

УДК 656.2

DOI: 10.25559/SITITO.15.201902.375-385

Об анализе пассажиропотоков Московского метрополитена

Ф. С. Поматилов*, Д. Е. Намиот

Московский государственный университет имени М.В. Ломоносова, г. Москва, Россия
119991, Россия, г. Москва, ГСП-1, Ленинские горы, д. 1

*pomatilov@gmail.com

Аннотация

В работе рассматривается модель анализа пассажиропотоков Московского метрополитена. Исходной информацией для анализа выступает матрица корреспонденции, которая описывает количество пассажиров, перемещавшихся между парами станций метро за указанный временной интервал. Такого рода данные собираются на основе информации мобильных операторов. Такого рода информация является широко используемым подходом к представлению транспортных данных. Статья содержит обзор моделей анализа транспортных данных. В работе предложен подход к классификации станций метрополитена по типу пассажиропотока. В рамках предложенного подхода сравниваются входы и выходы пассажиров для станций. Именно соотношение этих параметров, развернутое по времени и является основанием для классификации станций (отнесением станций к одному из классов, различающихся характером соотношения входов и выходов пассажиров). Это позволяет, например, определить рабочие и жилые зоны города, как станции, куда приезжают на работу, и станции, откуда едут на работу. Такой анализ позволяет определить станции, которые являются местом приема внешнего пассажиропотока. В каких-то случаях это будет объяснимо наличием рядом железнодорожных вокзалов и транспортно-пересадочных узлов, в каких-то случаях это будет свидетельством наличия каких-то относительно неярких (неизвестных) процессов в городе (например, автобусные линии из пригородов). Также изменения в найденных закономерностях могут служить свидетельством каких-то изменений в жизни города.

Ключевые слова: матрица корреспонденции, анализ пассажиропотоков, метро.

Для цитирования: Поматилов Ф. С., Намиот Д. Е. Об анализе пассажиропотоков Московского метрополитена // Современные информационные технологии и ИТ-образование. 2019. Т. 15, № 2. С. 375-385. DOI: 10.25559/SITITO.15.201902.375-385

© Поматилов Ф. С., Намиот Д. Е., 2019



Контент доступен под лицензией Creative Commons Attribution 4.0 License.
The content is available under Creative Commons Attribution 4.0 License.



On the Analysis of Moscow Metro Passenger Flows

F. S. Pomatilov*, D. E. Namiot

Lomonosov Moscow State University, Moscow, Russia

1, Leninskie gory, Moscow 119991, Russia

*pomatilov@gmail.com

Abstract

The paper considers the model of analysis of passenger traffic of the Moscow Metro. The initial information for the analysis is the correspondence matrix describing the number of passengers traveling between pairs of metro stations for a specified time interval. This kind of data is collected on the basis of information from mobile operators. This kind of information is a widely used approach to the presentation of transport data. The article contains a detailed review of transport data analysis models. The paper proposes an approach to the classification of metro stations by type of passenger traffic. The proposed approach compares the entrances and exits of passengers for stations. It is the ratio of these parameters, developed in time, presents the basis for the classification of stations (assigning stations to one of the classes that differ in the nature of the ratio of passenger entry and exit). This allows us, for example, to determine the working and residential areas of the city, as stations where they come to work, and stations from where they go to work. Such an analysis makes it possible to determine the stations that are the receiving point of the external passenger traffic. In some cases, this will be explained by the presence of a number of railway stations and transport hubs, in some cases this will be evidence of the presence of some relatively implicit (unknown) processes in the city (for example, bus lines from the suburbs). Also, changes in the found patterns can serve as evidence of some changes in the life of the city.

Keywords: origin-destination matrix, passenger flow analysis, metro.

For citation: Pomatilov F.S., Namiot D.E. On the Analysis of Moscow Metro Passenger Flows. *Sovremennye informacionnye tehnologii i IT-obrazovanie* = Modern Information Technologies and IT-Education. 2019; 15(2):375-385. DOI: 10.25559/SITITO.15.201902.375-385



Введение

Транспортная составляющая (мобильность, умная мобильность) – важная составная часть Умного Города [1]. Соответственно, анализу транспортных проектов уделяется большое внимание [2]. В условиях Москвы метрополитен является, конечно, основным видом транспорта. В статье рассматривается ряд моделей, которые позволяют анализировать пассажиропоток в Московском метрополитене. В качестве исходной информации рассматривается так называемая матрица корреспонденции (в англоязычной литературе она называется, обычно, origin-destination matrix или OD matrix [3, 17]). Это достаточно широко распространенный подход к описанию транспортных потоков. В такой матрице представлены поездки (количество пассажиров) для пар станций. Матрица существует для определенного временного среза (например, 15 минут, 1 час, 1 день) и показывает количество человек, которые отправились с одной станции на другую в течение соответствующего временного интервала. Теоретически, такая информация полностью описывает пассажиропоток.

Существуют различные способы получения такой информации. Например, такого рода данные могут собираться по информации об использовании проездных карт (например, карта Тройка в Москве) [4] или по информации полученной от мобильных операторов [21]. В данном случае источником данных послужила информация мобильных операторов. В московском метро работает мобильная связь. Соответственно, мобильные абоненты (пассажиры), находящиеся в метро обслуживаются станциями сотовой связи, находящимися в метро. Также, мобильный оператор может определять моменты переключения пользователей на обслуживание станциями в метро и обратно. Таким образом, мобильный оператор может полностью построить матрицу корреспонденции, зная станции входа и выхода. Естественно, что данные полностью анонимны, оператор указывает просто количество человек, начавших в заданный временной интервал поездку от одной станции к другой. При времени от 15 минут восстановить отдельные поездки по таким данным нельзя.

Исходные данные представлены в виде .CSV файла и справочника с названиями станций. В целях анализа данные были загружены в базу MySQL. Таблицы для анализируемых данных выглядят следующим образом:

Станции:

```
CREATE TABLE `stations` (
  `id` TINYINT(3) UNSIGNED NULL DEFAULT NULL,
  `name` CHAR(50) NULL DEFAULT NULL,
  `line` CHAR(10) NULL DEFAULT NULL
)
COLLATE='utf8_general_ci'
```

ENGINE=InnoDB;

id - для идентификационного номера станции;

name - название станции;

line – название линии метрополитена, на которой находится соответствующая станция.

