из задач, поставленных в этой работе, средство визуализации картографической информации должно обеспечивать следующие возможности:

- отображение базового слоя двумерной картографической подложки со схематическим покрытием;
- масштабирование карты;
- наложение дополнительных слоев на базовый с предоставлением методов их группировки и управления; отображение на карте большого числа (до нескольких тысяч) интерактивных геометрических объектов (меток и полилиний) с различными свойствами (размер, цвет), а также иерархическое взаимодействие между ними;
- поддержка кластеризации и анимации;
- бесплатное использование в закрытых системах.

Можно отметить следующие популярные на данный момент инструменты для отображения и работы с картографическими данными [15]:

- Leaflet³. Особенности этой JavaScript библиотеки являются компактностью; модульная система, обеспечивающая эффективность решений, что положительно сказывается на скорости отображения данных; наличие большого числа плагинов, с помощью которых можно придавать решениям дополнительный функционал, отсутствующий изначально в базовой версии (например, реализовано несколько программных модулей для удобного построения тепловых карт, что не относится к вышеописанным требованиям к средствам визуализации, но может быть полезным при расширении приложения). По умолчанию в качестве картографической подложки библиотека использует тайловые слои от ресурса OSM (OpenStreetMap)⁴, есть возможность пользовательского выбора поставщиков картографических изображений, поддерживается работа с сервисами OGC (Open Geospatial Consortium)⁵. Библиотека обеспечивает всю необходимую для данной работы функциональность, она проста в использовании и снабжена полной и актуальной документацией. Минусом Leaflet является возможное возникновение проблем с совместимостью плагинов, так как каждый из них может создаваться отдельными разработчиками независимо.
- Mapbox GL⁶. Основными преимуществами этого инструмента являются хорошая оптимизация и высокая производительность, а также более широкие функциональные возможности по сравнению с библиотекой Leaflet. Mapbox GL ориентируется на работу с векторными слоями и 3D моделирование, особое внимание уделяется визуальным стилям карт и отображаемых на них объектов, а также детализации. Mapbox GL удовлетворяет функциональным требованиям данной работы, однако с декабря 2020 года использование последней версии библиотеки (v.2) является платным, а

³URL: <https://leafletjs.com/> (дата обращения: 01.04.2021)

⁴URL: <https://www.openstreetmap.org/> (дата обращения: 01.04.2021)

⁵URL: <https://www.ogc.org/standards/owc> (дата обращения: 01.04.2021)

⁶URL: <https://docs.mapbox.com/mapbox-gl-js/api/> (дата обращения: 01.04.2021)

предыдущая бесплатная версия (v.1) больше не развивается.

- OpenLayers⁷. Еще одна JavaScript библиотека, имеет широкую поддержку протоколов картографических данных, полную (но несколько устаревшую) документацию, обеспечивается хорошая скорость работы для последней версии (v.3), однако по производительности OpenLayers уступает библиотекам Leaflet и Mapbox GL, имеет большие функциональные возможности, которые покрывают установленные требования.
- Google Maps API⁸. Проприетарная библиотека без возможности стороннего получения картографических данных и со значительными ограничениями на бесплатное использование, уступает первым трем библиотекам из списка в гибкости применения. К плюсам этого средства визуализации можно отнести простую и быструю интеграцию карты, что удобно при необходимости отображения ограниченного набора данных для небольшой системы.
- Yandex.Maps API⁹. В целом данная JavaScript библиотека по своим особенностям близка к предыдущей, она также является закрытой и задает ограничения на использование, лицензионное соглашение запрещает бесплатно применять библиотеку в системах без свободной регистрации.

Таким образом, из рассмотренных средств визуализации картографических данных для использования в данной работе наиболее подходящими являются библиотеки Leaflet и OpenLayers. Существенных (в контексте поставленных требований) различий между ними нет, однако применение Leaflet предпочтительнее, так как она более компактна и проста в использовании.

