

О моделях пассажирского потока для городских железных дорог

Д.Е. Намиот, О.Н. Покусаев, В.С. Лазуткина

Аннотация— В статье рассматриваются вопросы, связанные с определением (прогнозированием) пассажиропотока для городских железных дорог. Статья представляет собой обзор исследований, связанных с прогнозированием количества пассажиров для новых линий городских железных дорог. Мы рассматриваем источники данных для построения прогноза, а также модели, которые могут быть использованы для получения цифровых оценок. К числу обсуждаемых источников данных относятся данные о миграции, которые могут быть собраны с помощью телекоммуникационных операторов, а также информация об использовании общественного транспорта, получаемая от валидации транспортных карт. Также в работе рассматриваются методы прогноза пассажиропотока, основанные на анализе топологии сети. Последний раздел посвящен визуализации транспортных данных.

Ключевые слова—железная дорога, пассажирский поток, моделирование, Умный город.

I. ВВЕДЕНИЕ

В данной статье рассматриваются вопросы моделирования пассажирского потока для городских железных дорог. Городские железные дороги, как, впрочем, и другие транспортные проекты в городах, являются довольно дорогостоящими предприятиями, которые требуют всесторонних расчетов и обоснований. Основой для расчета социально-экономических эффектов от дороги [1] является именно информация об ожидаемом трафике. Естественно, что именно оценка трафика позволяет оценивать выручку от перевозок, но она также позволяет оценивать и экономию времени пассажиров от введения новых маршрутов, снижение транспортной усталости, достижимость новых районов и т.д. В этой связи упомянем еще раз работу [1], где мы и рассматривали, собственно говоря, различные задачи оценки социально-экономических эффектов.

В данной же работе мы хотели остановиться именно на том, как оценивают (моделируют) этот трафик. Статья представляет собой, в первую очередь, обзор источников данных и моделей данных, используемых в данной области. Идея состоит в том, что в части

моделирования пассажирских потоков уже опубликовано довольно много работ, есть большое количество готовых моделей и, соответственно, для новых проектов типа диаметральных линий железных дорог в Москве, с большой вероятностью, можно использовать какую-либо из уже существующих моделей. Ну или модифицировать ее для конкретных условий. Абсолютно то же самое относится и к сбору данных. Анализ перемещений по данным мобильных операторов выглядит, очевидно, одинаково во всех странах.

Оставшаяся часть работы структурирована следующим образом. В разделе II мы приводим обзор моделей пассажирских потоков. В разделе III рассмотрены системы визуализации данных для задач моделирования потоков.

II. МОДЕЛИ ПАССАЖИРСКИХ ПОТОКОВ

Базовый этап для оценки пассажиропотоков – это данные о миграции. Естественно, что пассажиры – это перемещающиеся жители, какая-то часть которых (именно это и будет оцениваться в модели) воспользуется нашим видом транспорта.

Два основных подхода к оценке миграции. Во-первых, это статистика по опросам (перепись), которую в США собирает статистическая служба. Ответы на опросы обязательны по закону. Эти опросы включают в себя и данные по миграции. Вот, например, в [2] в качестве исходной информации используется собираемая таким образом статистика по Сан-Франциско [3]. В работе [4] можно посмотреть детальную информацию по опросам в Bay Area.

При этом точность информации достаточно высокая. Информация отслеживается на уровне кварталов (их около 4000 в Сан-Франциско). Соответственно, матрица корреспонденции отмечает перемещения из квартала в квартал.

Но данные опросов именно по перемещениям доступны даже в США не во всех местах. В таком случае необходимы какие-то автоматические методы сбора информации о перемещениях. Безусловным лидером здесь являются данные мобильных операторов. Типичные примеры такого рода исследований представлены в работах [5-7]. Данные, которые собираются телекоммуникационными операторами, технически, могут описывать разные аспекты

Статья получена 21 января 2018.

Д.Е. Намиот - МГУ имени М.В. Ломоносова (e-mail: dnamiot@gmail.com).

О.Н. Покусаев - Центр цифровых высокоскоростных транспортных систем РУТ (МИИТ) (email: o.pokusaev@rut.digital)

В.С. Лазуткина - Центр цифровых высокоскоростных транспортных систем РУТ (МИИТ) (email: v.lazutkina@vsmexpert.ru)

функционирования мобильных телефонов в сетях связи. Одно из первых применений этих данных было связано с анализом активности пользователей, измеряемой в Эрлангах (1 эрланг – это использование абонентом сети 1 час).