Маршруты движения:

```
CREATE TABLE `routing` (
  `ts` DATETIME NULL DEFAULT NULL,
  `departure_zid` TINYINT
(3) UNSIGNED NULL DEFAULT NULL,
  `arrival_zid` TINYINT(3) UNSIGNED NULL DEFAULT NULL,
  `customers_cnt` SMALLINT(5) UNSIGNED NULL DEFAULT
NULL
)
COLLATE='utf8_general_ci'
ENGINE=InnoDB;
```

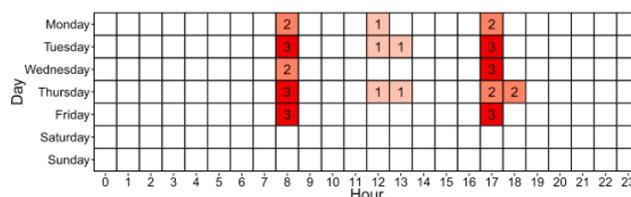
ts - содержит информацию о дате и времени отправления;

departure_zid – идентификационный номер станции отправления;

arrival_zid - идентификационный номер станции прибытия;

customers_cnt – число пассажиров, отправившихся от станции с *departure_zid* на станцию с *arrival_zid* за период времени *ts*.

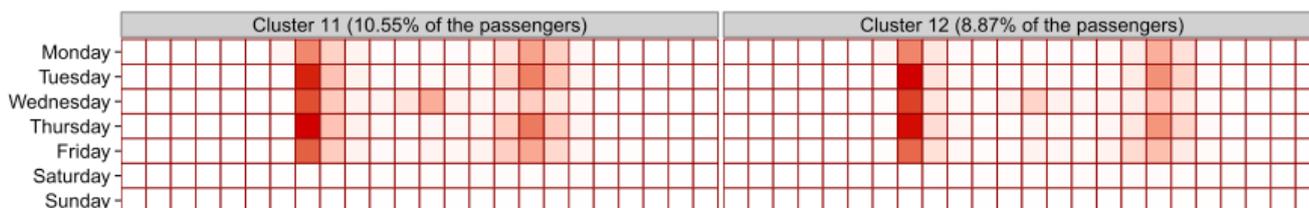
Исследовались часовые срезы за период 01.06.2016 – 30.06.2016. Основой для анализа служила развитая система SQL скриптов. Сама статья написана по результатам выпускной квалификационной работы, выполненной в лаборатории ОИТ факультета ВМК МГУ имени М.В. Ломоносова.



Р и с. 1. Временной профиль

F i g. 1. Time profile

После этого, используя временные профили пассажиров, их делят на 16 кластеров. Примеры изображения кластеров показаны на рисунке 2. Авторы делают выводы по каждому из кластеров. Например, на рисунке 2 изображены кластеры пассажиров, работающих по графику 5/2. Характерной особенностью для этих кластеров также является то, что время утреннего отправления на работу выражено четко и однозначно, при этом, время возвращения с работы несколько размыто.



Р и с. 2. Кластеры пассажиров

F i g. 2. Passenger clusters



Кроме того, рассмотрены социоэкономические кластеры (таблица 5.2). Как видно, разбиение осуществляется на основе общей площади домохозяйства, подушного дохода и числа жителей.

Таблица 1. Социоэкономические кластеры
Table 1. Socioeconomic Clusters

Cluster	Area (m ²)	Income (€)	Population
Collective housing, Medium+ income	15696 (7763)	21889 (1791)	392 (194)
Collective housing, Low income	12510 (7553)	14499 (2364)	397 (244)
Individual housing, High density	8495 (2452)	20804 (2502)	239 (71)
Individual housing, High income	4359 (1847)	24422 (1899)	99 (45)
Individual housing, Medium inc. & dens.	3721 (1576)	20761 (2379)	103 (45)
Individual housing, Low density	695 (337)	21655 (2794)	17 (9)
Individual housing, Low density	167 (85)	21361 (2749)	4 (2)

N.B. standard deviation is indicated in parentheses.

На рисунке 3 представлено распределение социоэкономического кластера на временной кластер.

В дополнение к вышеупомянутым характеристикам были рассмотрены типы билетов, популярные у тех или иных временных или социоэкономических кластеров. В результате авторам удалось понять, какие социоэкономические кластеры чаще прочих пользуются общественным транспортом, какие типы билетов пользуются популярностью в определенных районах города и т.д. Авторы исследовали только пассажиров автобусов и метро, поэтому как направление дальнейших исследований они рассматривают изучение связки этих видов транспорта с городской системой проката велосипедов.

Авторы работы [6] рассматривают вопрос о давке в часы пик на станциях Лондонского метро. Не секрет, что основной пас-

сажиропоток рабочего дня направляется утром на работу, а вечером с работы домой, именно в это время транспортная система испытывает пиковую нагрузку. Авторы ссылаются на то, что большинство систем, прокладываящих маршрут через метро, не учитывают такую важную составляющую поездки как комфорт в часы пик. На рисунке 4 показано недельное распределение поездок во всей сети метрополитена: среднее значение входов на станцию, выходов со станции.

Переходя к примерам, авторы рассматривают типичные картины будничных пассажиропотоков входов/выходов для станции, находящейся в отдаленном от центра, так называемом «спальном районе» - Финчли Сентрал (рисунк 5) и станцию в центре города - Канэри-Уорф (рисунк 6). На каждом рисунке присутствуют утренние и вечерние пики. Наблюдается следующий шаблон: основной утренний пассажиропоток направлен из «спального района» в центр города, а вечерний в обратную сторону.