⁷URL: <https://openlayers.org/> (дата обращения: 01.04.2021)

⁸URL: <https://developers.google.com/maps/documentation/javascript/overview> (дата обращения: 01.04.2021)

⁹URL: <https://yandex.ru/dev/maps/> (дата обращения: 01.04.2021)

5. Реализация системы

5.1. Исходные данные

Исходные данные сети общественного транспорта Санкт-Петербурга предоставляются Учреждением. Источниками данных служат системы АСУ ГПТ (Автоматическая система управления городским пассажирским транспортом), осуществляющая геоинформационный мониторинг транспортных средств и его предобработку, и СЭКОП (Система электронного контроля оплаты проезда), с помощью которой осуществляется электронный контроль оплаты проезда пассажирами. Расчеты по этим данным, необходимые для получения общих характеристик транспортной сети, уже были выполнены в системе анализа пассажиропотоков Санкт-Петербурга [16], а их результаты используются в данной работе.

Информация о времени прибытия транспортных средств на остановочные пункты содержится в таблице треков, структура которой определяется следующим образом:

1. Вид транспорта;
2. Номер маршрута;
3. Бортовой номер транспортного средства;
4. Номер наряда;
5. Номер рейса;
6. Идентификатор остановочного пункта;
7. Дата и время прибытия;

Месячный объем данных в этой таблице 31-34 млн. записей, однако среди данных о времени прибытия содержится значительное число пропусков (около 25-30% от общего числа).

По информации из таблицы треков вычислена средняя скорость движения транспортных средств на определенных перегонах в различные периоды дня. Структура таблиц, содержащих скорости движения транспортных средств по отдельным маршрутам следующая:

1. Вид транспорта;
2. Номер маршрута;
3. Идентификатор начального остановочного пункта;
4. Идентификатор конечного остановочного пункта;
5. Тип дня (будний или выходной);
6. Значения скорости для различных интервалов времени (всего имеется три таблицы, с часовыми, 15-минутными и общими интервалами);

Дополнительно описывается средняя скорость движения на каждом перегоне для всех соответствующих маршрутов одного вида транспорта. Аналогичным образом определяются таблицы со средним временем прохождения перегонов. Каждая таблица содержит чуть более 45 тыс. записей.

Математическая обработка данных, взятых при совместной обработке треков и системы оплаты проезда, позволяет получить информацию об использовании уникального проездного документа на конкретном транспортном средстве, которая помещается в таблицу поездок. Информация о каждой уникальной поездке содержит следующие сведения:

1. Уникальный идентификатор проездного документа;
2. Тип идентификатора проездного документа;
3. Идентификатор начального остановочного пункта;
4. Идентификатор конечного остановочного пункта;
5. Дата и время начала поездки;
6. Дата и время завершения поездки;
7. Вид транспорта;
8. Номер маршрута (по которому осуществлялась поездка);
9. Бортовой номер соответствующего транспортного средства;
10. Номер наряда;
11. Номер рейса;

Общее количество записей за один месяц колеблется в пределах 85-105 млн.

На основании данной таблицы, которая является основным источником информации при расчётах, и производятся вычисления большинства аналитических характеристик сети и производных таблиц. Таблица суммарного пассажиропотока за месяц содержит около 50 тыс. записей, а таблица суммарного пассажирооборота за месяц около 100 тыс. Дополнительно используются таблицы со значениями для ежедневного пассажиропотока и пассажирооборота по маршрутам, а также пассажирооборота по остановочным пунктам (с обобщением по всем маршрутам для одного вида транспорта).

Сведения о числе корреспонденций между остановочными пунктами с разбиением по временным промежуткам (с 15-минутным и часовым) содержатся в двух таблицах матриц корреспонденций (количество записей колеблется в пределах 9-11 млн.). Сведения о траекториях движения транспортных средств находятся в таблице, содержащей данные о геометрии маршрутов (25 тыс. записей). Последовательные наборы координат соответствуют прохождению транспортным средством варианта маршрута на конкретном перегоне, определяемым идентификаторами начальной и конечной остановок настоящего перегона. Данные о траекториях являются полными, однако среди значений координат имеются сдвиги, особенно выраженные на конечных узлах, что обусловлено выбросами систем геопозиционирования, возникающими при трассировке маршрута. Отдельными таблицами представлены подробные сведения об остановочных пунктах (запись для каждого остановочного пункта среди прочего содержит его идентификатор, название и координаты, всего около 10 тыс. записей), маршрутах, вспомогательная и иная информация о транспортной сети.

