

СПЕЦИАЛЬНЫЙ ВЫПУСК

УДК: 519.8

Облачная интерпретация энтропийной модели расчета матрицы корреспонденций

И. В. Подлипнова^{1,а}, Ю. В. Дорн^{1,2,б}, И. А. Склонин^{1,с}

¹Национальный исследовательский университет «Московский физико-технический институт»,
Россия, 141701, г. Долгопрудный, Институтский пер., д. 9

²Институт перспективных исследований проблем искусственного интеллекта и интеллектуальных
систем, Московский государственный университет имени М. В. Ломоносова,
Россия, 119991, г. Москва, Ломоносовский проспект, д. 27, корп. 1

E-mail: ^а podlipnova.iv@phystech.edu, ^б dornyv@yandex.ru, ^с sklonin.ilya@phystech.edu

Получено 16.12.2023, после доработки – 23.12.2023.
Принято к публикации 23.12.2023.

С ростом населения городов сильнее ощущается необходимость планирования развития транспортной инфраструктуры. Для этой цели создаются пакеты транспортного моделирования, которые обычно содержат набор задач выпуклой оптимизации, итеративное решение которых приводит к искомому равновесному распределению потоков по путям. Одно из направлений развития транспортного моделирования – это построение более точных обобщенных моделей, которые учитывают различные типы пассажиров, их цели поездок, а также специфику личных и общественных средств передвижения, которыми могут воспользоваться агенты. Другим не менее важным направлением является улучшение эффективности производимых вычислений, так как в связи с большой размерностью современных транспортных сетей поиск численного решения задачи равновесного распределения потоков по путям является довольно затратным. Итеративность всего процесса решения лишь усугубляет это. Одним из подходов, ведущим к уменьшению числа производимых вычислений, и является построение согласованных моделей, которые позволяют объединить блоки 4-стадийной модели в единую задачу оптимизации. Это позволяет исключить итеративную прогонку блоков, перейдя от решения отдельной задачи оптимизации на каждом этапе к некоторой общей задаче. В ранних работах было доказано, что такие подходы дают эквивалентные решения. Тем не менее стоит рассмотреть обоснованность и интерпретируемость этих методов. Целью данной статьи является обоснование единой задачи, объединяющей в себе как расчет матрицы корреспонденций, так и модальный выбор, для обобщенного случая, когда в транспортной сети присутствуют различные слои спроса, типы агентов и классы транспортных средств. В статье приводятся возможные интерпретации для калибровочных параметров, применяемых в задаче, а также для двойственных множителей, ассоциированных с балансовыми ограничениями. Авторы статьи также показывают возможность объединения рассматриваемой задачи с блоком определения загрузки сети в единую задачу оптимизации.

Ключевые слова: мультиномиальный логит, модель дискретного выбора, модальный выбор, энтропийная модель

Исследование выполнено при поддержке Министерства науки и высшего образования Российской Федерации (госзаказ), номер проекта FSMG-2024-0011.

UDC: 519.8

Cloud interpretation of the entropy model for calculating the trip matrix

I. V. Podlipnova^{1,a}, Yu. V. Dorn^{1,2,b}, I. A. Sklonin^{1,c}

¹National Research University Moscow Institute of Physics and Technology,
9 Institute lane, Dolgoprudny, 141701, Russia

²MSU Institute for Artificial Intelligence,
27/1 Lomonosov's prospect, Moscow, 119991, Russia

E-mail: ^a podlipnova.iv@phystech.edu, ^b dornvv@yandex.ru, ^c sklonin.ilya@phystech.edu

Received 16.12.2023, after completion – 23.12.2023.

Accepted for publication 23.12.2023.