Рис. 1. Матрица корреспонденции [2]

Это позволяло определять районы активности пользователей в городе (так называемая, социальная динамика). Другие способы сбора параметров (например, Location update, Handover) позволяют прийти, в итоге, к той же самой матрице корреспонденций (начальные и конечные точки в перемещениях мобильного телефона). “Окончание” маршрута может быть определено эмпирически, как некоторый временной интервал, в течение которого мобильное устройство оставалось на месте. В принципе, здесь есть полная аналогия с понятием веб-сессии. Сессия посещения веб-сайта конкретным пользователем также определяется (завершается) по анализу времени неактивности пользователя [8]. В работе [2] использовался другой подход – анализировалось место совершения звонка по телефону и предполагалось, что время с 21:00 до 6:00 соответствует звонку из дома, а время 9:00 до 17:00 – звонку с работы. Это, соответственно, давало один маршрут. В работе [9] авторы используют несколько другой подход. Например, их алгоритм предполагает, что местоположение, из которого пользователь уходит утром и к которому он или она возвращается ночью, является домом. Также предполагается, что местонахождение самого длинного повторяющегося пребывания в будние дни является рабочим местом данного мобильного абонента. В данной работе описывается построение общей модели мобильности в городе Бостон, где по данным операторов установлены пары дом-работа для 75% абонентов, и модель может предсказывать перемещения пользователей с 10 минутным интервалом.

Получив в итоге информацию о перемещениях (миграции), мы должны распределить эти поездки по видам транспорта.

Как первый шаг (он же самый простой), можно

отделить пешеходов. Это, как правило, основано на простых предположениях о длительности маршрута. Если эта длительность (а маршруты начинаются в некотором квартале и заканчиваются в некотором квартале – см. рис. 1) меньше некоторого граничного значения (например, 1 км), можно считать, что это пешеходный маршрут.

Оставшиеся “мигранты” будут либо пассажирами общественного транспорта, либо воспользуются автомобилями. Вопрос с долей автомобилистов в этом потоке является, конечно, дискуссионным. В упоминавшихся [2] американских работах (Бостон, Сан-Франциско) использовались данные статических опросов, где жители прямо должны были отвечать о способах перемещения по городу. Для Москвы такие данные недоступны. Наиболее точным способом определения количества людей, воспользовавшихся автомобилями, будет исключение из числа мигрантов тех, кто воспользовался общественным транспортом. Транспортные карты в Москве (Тройка, например) используются при посадке (входе). Аналог того, что в социальных сетях называется check-in (отметка о присутствии) [10]. Соответственно, из общего числа маршрутов можно вычесть пешеходов и тех, кто отметился в транспорте. С достаточной степенью точности это и даст количество тех, кто воспользовался автомобилем.

Информация об автомобильных поездках может быть и привязана к районам. Если объединить информацию об использовании транспорта в нескольких кварталах (объединение, очевидно, нужно, потому что жители могут использовать остановку транспорта/станцию метро в соседнем квартале), то по разности общего количества поездок и поездок на транспорте можно оценить уровень использования автомобилей в данной маленькой агломерации.

Соответственно, такой подход определяет и требования к информационной системе, на которой можно проводить такие расчеты. Основой должна быть ГИС-система, основной элемент отображения – входные точки (остановки, станции) общественного транспорта. На рисунке 2 изображен пример карты с интеграцией маршрутов наземного транспорта и планируемого железнодорожного диаметра.

Дальнейший подсчет потоков именно для железной дороги (чтобы использовать расчеты социально-экономических эффектов [1]) происходит следующим образом. Мы предполагаем, что пользователи общественного транспорта всегда будут выбирать наиболее оптимальный по времени маршрут. В предположении, что городская железная дорога построена (жирная линия на рисунке 2) мы можем распределить их маршруты и определить тех, кто воспользуется железной дорогой. Это даст трафик и загрузку станций. Также можно подсчитать суммарное время, которое будет затрачено на поездки. Далее, точно

также можно подсчитать время без учета этой дороги, и разность этих времен даст экономию времени пассажиров при наличии железной дороги. Эта экономия (в часах) может быть умножена на среднюю почасовую зарплату и т.д. О деталях расчетов можно посмотреть работу [1].