5.2. Программные средства, используемые при разработке системы

Для разработки приложения была выбрана среда Oracle Application Express, как наиболее оптимальная для выполнения поставленных задач. Хранение и управление данными, которые используются в приложении, осуществляется через объектно-реляционную систему управления базами данных Oracle Database.

В соответствии с обзором средств визуализации картографической информации (см. выше), инструментом отображения геоданных транспортной сети выбрана программная библиотека JavaScript Leaflet. Для создания пользовательского интерфейса и определения логики обработки запросов используются средства Oracle APEX и язык JavaScript, а для работы с данными на стороне сервера язык PL/SQL. Построение диаграмм выполняется с помощью библиотеки D3.js¹⁰ и web-сервиса Google Charts¹¹.

5.3. Серверная часть и преобразование исходных данных

Приложение классическим образом разделяется на серверную и клиентскую части – на стороне сервера организуется хранение и обработка всех данных о транспортной системе, а также весь массив промежуточных вычислений, необходимых для работы пользователя; на стороне клиента, в соответствии с конкретными запросами, выполняется обращение к серверу за этими данными по технологии AJAX, а после их визуальное представление.

Исходными таблицами при построении картографической модели транспортной сети (графическое отображение отдельных маршрутов

¹⁰URL: <https://d3js.org/> (дата обращения: 01.04.2021)

¹¹URL: <https://developers.google.com/chart> (дата обращения: 01.04.2021)

и транспортной сети в целом) являются таблицы остановок и геометрий маршрутов. Данные из этих таблиц очищаются и приводятся в единообразный вид для дальнейшей визуализации (например, значения координат конечных точек геометрий для некоторых траекторий движения не всегда точны, что корректируется замещением координатами соответствующих остановочных пунктов). Создаются два обработчика AJAX-запросов, через которые на стороне клиента получают преобразованные данные в формате JSON, которые в дальнейшем используются либо для построения модели транспортной сети (по какому-либо одному виду транспорта), либо для построения модели отдельного маршрута (в качестве параметров запроса выступают номер маршрута, его вариант и направление движения – прямое или обратное).

На рис. 13 приводятся варианты характеристик движения транспортных средств и перемещения пассажиров, значения по которым необходимо предоставить клиентской стороне. Для каждого такого варианта описаны параметры соответствующих AJAX-запросов. Данные о скорости, времени и корреспонденциям, содержащиеся в исходных таблицах, не требуют дополнительных преобразований. Дополнительное предварительное преобразование с использованием уже написанных скриптов необходимо лишь для табличных значений пассажиропотока и пассажирооборота. В качестве примера можно привести ситуацию, когда при вычислении усредненной величины наполнения салона транспортного средства по определенному маршруту необходимо разделить соответствующее значение из таблицы суммарного пассажиропотока на количество дней (к-во будних или выходных дней в расчетном месяце), а также учесть долю пассажиров, не пользующихся электронными проездными билетами. Создаются обработчики клиентских запросов, использующие преобразованные данные для формирования отображаемых данных.