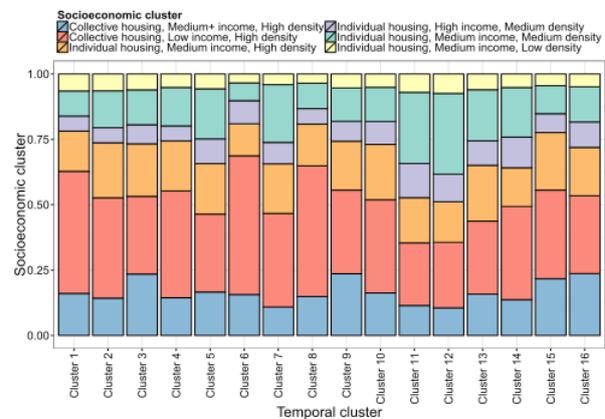


Рис. 3. Пропорции кластеров
Fig. 3. Cluster proportions

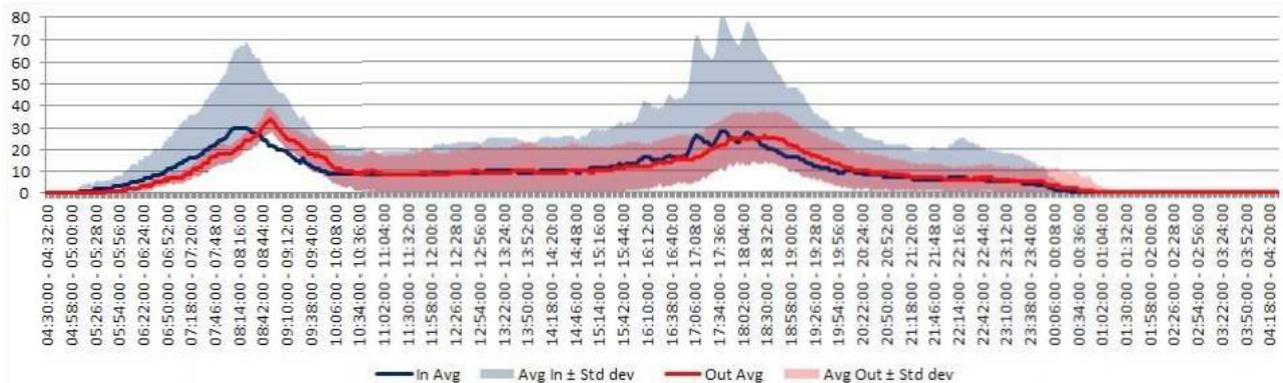


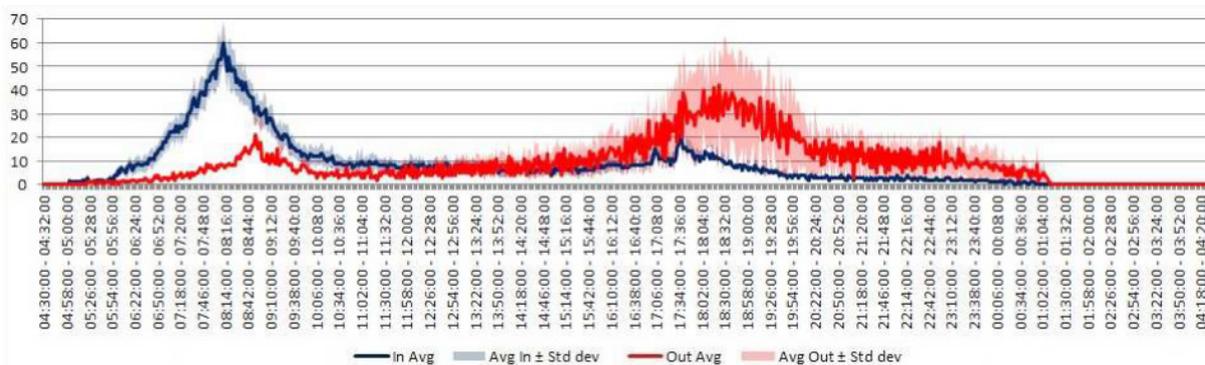
Рис. 4. Недельное распределение поездок
Fig. 4. Weekly travel distribution

Далее авторы объединяют станции в кластеры, основываясь на похожести направлений их пассажиропотоков. На рисунке 7 представлено распределение кластеров станций. На вертикальной оси представлены коэффициенты вычисленные путем деления количества входов/выходов в конкретное время, на максимальное количество входов/выходов соответственно, которые были зафиксированы за время наблюдений. Отрицательные значения означают, что станции используются, преимущественно, на вход, а положительные, что станция исполь-

зуется преимущественно на выход.

В работе [7] авторы имеют дело с краткосрочным прогнозом пассажиропотока. Они собирают информацию об отправителе и пункте назначения из данных на смарт-карте, а затем исследуют схемы пассажиропотока в системе метро. Они изучают закономерности притока и оттока на каждой станции со средними значениями в каждом временном интервале. Авторы исследовали комбинацию входных и выходных потоков, а также время их появления. Согласно их исследованию, станции мо-





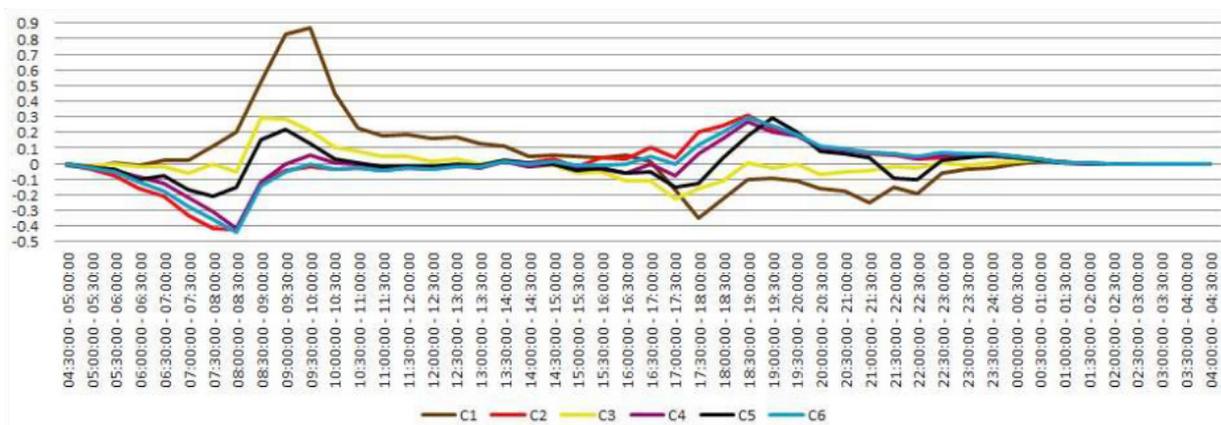
Р и с. 5. Распределение пассажиропотоков на станции Финчли Централ