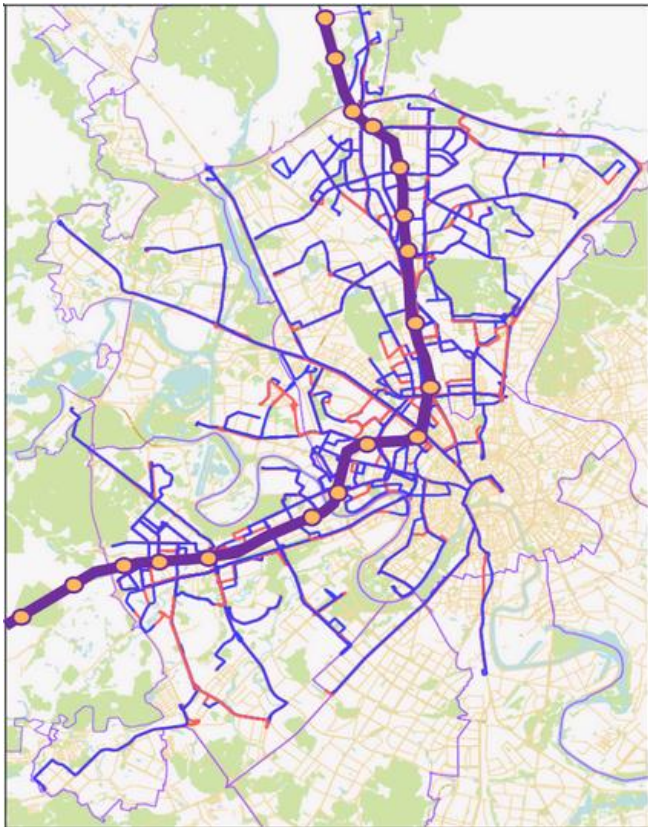


Рис. 2. Маршруты городского транспорта и городская железная дорога.

Следующий важный момент – это получить метрики для пассажирского потока на железной дороге. Вот, например, в работе [2] приводятся характеристики длительности путешествий в Сан-Франциско и Бостоне (в обоих городах есть Urban Rail Transit – URT):

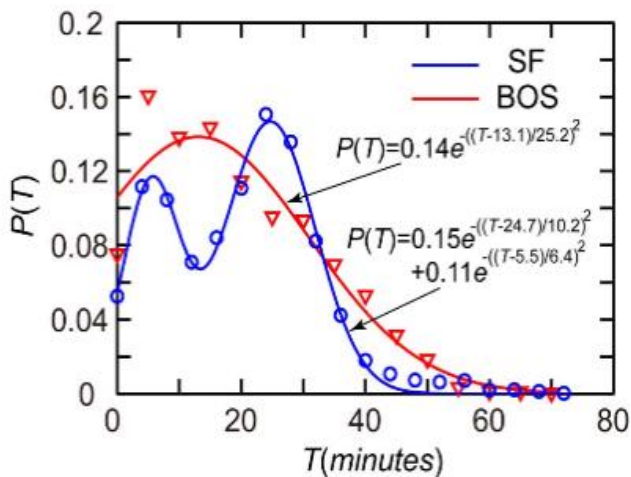


Рис. 3. Распределение поездок по длительности [2]

В качестве следующей метрики может быть

использовано количество наикратчайших маршрутов, проходящих через конкретный сегмент железной дороги. Это показатель важности данного сегмента в транспортной системе города. Или то, насколько чувствительным для города будут проблемы с движением на данном сегменте.

Еще несколько элементов в транспортном моделировании, на которые хотели бы обратить внимание. В работе [11] речь идет о своеобразном моделировании на микро-уровне – как быстро пассажиры могут осуществлять переходы и пересадки на станциях. Это актуально, в частности, для железнодорожных диаметров в Москве, где будет использоваться существующая железнодорожная инфраструктура, а вот станции будут строиться (модифицироваться), и последнее должно делаться, естественно, с учетом анализа пассажиропотоков.

В работе [12] приведен анализ транспортной сети на основе центричности. Первоначально, понятие центричности было впервые введено в работе [13] для измерения важности узла в больших социальных сетях, которые не были полностью связаны. Примером использования этого понятия для анализа транспортных потоков является работа [14]. Идея в том, что структурные свойства сети определяют поток через сеть. Центричность для узла v определяется как:

$$b_c(v) = \sum_{s,t \in V, s \neq t} \frac{\sigma(s \rightarrow t|v)}{\sigma(s \rightarrow t)}$$

где $\sigma(s \rightarrow t|v)$ – число кратчайших путей от станции s до станции t , проходящих через v , а $\sigma(s \rightarrow t)$ – общее число кратчайших путей от s до t .