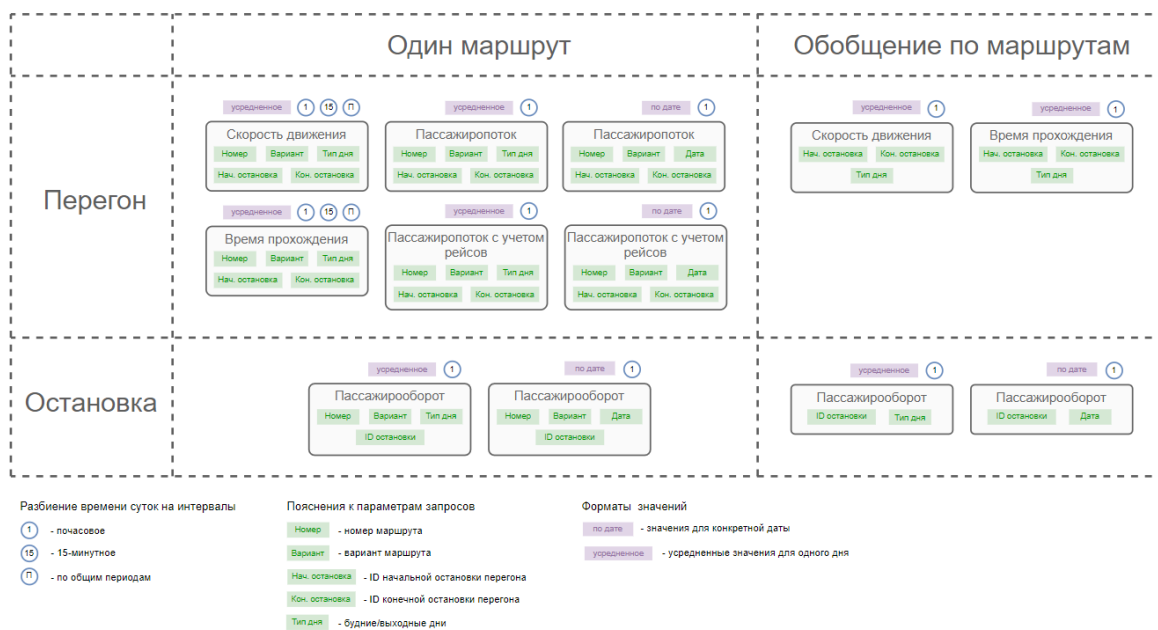


Рис. 13: Характеристики движения транспортных средств и перемещения пассажиров, необходимые для отображения

5.4. Клиентская часть и визуализация

На стороне клиента выполняется построение и отображение объектов транспортной сети, вывод подробных характеристик для этих объектов, а также предоставление сведений в соответствии с пользовательскими запросами. В качестве базового средства визуализации объектов выступает интерактивная двумерная картографическая подложка, при создании которой используются тайлы с ресурса OSM (OpenStreetMap), для отображения же характеристик транспортных потоков используются диаграммы и табличные представления, а также средства цветовой дифференциации и масштабирования при отображении моделей.

Каждая такая модель определяется JavaScript-объектом, для которого заданы свойства с описанием отдельных элементов модели. Построение модели осуществляется в соответствии с описанием её составляющих, каждой из которых берётся в соответствие векторный элемент (полилиния, круговая метка), все такие элементы комбинируются в единый слой, который накладывается на базовый слой подложки. Включение/отключение видимости настоящего слоя осуществляется его до-

бавлением/удалением. Далее приводится обзор моделей для некоторых объектов транспортной системы.

Как уже было отмечено, структура транспортной сети, а также отдельных ее составляющих (маршрутов), задается множеством остановочных пунктов и перегонов, которые их связывают. Для ее моделирования используется представление с помощью графов – перегоны сопоставляются ребрам (полилиниям), а остановочные пункты вершинам (меткам). Для каждого элемента сети (остановочного пункта или перегона) создается специальный объект, в структуре которого описываются как общие сведения об этом элементе (идентификатор, координаты), так и свойства его отображения (цвет, толщина линии или радиус метки) (рис. 14), все эти сведения заносятся в структуру свойств объекта модели транспортной сети или отдельного маршрута и используются при их построении. На рис. 17а приводится пример отображения в приложении модели автобусной транспортной сети, а на рис. 18а модели отдельного маршрута.



Рис. 14: Пример объектов, в которых описываются сведения об остановочных пунктах и перегонах

Интерактивность при работе с данными достигается за счёт возможности выборки как путем символьного ввода, так и непосредственным

выбором графических объектов (через поиск и нажатием). В том случае, если отображается полная модель транспортной сети, при выборе конкретного элемента (перегона или остановочного пункта) выполняется считывание данных с сервера о проходящих через этот элемент маршрутах с последующим предоставлением по запросу пользователя подробной информации о характеристиках движения транспортных средств и перемещении пассажиров на данном участке. При отображении на карте модели отдельного маршрута выбор того или иного его объекта аналогичным образом инициирует считывание соответствующих сведений о наполнении салона, пассажирообороте, скорости и времени движения (рис. 15).