F i g. 5. Passenger traffic distribution at Finchley Central Station



Р и с. 6. Распределение пассажиропотоков на станции Канэри-Уорф

F i g. 6. Passenger traffic distribution at Canary Wharf station



Р и с. 7. Кластеризация станций

F i g. 7. Station Clustering

гут быть разделены на следующие четыре группы в соответствии с их схемами потока:

1. Станции, которые имеют двухпиковую схему как притока, так и оттока в будние дни. Наиболее очевидное объяснение состоит в том, что эти станции расположены в зонах, которые служат как жилыми, так и рабочими. Количество людей, покидающих эти зоны, сопоставимо с количеством людей, прибывающих в эти зоны в часы пик.
2. Станции, которые имеют только один пик притока в утренний час пик и один пик оттока в вечерний час пик. Это типично

для станций, расположенных в жилых районах. Пассажиры отправляются из дома на работу в час пик и возвращаются домой в час пик.

3. Станции с утренним пиком на оттоке и вечерним пиком на притоке. Эти станции находятся в нежилых зонах, в деловых или промышленных зонах. Пассажиры работают здесь.

4. Станции, имеющие пики в полдень как притока, так и оттока. Такие станции могут быть расположены в популярном коммерческом районе. Это трафик для покупок.

Мы находим эту конфигурацию довольно реалистичной. В [7]



авторы также предлагают основанную на данных структуру для краткосрочного прогнозирования пассажиропотока в метро с возможностью использования как пространственной, так и временной информации. Для пространственных данных ключевым моментом является ожидаемая длина поездок (статистическое распределение длины поездок). Подобные задачи обсуждаются также в работах [8] и [9].

Интересным вопросом является закономерность в схемах путешествий. Согласно [10], выявление закономерностей в мобильности людей имеет основополагающее значение для нашего понимания динамики городов и имеет важное значение для городского и транспортного планирования, городского управления и выработки политики. В этой статье рассматриваются два критических аспекта моделей мобильности, а именно: временное распределение времени начала поездок на станциях и шаблон потоков для входов и выходов, в разные моменты времени, взятые из матрицы OD. Разнообразие, которое получают эти данные, показывает важные различия между городами. Подобные вопросы обсуждаются также в работах [11] и [12].

Одним из трендов концепции Индустрия 4.0 является переход от владения к аренде, в том числе, использованию разделяемых индивидуальных транспортных средств (автомобилей, велосипедов и т.д.). Авторы работы [18] анализируют влияние разделяемых велосипедов, которым не требуется специальная парковка (dockless bikes) на увеличение области охвата обслуживанием станции.

Направление пассажиропотоков, как и транспортных потоков в принципе, является важным предметом исследования в урбанистике. Авторы работы [19] рассматривают факторы, которые заставляют людей совершать поездки: социальноэкономические (средний возраст, уровень безработицы, число домовладений, число владельцев авто и т.д.), земельные (плотность населения, плотность рабочих мест и т.д.), доступность (расстояние до станции), транспортной системы (плотность станций метро, число автобусных остановок). Подобные вопросы рассмотрены в работе [20].

Основной областью применения результатов анализа пассажиропотоков являются различного рода оптимизации: оптимизация выбора маршрутов [13], оптимизация пересадочных пассажиропотоков [14], оптимизация синхронизации движения поездов пересекающихся линий метро [15], оптимизация расписания движения поездов с учетом пассажиропотока и стоимости электроэнергии в зависимости от времени суток [16].

Пассажиропотоки относительно станций

Рассмотрим, как распределены пассажиропотоки в разное время дня. На рисунке 8 представлен суммарный пассажиропоток во всем метрополитене в будние дни, сгруппированный по часам. Видно, что с момента открытия метрополитена пассажиропоток возрастает – сотрудники едут к своим рабочим местам, а студенты на учебу. Далее пассажиропоток уменьшается, но после 12 часов дня начинает расти и достигает пика с 18 до 19 часов.



Р и с. 8. Пассажиропоток в будние дни

Fig. 8. Passenger traffic on weekdays

Важной характеристикой пассажиропотока является баланс количества приехавших на станцию и уехавших со станции пассажиров. Для того чтобы лучше понять баланс пассажиропотоков на станции введем коэффициент перекоса пассажиропотока (далее КПП), который вычисляется по формуле 1:

$$КПП = \frac{(П - У)}{П} \times 100\% \quad (1)$$

где,

P – число приехавших пассажиров за единицу времени,

U – число уехавших пассажиров за единицу времени

Число приехавших пассажиров подразумевает число пассажиров, уехавших с других станций и приехавших на целевую. Число уехавших пассажиров подразумевает число уехавших с целевой станции. $КПП$ позволяет понять, на сколько процентов число приехавших на станцию пассажиров превосходит число уехавших за некоторый период времени. Отрицательное значение $КПП$ свидетельствует о том, что число уехавших пассажиров превосходит число приехавших. $КПП$ удобен для относительной оценки преобладающего пассажиропотока, для абсолютной оценки удобно пользоваться значением разницы пассажиропотоков (далее РП), которую легко вычислить по формуле 2:

$$РП = П - У \quad (2)$$

где,

P – число приехавших пассажиров за единицу времени,

U – число уехавших пассажиров за единицу времени,

Значение $РП$ позволяет понять, на сколько человек за единицу времени отличались пассажиропотоки на станцию / со станции. Как и для $КПП$, знак показывает направление преобладающего пассажиропотока: положительное значение свидетельствует о преобладании числа приехавших пассажиров на станцию над числом уехавших со станции, в то время как отрицательное значение свидетельствует об обратном.