И согласно исследованию, трафик, проходящий через станцию, коррелирует с ее центричностью. Чем больше кратчайших путей проходит через станцию, тем больше через нее ездит пассажиров. Что дает возможность предсказывать объемы пассажиров, поскольку характеристики узлов могут быть вычислены статически (исходя только из топологии сети).

III. ВИЗУАЛИЗАЦИЯ ТРАНСПОРТНЫХ ПОТОКОВ

Это большая и, естественно, важная тема, которая, возможно, потребует отдельного рассмотрения. Здесь же приведем некоторые примеры того, как это может быть реализовано [15].

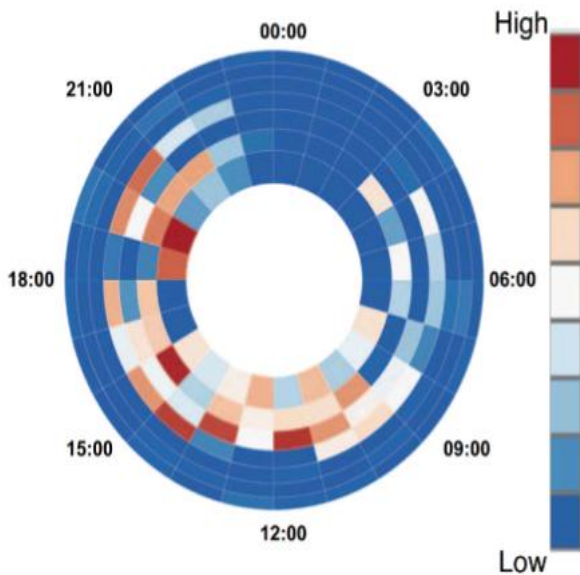


Рис. 4. Отображение периодических значений [15]

Например, рисунок 4 показывает интересную схему иллюстрации периодических событий. Каждое кольцо соответствует одному дню, время дня обозначено по внешнему диаметру, значения (например – загрузка) показано цветом в соответствии с легендой справа. Весьма интересный формат для отображения трафика, например.

На рисунке 5 представлено типичное отображение положения железнодорожных поездов (показаны цветом) на карте дорог [16].

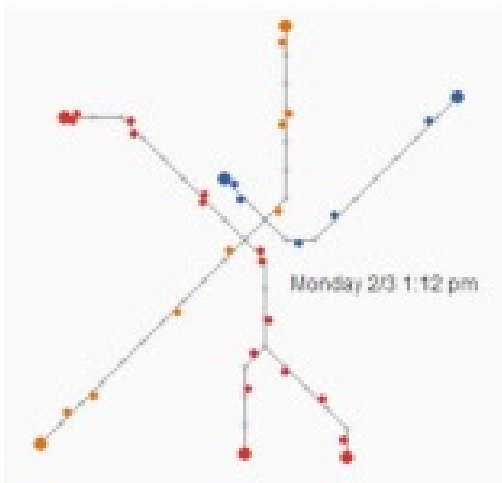


Рис. 5. Текущее положение поездов [17]

А на рисунке 6 представлено параллельное (одновременное с картой) движение поездов между станциями. Ось времени – вертикальна. На этой схеме можно показывать задержки и перерывы в движении.

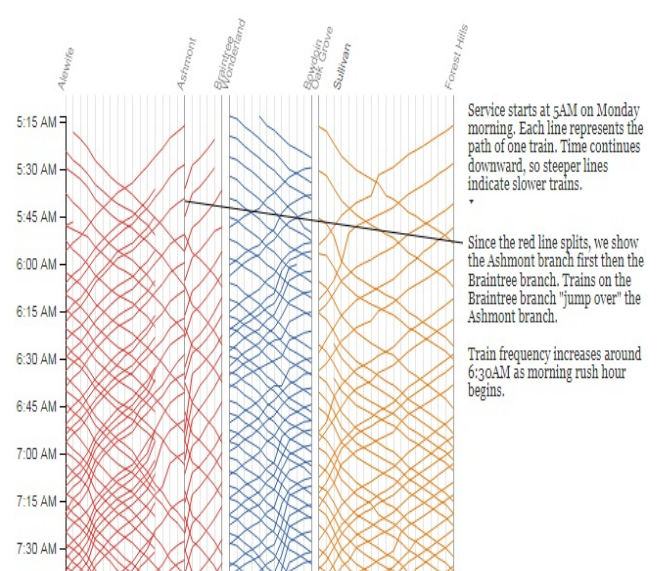


Рис.6. Схема движения [17]