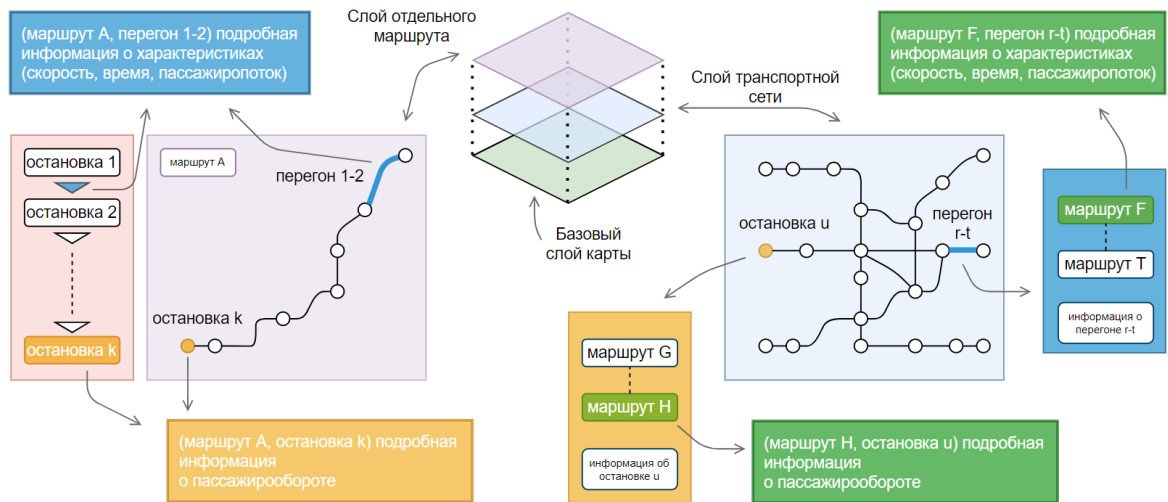


Рис. 15: Демонстрация взаимодействия с отображаемыми на карте моделями транспортных объектов

Подробное описание считываемых характеристик приводится на рис. 13. Пользователю предоставляется возможность указать параметры отображения характеристик, после чего в соответствии с выбранным на карте участком сети и с учетом указанных параметров выполняется построение гистограмм (рис. 17б), на которых демонстрируются значения конкретных характеристики в зависимости от временного промежутка для одного маршрута на этом участке, а также динамика их изменения, и матрицы тепловой карты (рис. 17в), позволяющей делать общий обзор значений характеристики сразу по нескольким маршру-

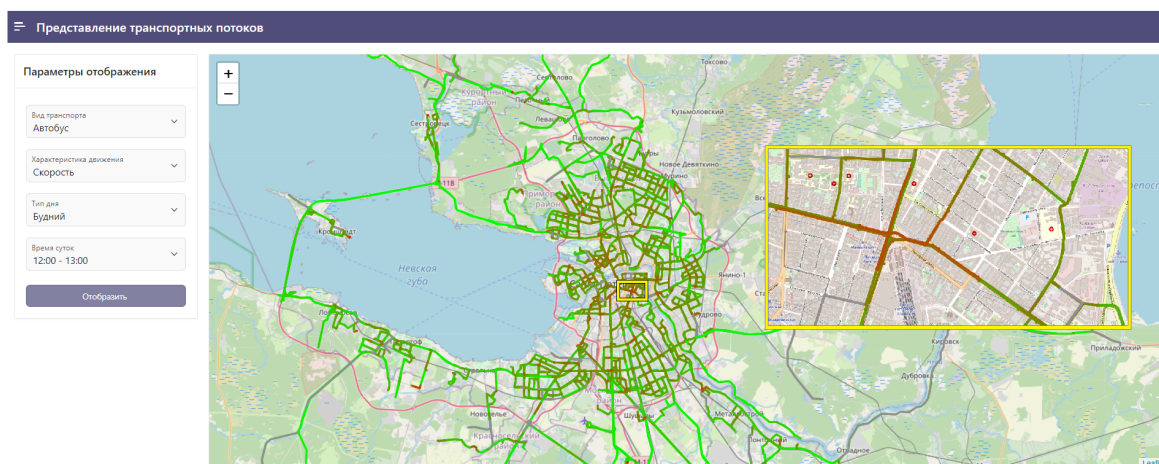
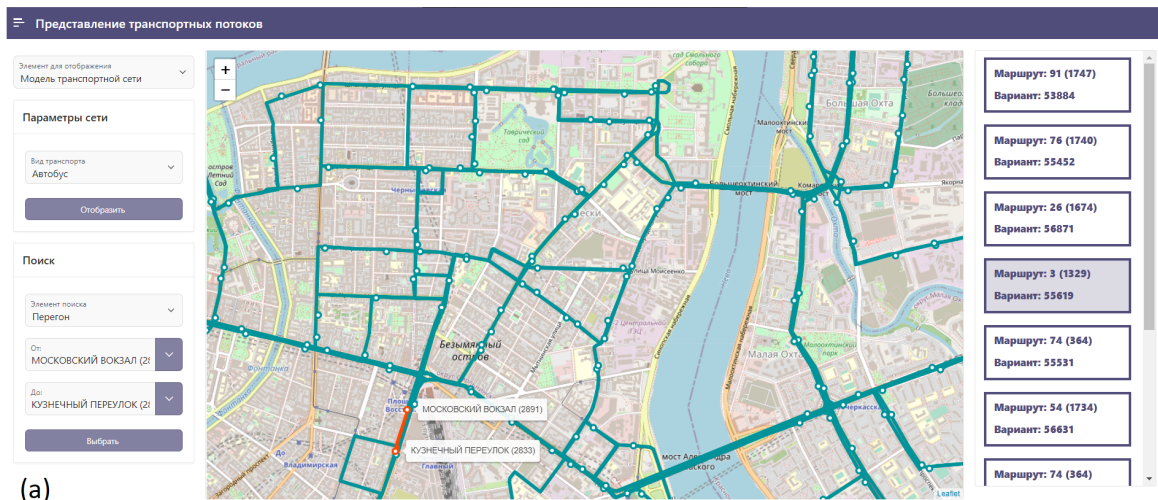