На рисунке 8 показано, что пассажиропоток в будние дни имеет утренний и вечерний пики, в 8 и 18 часов соответственно. Для того чтобы провести кластеризацию, введем признак станции – группу станции. Признак будет основан на комбинации значений $КПП$ в пиковые часы утром и вечером, а именно на знаке («+» или «-») значения $КПП$. Количество размещений с повторениями для двух значений равно четырём. В таблице 2.1 представлены комбинации, номера групп и краткая характеристика, помогающая понять суть группы. Буква «u» или «v» после $КПП$ обозначает значение в утренний или вечерний час пик соответственно.



Т а б л и ц а 2. Комбинации КПП и соответствующие им группы

Комбинация КПП	Номер группы	Характеристика группы
КПП _у < 0 & КПП _в < 0	0	«Полюс отъездов»
КПП _у > 0 & КПП _в < 0	1	«Рабочая зона»
КПП _у < 0 & КПП _в > 0	2	«Жилая зона»
КПП _у > 0 & КПП _в > 0	3	«Полюс приездов»

Группа 0 – «Полюс отъездов» - представляет станции, на которых КПП отрицательный как в утренний, так и в вечерний пик, следовательно, количество уехавших пассажиров превышает количество приехавших.

Группа 1 – «Рабочая зона» - представляет станции, у которых утром пассажиропоток направлен на станцию, а вечером со станции.

Группа 2 – «Жилая зона» - представляет станции, у которых утром пассажиропоток направлен от станции, а вечером на станцию.

Группа 3 – «Полюс приездов» - представляет станции, на которых КПП в пиковые часы постоянно больше нуля.

Следует отметить, что приведенная выше группировка сделана из предположения, что большинство пассажиров по будням в период с 8 до 9 часов утра едут к рабочим местам, а с 18 до 19 часов едут обратно. Значения КПП и РП равные нулю приняты условно невероятными. Станции, относящиеся к разным группам проиллюстрированы на следующих рисунках.

В первую группу (рисунок 10), ожидаемо, вошли станции, большинство из которых расположено ближе к центру города (Кропоткинская, Кузнецкий Мост, Чистые пруды и т.д.), или имеет в непосредственной близости офисные центры или производственные организации (Шоссе Энтузиастов и др.)



Р и с. 10. Станции группы 1
F i g. 10. Group 1 stations



Р и с. 9. Станции группы 0
F i g. 9. Group 0 stations



Р и с. 11. Станции группы 2
F i g. 11. Group 2 stations

Во вторую группу (рисунок 11), в большинстве, вошли станции находящиеся ближе к периферии – «спальные районы». Интересно отметить, что некоторые станций, расположенные в непосредственной близости от вокзалов: Белорусского, Савеловского, Павелецкого входят в группу 1, хотя можно было предположить, что привокзальные станции будут иметь ре-



жим использования похожий на станции «жилой зоны», так как именно с них пассажиры попадают в систему метрополитена, а потом возвращаются на электричках домой. Такое распределение говорит о том, что количество рабочих мест в районе этих станций выше, чем количество пассажиров, которые приезжают на вокзалы и потом пользуются для проезда метро. При этом станции Комсомольская Кольцевой и Сокольнической линий находятся во второй группе, так как там расположено три вокзала, и их суммарный пассажиропоток превышает количество рабочих мест вблизи этих станций.



Р и с. 12. Станции группы 3
F i g. 12. Group 3 stations

Особый интерес представляют станции групп 0 (рисунок 9) и 3 (рисунок 12). Например, станция Киевская Арбатско-Покровской линии, в основном, используется для того, чтобы уехать, в то время как одноименные станции Филевской и Кольцевой линий используются для приезда. Станции Римская и площадь Ильича имеют переход и находятся в непосредственной близости от железнодорожной платформы Серп и молот. При этом станции имеют диаметрально противоположный режим использования: со станции Площадь Ильича пассажиры обычно уезжают, а на станцию Римская приезжают. Похожая картина со станцией Парк Победы. В третью группу вошли станции рекреации: Партизанская, ВДНХ. В целом, можно сказать, что распределение станций сходится с ожидаемым.

Далее вычислим средние значения КПП и РП для каждой из групп в 8 и 18 часов. Данные, полученные в результате запроса, представлены в таблице 3. Для группы 0 характерен высокий модуль значений КПП утром, что говорит о значительном преобладании пассажиропотока от станции над пассажиропотоком на станцию, вечером КПП не достигает таких значений. Для группы 1, утренний пассажиропоток на станции, в среднем, превышает пассажиропоток от станций на 65%, в то вре-

мя как в вечерний пик, отъезд пассажиров превышает прибытие на 150%. Станции группы 2 имеют утром, в среднем, значительное преобладание пассажиропотока от станции - почти 250%, а вечером пассажиропоток на станцию преобладает на 45%. Станции группы 3 имеют незначительное преобладание как в утренний, так и вечерний пик.

Т а б л и ц а 3. Средние значения РП и КПП
T a b l e 3. Average RP and KPP

Группа	Время дня	Средняя РП	Средний КПП
0	8	-43101,3	-209,674
0	18	-33012,6	-77,1042
1	8	85197,42	65,15854
1	18	-69207,3	-150,071
2	8	-97988,6	-245,747
2	18	88847,07	45,03324
3	8	56409,77	40,24792
3	18	33055,08	17,09725

Рассмотрим вопрос распределение станций разных линий по группам, а именно вычислим процент станций линии, принадлежащий той или иной группе.

Т а б л и ц а 4. Линии группы 0
T a b l e 4. Group lines 0

% от всех станций линии	Всего станций	Из них в группе	Линия
22,7273	22	5	АПЛ
14,2857	7	1	БутЛ
9,0909	22	2	ЗамЛ
-	10	-	КалЛ
-	3	-	КахЛ
-	12	-	КолЛ
8,3333	24	2	КРЛ
10	20	2	ЛДЛ
13,6364	22	3	СокЛ
4	25	1	СТЛ
8,6957	23	2	ТКЛ
53,8462	13	7	ФилЛ

Прочерк означает, что станции соответствующей линии не представлены в группе. Первая колонка показывает, какой процент от станций линии содержится в этой группе, вторая - общее количество станций линии, третья - число станций, принадлежащих группе. Ниже, в таблицах 5, 6, 7 представлены результаты для групп 1, 2 и 3 соответственно.