БИБЛИОГРАФИЯ

- [1] Namiot D. et al. On the assessment of socio-economic effects of the city railway //International Journal of Open Information Technologies. – 2018. – Т. 6. – №. 1. – С. 92-103.
- [2] Wang J. et al. Vulnerability analysis and passenger source prediction in urban rail transit networks //PloS one. – 2013. – Т. 8. – №. 11. – С. e80178.
- [3] US Census Bureau. Available: <http://www.census.gov/geo/www/tiger/tgrshp2010/tgrshp2010.html>. Retrieved: Jan, 2018
- [4] Bay Area Census <http://www.bayareacensus.ca.gov/faq.htm>. Retrieved: Jan, 2018
- [5] Steenbruggen J. et al. Mobile phone data from GSM networks for traffic parameter and urban spatial pattern assessment: a review of applications and opportunities //GeoJournal. – 2013. – Т. 78. – №. 2. – С. 223-243.
- [6] Ratti C. et al. Mobile landscapes: using location data from cell phones for urban analysis //Environment and Planning B: Planning and Design. – 2006. – Т. 33. – №. 5. – С. 727-748.
- [7] Calabrese F. et al. Real-time urban monitoring using cell phones: A case study in Rome //IEEE Transactions on Intelligent Transportation Systems. – 2011. – Т. 12. – №. 1. – С. 141-151.
- [8] Namiot D., Sneps-Sneppe M. On the analysis of statistics of mobile visitors //Automatic Control and Computer Sciences. – 2014. – Т. 48. – №. 3. – С. 150-158.
- [9] Ford, MIT use Bostonians' cellphone location data for traffic planning <https://www.computerworld.com/article/3112845/cartech/ford-mit-use-bostonians-cellphone-location-data-for-city-traffic-planning.html> Retrieved: Jan, 2018
- [10] Namiot D., Sneps-Sneppe M. Customized check-in procedures //Smart Spaces and Next Generation Wired/Wireless Networking. – 2011. – С. 160-164.
- [11] Zhou Y. et al. Time prediction model of subway transfer //SpringerPlus. – 2016. – Т. 5. – №. 1. – С. 44.
- [12] Ramli M. A. et al. A method to ascertain rapid transit systems' throughput distribution using network analysis //Procedia Computer Science. – 2014. – Т. 29. – С. 1621-1630.
- [13] Freeman L. C. A set of measures of centrality based on betweenness //Sociometry. – 1977. – С. 35-41.
- [14] Altshuler Y. et al. Augmented betweenness centrality for mobility prediction in transportation networks //International Workshop on Finding Patterns of Human Behaviors in NETworks and MObility Data, NEMO11. – 2011.

- [15] Chen W., Guo F., Wang F. Y. A survey of traffic data visualization //IEEE Transactions on Intelligent Transportation Systems. – 2015. – T. 16. – №. 6. – С. 2970-2984
- [16] K. Kloeckl, X. Chen, C. Sommer, C. Ratti, and A. Biderman, “Trains of data.” Available: <http://senseable.mit.edu/trainsofdata/>
- [17] Mike Barry and Brian Card Visualizing MBTA Data. An interactive exploration of Boston's subway system <http://mbtaviz.github.io/>
Retrieved: Jan, 2018

On passenger flow data models for urban railways

Dmitry Namiot, Oleg Pokusaev, Varvara Lazutkina

Abstract— The article deals with issues related to the calculation (forecasting) of passenger traffic for urban railways. The article is a survey of research related to forecasting the number of passengers for new lines of urban railways. We consider the data sources for building the forecast, as well as models that can be used to obtain digital estimates. The sources of data discussed in this paper include migration data that can be collected using telecommunications operators, as well as information on the use of public transport, obtained from the validation of transport cards. Also, the paper considers the methods of forecasting passenger traffic, based on the analysis of network topology. The last section is devoted to the visualization of transport data.

Keywords— urban railway, passenger flow, modeling, Smart City.