Рис. 16: Интерфейс разрабатываемого приложения, модель автобусной сети с цветовым выделением перегонов (в зависимости от средней скорости движения транспортных средств)

там, дополнительно приводятся соответствующие числовые значения в табличном виде (рис. 18б).

При использовании отображения характеристик(и) объектов методами цветовым выделением, выбранный объект дополняется параметром(ами), определяющим(и) текущие значения задаваемой характеристики (например, средней скорости движения) для различных элементов объекта (например, перегонов). Эти значения используются при окрашивании в заданный цвет заданных элементов отображения. На рис. 16 демонстрируется отображение в приложении цветовой модели для автобусной сети, характеристикой выступает средняя скорость движения транспортных средств в будние дни с 12:00 до 13:00. Подобным образом возможны прочие расширения базового объекта для модели, например, с использованием толщины полилиний линий вместо их цветового выделения, однако в настоящий момент такие возможности в системе не реализованы.

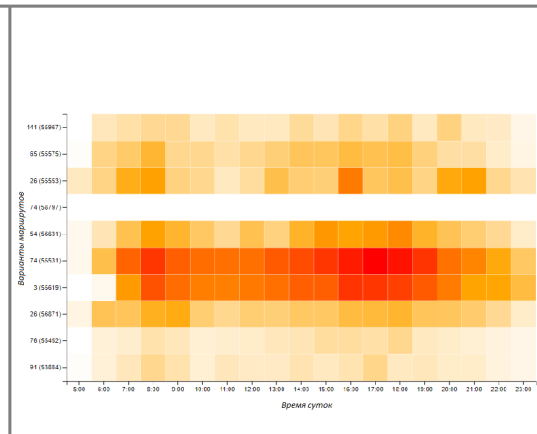
Структурное представление модели межостановочной матрицы корреспонденций задается множеством меток, сопоставляемых всем остановочным пунктам транспортной сети, и множеством линий, каждая из которых соединяет две метки, если между соответствующими им



(a)