Станции группы 0 составляют незначительный процент от станций всех линий, кроме Филевской. Можно сказать, что больше половины станций Филевской линии принадлежат этой группе. На трех линиях: Калининской, Каховской и Кольцевой нет ни одной станции, принадлежащей к группе 0.



Таблица 5. Линии группы 1

Table 5. Group lines 1

% от всех станций линии	Всего станций	Из них в группе	Линия
27,2727	22	6	АПЛ
-	7	-	БутЛ
45,4545	22	10	ЗамЛ
50	10	5	КалЛ
33,3333	3	1	КахЛ
75	12	9	КолЛ
50	24	12	КРЛ
35	20	7	ЛДЛ
54,5455	22	12	СокЛ
40	25	10	СТЛ
47,8261	23	11	ТКЛ
38,4615	13	5	ФилЛ

Таблица 6. Линии группы 2

Table 6. Group lines 2

% от всех станций линии	Всего станций	Из них в группе	Линия
40,9091	22	9	АПЛ
85,7143	7	6	БутЛ
36,3636	22	8	ЗамЛ
30	10	3	КалЛ
33,3333	3	1	КахЛ
8,3333	12	1	КолЛ
33,3333	24	8	КРЛ
35	20	7	ЛДЛ
27,2727	22	6	СокЛ
56	25	14	СТЛ
43,4783	23	10	ТКЛ
-	13	-	ФилЛ

Таблица 7. Линии группы 3

Table 7. Group lines 3

% от всех станций линии	Всего станций	Из них в группе	Линия
9,0909	22	2	АПЛ
-	7	-	БутЛ
9,0909	22	2	ЗамЛ
20	10	2	КалЛ
33,3333	3	1	КахЛ
16,6667	12	2	КолЛ
8,3333	24	2	КРЛ
5	20	1	ЛДЛ
-	22	-	СокЛ
-	25	-	СТЛ
-	23	-	ТКЛ
7,6923	13	1	ФилЛ

В таблице 5 представлены станции группы 1. Наибольший процент от станций, принадлежащих этой группе имеют: Калининская, Калужско-Рижская, Кольцевая, Замоскворецкая, Сокольническая и Таганско-Краснопресненская линии. Станции этой линии относятся к так называемой, «рабочей зоне», поэтому можно предположить, что именно вблизи станций указанных выше линий высокая концентрация рабочих мест. В таблице 6 содержатся данные, относящиеся к группе 2. Здесь лидерами являются Бутовская и Серпуховско-Тимирязевская линии. Филевская линия не представлена. При этом 85% станций Бутовской линии принадлежит этой группе.

В таблице 7 содержатся данные, относящиеся к группе 3. В этой группе отсутствуют станции: Бутовской, Серпуховско-Тимирязевской, Сокольнической, Таганско-Краснопресненской линий. Процент станций остальных линий невысок.

Заключение

В статье предложен метод классификации станций на основе коэффициента перекоса пассажиропотока (КПП), отражающего преобладающий относительно станции (со станции / на станцию) пассажиропоток за заданный период времени. Предлагается классификация, основанная на комбинации знаков значений КПП в утренний и вечерний часы пик. Знак КПП показывает преобладающее направление пассажиропотока: «-» со станции, «+» на станцию. Получено четыре группы: утренний и вечерний КПП меньше нуля («Полнос отъездов»); утренний КПП больше нуля, вечерний КПП меньше нуля («Рабочая зона»); утренний КПП меньше нуля, вечерний КПП больше нуля («Жилая зона»); утренний и вечерний КПП больше нуля («Полнос приездов»). Посчитан процент станций каждой из линий, принадлежащий каждой из групп.

Список использованных источников

- [1] *Намиот Д. Е., Д. Е., Шнепс-Шнеппе М. А.* Об общественных стандартах для Умного Города // International Journal of Open Information Technologies. 2016. Т. 4, № 7. С. 32-37. URL: <https://elibrary.ru/item.asp?id=26252117> (дата обращения: 21.04.2019).
- [2] *Намиот Д. Е., Кутузманов З. М., Федоров Е. А., Покусаев О. Н.* Об оценке социально-экономических эффектов городской железной дороги // International Journal of Open Information Technologies. 2018. Т. 6, № 1. С. 92-103. URL: <https://elibrary.ru/item.asp?id=32314918> (дата обращения: 21.04.2019).
- [3] *Намиот Д. Е., Покусаев О. Н., Лазуткина В. С.* О моделях пассажирского потока для городских железных дорог // International Journal of Open Information Technologies. 2018. Т. 6, № 3. С. 9-14. URL: <https://elibrary.ru/item.asp?id=32595087> (дата обращения: 21.04.2019).
- [4] *Namiot D., Sneps-Snepppe M.* A Survey of Smart Cards Data Mining // CEUR Workshop Proceedings. Supplementary Proceedings of the Sixth International Conference on Analysis of Images, Social Networks and Texts (AIST 2017). Moscow, Russia, 2017. Vol. 1975. URL: <http://ceur-ws.org/Vol-1975/paper33.pdf> (дата обращения: 21.04.2019).
- [5] *El Mahrsi M. K. et al.* Understanding passenger patterns in public transit through smart card and socioeconomic data // ACM SIGKDD Workshop on Urban Computing. 2014. URL: <https://www.comeetie.fr/pdfreposit/urbcomp2014.pdf> (дата обращения: 21.04.2019).