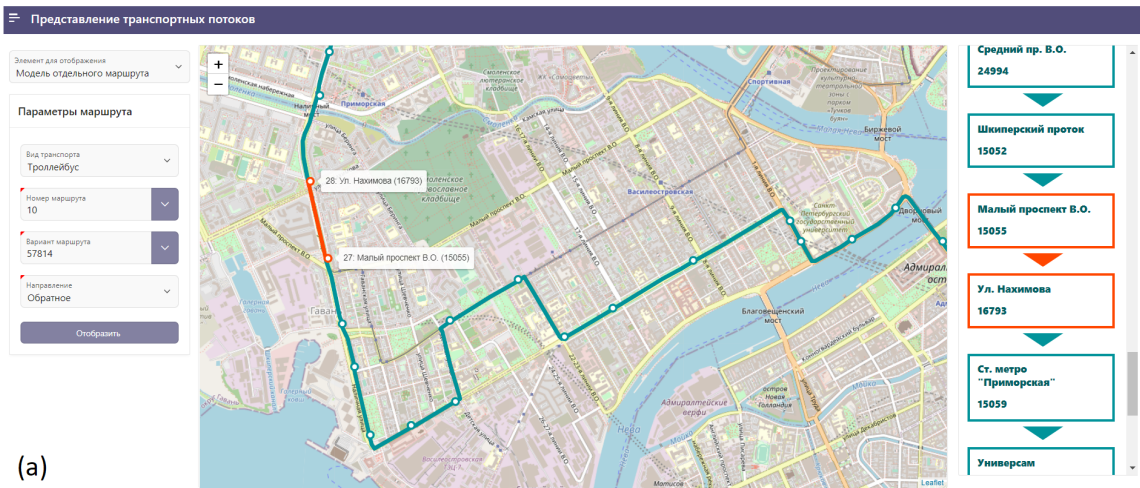


(б)

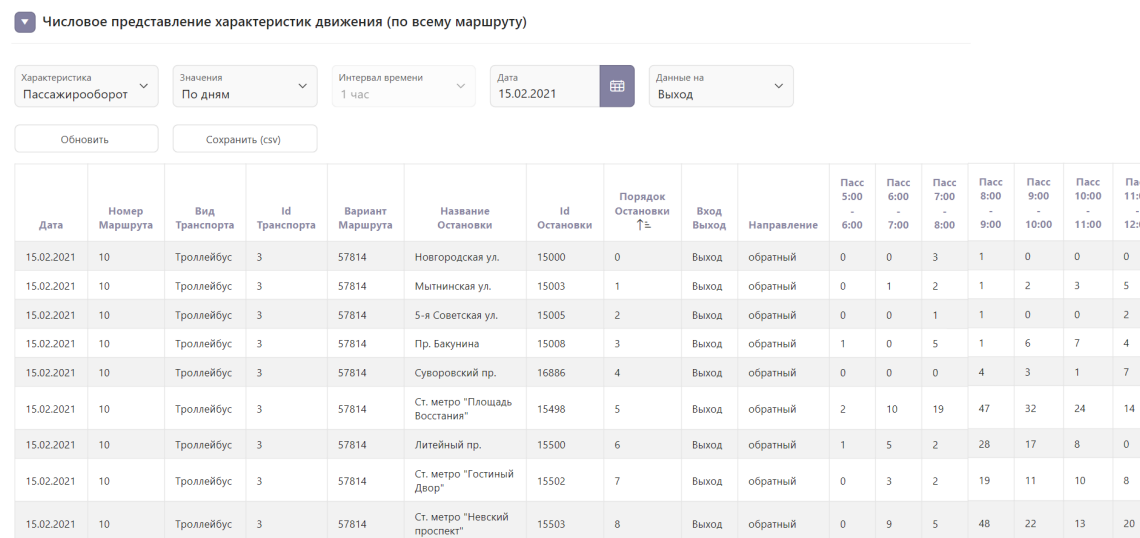


(в)

Рис. 17: Интерфейс разрабатываемого приложения, (а) представление интерактивной модели автобусной транспортной сети, (б) демонстрация с использованием гистограммы величины усредненного пассажиропотока в будние дни для автобусов 3-го маршрута на перегоне от "Московского вокзала" до "Кузнеченого переуллка", (в) матрица тепловой карты, показывающая наполнение салонов транспортных средств сразу для всех маршрутов на этом перегоне, горизонтальная ось определяет время суток, вертикальная – варианты маршрутов, цветом обозначается интенсивность потока



(a)



(б)

Рис. 18: Интерфейс разрабатываемого приложения, (а) представление интерактивной модели отдельного маршрута (на примере 10-ого маршрута троллейбуса с обратным направлением), (б) демонстрация в табличном виде числовых значений для пассажирооборота на выход для этого маршруте за 15 февраля 2021 года

остановочными пунктами имеются пассажирские поездки. Общее число всех таких линий значительное, поэтому для ухода от нагромождения объектов на карте в конкретный момент времени отображаются только некоторые из них, имеющие общую начальную метку, которая задается пользователем. Такое представление сопровождается числовой информацией в табличном виде о количестве соответствующих поездок.