- [6] *Ceapa I., Smith C., Capra L.* Avoiding the crowds: understanding Tube station congestion patterns from trip data // Proceedings of the ACM SIGKDD International Workshop on Urban Computing (UrbComp '12). ACM, New York, NY, USA, 2012. Pp. 134-141. DOI: 10.1145/2346496.2346518
- [7] *Dai X., Lijun S., Yanyan X.* Short-Term Origin-Destination Based Metro Flow Prediction with Probabilistic Model Selection Approach // Journal of Advanced Transportation. 2018. Vol. 2018. Article ID 5942763, 15 pp. DOI: 10.1155/2018/5942763
- [8] *Liu Y., Qin Y., Guo J., Cai C., Wang Y., Jia L.* Short-term forecasting of rail transit passenger flow based on long short-term memory neural network // 2018 International Conference on Intelligent Rail Transportation (ICIRT), Singapore, 2018. Pp. 1-5. DOI: 10.1109/ICIRT.2018.8641683
- [9] *Mínguez R., Sánchez-Cambronero S., Castillo W., Jiménez P.* Optimal traffic plate scanning location for OD trip matrix and route estimation in road networks // Transportation Research Part B: Methodological. 2010. Vol. 44, Issue 2. Pp. 282-298. DOI: 10.1016/j.trb.2009.07.008
- [10] *Zhong C., Batty M., Manley E., Wang J., Wang Z., Chen F., Schmitt G.* Variability in Regularity: Mining Temporal Mobility Patterns in London, Singapore and Beijing Using Smart-Card Data // PLoS ONE. 2016. Vol. 11, Issue 2. e0149222. DOI: 10.1371/journal.pone.0149222
- [11] *Long Y., Thill J.-C.* Combining smart card data and household travel survey to analyze jobs-housing relationships in Beijing // Computers, Environment and Urban Systems. 2015. Vol. 53. Pp. 19-35. DOI: j.compenvurbsys.2015.02.005
- [12] *Hong C., Manley E., Arisona S.M., Batty M., Schmitt G.* Measuring variability of mobility patterns from multiday smart-card data // Journal of Computational Science. 2015. Vol. 9. Pp. 125-130. DOI: 10.1016/j.jocs.2015.04.021
- [13] *Zhu X., Xu R.* Generating route choice sets with operation information on metro networks // Journal of Traffic and Transportation Engineering (English Edition). 2016. Vol. 3, Issue 3. Pp. 243-252. DOI: 10.1016/j.jtte.2016.05.001
- [14] *Chang J.Y., Yang H.L.* Improving the transfer flow of the Taipei metro system for route development: Using open data of the origin and destination station // Proceedings of The 18th International Conference on Electronic Business. ICEB, Guilin, China, 2018. Pp. 766-772. URL: <https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1084&context=iceb2018> (дата обращения: 11.04.2019).
- [15] *Guo X. et al.* Multiperiod-based timetable optimization for metro transit networks // Transportation Research Part B: Methodological. 2017. Vol. 96. Pp. 46-67. DOI: 10.1016/j.trb.2016.11.005
- [16] *Mo Y. et al.* A flexible metro train scheduling approach to minimize energy cost and passenger waiting time // Computers & Industrial Engineering. 2019. Vol. 132. Pp. 412-432. DOI: 10.1016/j.cie.2019.04.031
- [17] *Leeuwen van D. et al.* Spatio-Temporal Clustering of Time-Dependent Origin-Destination Electronic Trace Data // Procedia Computer Science. 2018. Vol. 130. Pp. 359-367. DOI: 10.1016/j.procs.2018.04.053
- [18] *Lin D., Zhang Y., Zhu R., Meng L.* The analysis of catchment areas of metro stations using trajectory data generated by dockless shared bikes // Sustainable Cities and Society. 2019. Vol. 49. Id: 101598. DOI: 10.1016/j.scs.2019.101598
- [19] *Calvo F., Eboli L., Forciniti C., Mazzulla G.* Factors influencing trip generation on metro system in Madrid (Spain) // Transportation Research Part D: Transport and Environment. 2019. Vol. 67. Pp. 156-172. DOI: 10.1016/j.trd.2018.11.021
- [20] *Martin D., Gale C., Cockings S., Harfoot A.* Origin-destination geodemographics for analysis of travel to work flows // Computers, Environment and Urban Systems. 2018. Vol. 67. Pp. 68-79. DOI: 10.1016/j.compenvurbsys.2017.09.002
- [21] *Bachir D., Khodabandelou G., Gauthier V., El-Yacoubi M. A., Puchinger J.* Inferring dynamic origin-destination flows by transport mode using mobile phone data // Transportation Research Part C: Emerging Technologies. 2019. Vol. 101. Pp. 254-275. DOI: 10.1016/j.trc.2019.02.013

Поступила 21.04.2019; принята к публикации 25.06.2019;
опубликована онлайн 25.07.2019.

Об авторах:

Поматилов Филипп Семёнович, выпускник факультета вычислительной математики и кибернетики, Московский государственный университет имени М.В. Ломоносова (119991, Россия, г. Москва, ГСП-1, Ленинские горы, д. 1), ORCID: <http://orcid.org/0000-0002-1997-3043>, pomatilov@gmail.com

Намиот Дмитрий Евгеньевич, старший научный сотрудник лаборатории открытых информационных технологий, факультет вычислительной математики и кибернетики, Московский государственный университет имени М.В. Ломоносова (119991, Россия, г. Москва, ГСП-1, Ленинские горы, д. 1), кандидат физико-математических наук, ORCID: <http://orcid.org/0000-0002-4463-1678>, dnamiot@gmail.com

Все авторы прочитали и одобрили окончательный вариант рукописи.

References

- [1] *Namiot D., Sneps-Snepp M.* On the domestic standards for Smart Cities. *International Journal of Open Information Technologies*. 2016; 4(7):32-37. Available at: <https://elibrary.ru/item.asp?id=26252117> (accessed 21.04.2019). (In Russ., abstract in Eng.)
- [2] *Namiot D., Kutuzmanov Z., Fedorov E., Pokusaev O.* On the assessment of socio-economic effects of the city railway. *International Journal of Open Information Technologies*. 2018; 6(1):92-103. Available at: <https://elibrary.ru/item.asp?id=32314918> (accessed 21.04.2019). (In Russ., abstract in Eng.)
- [3] *Namiot D., Pokusaev O., Lazutkina V.* On passenger flow data models for urban railways. *International Journal of Open Information Technologies*. 2018; 6(3):9-14. Available at: <https://elibrary.ru/item.asp?id=32595087> (accessed 21.04.2019). (In Russ., abstract in Eng.)
- [4] *Namiot D., Sneps-Snepp M.* A Survey of Smart Cards Data Mining. *CEUR Workshop Proceedings*. Supplementary Proceedings of the Sixth International Conference on Analysis of Images, Social Networks and Texts (AIST 2017). Moscow, Russia, 2017; 1975. Available at: <http://ceur-ws.org/Vol-1975/paper33.pdf> (accessed 21.04.2019). (In Eng.)
- [5] *El Mahrsi M.K. et al.* Understanding passenger patterns in public transit through smart card and socioeconomic data. *ACM SIGKDD Workshop on Urban Computing*. 2014. Available at: <https://www.comeetie.fr/pdfrepos/urbcomp2014.pdf> (accessed 21.04.2019). (In Eng.)
- [6] *Ceapa I., Smith C., Capra L.* Avoiding the crowds: understanding Tube station congestion patterns from trip data. *Proceedings of the ACM SIGKDD International Workshop on Urban Computing (UrbComp '12)*. ACM, New York, NY, USA, 2012; 134-141. (In Eng.) DOI: 10.1145/2346496.2346518