Заключение

В рамках данной дипломной работы было реализовано web-приложение, обеспечивающее быстрый и простой доступ к большим массивам данных, используемым для анализа сети городского пассажирского транспорта Санкт-Петербургской агломерации.

В частности, были решены поставленные задачи:

- Сформированы требования к разрабатываемой системе в соответствии с запросами пользователей;
- Изучены релевантные методы представления данных транспортной сети и дан краткий обзор существующих систем, посредством которых такого рода методы реализуются;
- Выполнен поиск и изучение средств визуализации картографической информации и соответствующих программных решений;
- Реализован комплексный доступ к данным о характеристиках транспортной сети.

Исходные коды были переданы в ведение Учреждения (СПб ГКУ "Организатор перевозок").

В настоящий момент разработанное приложение запущено на сервере Учреждения и используется в тестовом режиме с целью анализа характеристик сети на выявление ошибок.

Предполагается дальнейшее сотрудничество с Учреждением.

Список литературы

- [1] Marchi Massimo De. Visualizing Public Transport Systems. — 2015.
- [2] W. Zeng. Visual Analytics for Massive Urban Public Transport Data. — 2015.
- [3] S. Ram, Y. Wang, F. Currim et al. SMARTBUS: A Web Application for Smart Urban Mobility and Transportation // Proceedings of the 25th International Conference Companion on World Wide Web. — 2016. — P. 363 – 368.
- [4] C. Stewart, E. Diab, R. Bertini. Perspectives on Transit: Potential Benefits of Visualizing Transit Data // Transportation Research Record Journal of the Transportation Research Board. — 2016. — P. 90 – 101.
- [5] Phan Doantam, Xiao Ling, Yeh Ron et al. Flow map layout // IEEE Symposium on Information Visualization. — 2005. — P. 219 – 224.
- [6] K. BUCHIN, B. SPECKMANN, K. VERBEEK. Flow map layout via spiral trees // Visualization and Computer Graphics, IEEE Transactions on 17. — 2011. — P. 2536 – 2544.
- [7] M.G. Demissie, de Almeida Correia G.H., C. Bento. Exploring cellular network handover information for urban mobility analysis // J. Transp. Geogr. — 2013. — P. 164 – 170.
- [8] Chen W., Guo F., Wang F. A Survey of Traffic Data Visualization // IEEE Transactions on Intelligent Transportation Systems. — 2015. — P. 2970 – 2984.
- [9] G. Andrienko, N. Andrienko, P. Bak et al. Visual Analytics of Movement // Visual Analytics Focusing on Spatial Events. — 2013. — P. 209 – 251.

- [10] Pu J., Liu S., Ding Y. et al. T-Watcher: A new visual analytic system for effective traffic surveillance // Proc. IEEE 14th Int. Conf. MDM. — 2013. — P. 127 – 136.
- [11] Y. Song, H.J. Miller. Exploring traffic flow databases using space-time plots and data cubes. Transportation. — 2012.
- [12] W. Zeng, C.W. Fu, S.M. Arisona et al. Visualizing mobility of public transportation system. — 2014.
- [13] T Sobral, T Galvão, J. Borges. Visualization of Urban Mobility Data from Intelligent Transportation Systems. — 2019.
- [14] A Anwar, A Odoni, N. Toh. BusViz: Big Data for Bus Fleets // Transportation Research Record. — 2016. — P. 102 – 109.
- [15] M.O. Khitrin. Comparison of JavaScript libraries for web-cartography // Вестник ЮУрГУ. Серия: Компьютерные технологии, управление, радиоэлектроника. — 2017.
- [16] N.G. Grafeeva, I. Tretyakov, E. Mikhailova. Quality Analysis of Urban Transit System in St. Petersburg // Advances in Intelligent Systems and Computing. — 2018. — P. 889 – 898.