- [7] Dai X., Lijun S., Yanyan X. Short-Term Origin-Destination Based Metro Flow Prediction with Probabilistic Model Selection Approach. *Journal of Advanced Transportation*. 2018; 2018:5942763. (In Eng.) DOI: 10.1155/2018/5942763
- [8] Liu Y., Qin Y., Guo J., Cai C., Wang Y., Jia L. Short-term forecasting of rail transit passenger flow based on long short-term memory neural network. *2018 International Conference on Intelligent Rail Transportation (ICIRT)*. Singapore. 2018; 1-5. (In Eng.) DOI: 10.1109/ICIRT.2018.8641683
- [9] Mínguez R., Sánchez-Cambronero S., Castillo W., Jiménez P. Optimal traffic plate scanning location for OD trip matrix and route estimation in road networks. *Transportation Research Part B: Methodological*. 2010; 44(2):282-298. (In Eng.) DOI: 10.1016/j.trb.2009.07.008
- [10] Zhong C., Batty M., Manley E., Wang J., Wang Z., Chen F., Schmitt G. Variability in Regularity: Mining Temporal Mobility Patterns in London, Singapore and Beijing Using Smart-Card Data. *PLoS ONE*. 2016; 11(2):e0149222. (In Eng.) DOI: 10.1371/journal.pone.0149222
- [11] Long Y., Thill J-C. Combining smart card data and household travel survey to analyze jobs-housing relationships in Beijing. *Computers, Environment and Urban Systems*. 2015; 53:19-35. (In Eng.) DOI: j.compenvurbsys.2015.02.005
- [12] Hong C., Manley E., Arisona S.M., Batty M., Schmitt G. Measuring variability of mobility patterns from multiday smart-card data. *Journal of Computational Science*. 2015; 9:125-130. (In Eng.) DOI: 10.1016/j.jocs.2015.04.021
- [13] Zhu X., Xu R. Generating route choice sets with operation information on metro networks. *Journal of Traffic and Transportation Engineering (English Edition)*. 2016; 3(3):243-252. (In Eng.) DOI: 10.1016/j.jtte.2016.05.001
- [14] Chang J.Y., Yang H.L. Improving the transfer flow of the Taipei metro system for route development: Using open data of the origin and destination station. *Proceedings of The 18th International Conference on Electronic Business*. ICEB, Guilin, China. 2018; 766-772. Available at: <https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1084&context=iceb2018> (accessed 21.04.2019). (In Eng.)
- [15] Guo X. et al. Multiperiod-based timetable optimization for metro transit networks. *Transportation Research Part B: Methodological*. 2017; 96:46-67. (In Eng.) DOI: 10.1016/j.trb.2016.11.005
- [16] Mo Y. et al. A flexible metro train scheduling approach to minimize energy cost and passenger waiting time. *Computers & Industrial Engineering*. 2019; 132:412-432. (In Eng.) DOI: 10.1016/j.cie.2019.04.031
- [17] Leeuwen van D. et al. Spatio-Temporal Clustering of Time-Dependent Origin-Destination Electronic Trace Data. *Procedia Computer Science*. 2018; 130:359-367. (In Eng.) DOI: 10.1016/j.procs.2018.04.053
- [18] Lin D., Zhang Y., Zhu R., Meng L. The analysis of catchment areas of metro stations using trajectory data generated by dockless shared bikes. *Sustainable Cities and Society*. 2019; 49:101598. (In Eng.) DOI: 10.1016/j.scs.2019.101598
- [19] Calvo F., Eboli L., Forciniti C., Mazzulla G. Factors influencing trip generation on metro system in Madrid (Spain). *Transportation Research Part D: Transport and Environment*. 2019; 67:156-172. (In Eng.) DOI: 10.1016/j.trd.2018.11.021
- [20] Martin D., Gale C., Cockings S., Harfoot A. Origin-destination geodemographics for analysis of travel to work flows. *Computers, Environment and Urban Systems*. 2018; 67:68-79. (In Eng.) DOI: 10.1016/j.compenvurbsys.2017.09.002
- [21] Bachir D., Khodabandelou G., Gauthier V., El-Yacoubi M.A., Puchinger J. Inferring dynamic origin-destination flows by transport mode using mobile phone data. *Transportation Research Part C: Emerging Technologies*. 2019; 101:254-275. (In Eng.) DOI: 10.1016/j.trc.2019.02.013

Submitted 21.04.2019; revised 25.06.2019;
published online 25.07.2019.

About the authors:

Filipp S. Pomatillov, graduate Faculty of Computational Mathematics and Cybernetics, Lomonosov Moscow State University (1, Leninskie gory, Moscow 119991, Russia), ORCID: <http://orcid.org/0000-0002-1997-3043>, pomatillov@gmail.com

Dmitry E. Namiot, Senior Researcher of the Laboratory of Open Information Technologies, Faculty of Computational Mathematics and Cybernetics, Lomonosov Moscow State University (1, Leninskie gory, Moscow 119991, Russia), Ph.D. (Phys.-Math.), ORCID: <http://orcid.org/0000-0002-4463-1678>, dnamiot@gmail.com

All authors have read and approved the final manuscript.