As the population of cities grows, the need to plan for the development of transport infrastructure becomes more acute. For this purpose, transport modeling packages are created. These packages usually contain a set of convex optimization problems, the iterative solution of which leads to the desired equilibrium distribution of flows along the paths. One of the directions for the development of transport modeling is the construction of more accurate generalized models that take into account different types of passengers, their travel purposes, as well as the specifics of personal and public modes of transport that agents can use. Another important direction of transport models development is to improve the efficiency of the calculations performed. Since, due to the large dimension of modern transport networks, the search for a numerical solution to the problem of equilibrium distribution of flows along the paths is quite expensive. The iterative nature of the entire solution process only makes this worse. One of the approaches leading to a reduction in the number of calculations performed is the construction of consistent models that allow to combine the blocks of a 4-stage model into a single optimization problem. This makes it possible to eliminate the iterative running of blocks, moving from solving a separate optimization problem at each stage to some general problem. Early work has proven that such approaches provide equivalent solutions. However, it is worth considering the validity and interpretability of these methods. The purpose of this article is to substantiate a single problem, that combines both the calculation of the trip matrix and the modal choice, for the generalized case when there are different layers of demand, types of agents and classes of vehicles in the transport network. The article provides possible interpretations for the gauge parameters used in the problem, as well as for the dual factors associated with the balance constraints. The authors of the article also show the possibility of combining the considered problem with a block for determining network load into a single optimization problem.

Keywords: multinomial logit, discrete choice model, modal choice, entropy model

Citation: *Computer Research and Modeling*, 2024, vol. 16, no. 1, pp. 89–103 (Russian).

The research is supported by the Ministry of Science and Higher Education of the Russian Federation (Goszadaniye), project No. FSMG-2024-0011.

Введение

Основной методологией моделирования транспортных потоков в целях долгосрочного планирования и при принятии решений по оценке крупных транспортных инфраструктурных проектов является 4-стадийная модель [Ortúzar, Willumsen, 2011; Гасников, Гасникова, 2020], объединяющая в себе сразу несколько моделей: статистическую модель генерации спроса, модель расчета матрицы корреспонденций, модель распределения корреспонденций по различным типам передвижения (модам) и модель определения порождаемой загрузки сети.

Каждый из блоков решает свою задачу. Модели и подходы к решению этих задач долгое время создавались, развивались и исследовались независимо от других. Хотя корни почти всех моделей, используемых для решения этих задач, лежат в теоретико-игровых и оптимизационных подходах и могут быть отслежены до работ начала XX века [Pigou, 1920], их современные версии существенно различны. Например, модели равновесных распределений потоков по маршрутам основываются на концепции равновесия Нэша – Вардропа [Wardrop, 1952], расщепление по модам, как правило, основано на моделях дискретного выбора [Anderson, Palma, Thisse, 1992], а модуль расчета матрицы корреспонденций использует гравитационную или энтропийную модель Вильсона [Вильсон, 1978], идеально вдохновленную статистической физикой.

Одним из актуальных направлений исследований является построение согласованных моделей для реализации 4-стадийной схемы. В частности, в работах [Evans, 1976; Гасников, Гасникова, 2020; Гасникова и др., 2023] строятся оптимизационные модели, решение которых задает согласованную матрицу корреспонденций, модальное расщепление и равновесную загрузку сети. При этом гарантируется, что полученные таким образом решения совпадают с решениями, полученными на выходе стандартных блоков в 4-стадийной модели. Это позволяет улучшать эффективность вычисления модели, однако ничего не говорит про обоснованность или интерпретируемость выбранной модели.

В работе [Гасников и др., 2014] была предпринята одна из первых попыток согласования моделей оценки в 4-стадийной схеме не только с точки зрения вычислений, но и с точки зрения обоснования и методологии. В частности, была построена согласованная модель расчета матрицы корреспонденций и равновесной загрузки сети с использованием единого подхода моделирования.

В этой работе мы предлагаем развить эти идеи и предлагаем подход для гибкого построения согласованных оптимизационных моделей, подменяющих многостадийную схему и допускающих интерпретацию основных блоков с единой позиции.

Описание 4-стадийной модели

Для того чтобы обозначить контекст данной статьи, а также место, которое занимает поиск матрицы корреспонденций (передвижений) в задаче поиска транспортного равновесия, опишем кратко один из основных фреймворков, применяемых в транспортном моделировании.

Как уже было сказано, основной методологией моделирования транспортных потоков является 4-стадийная модель. Она состоит из четырех блоков: оценка общего трафика, построение соответствующей матрицы корреспонденций, распределение по модам и, наконец, выбор маршрутов. Модули, отвечающие за генерацию матрицы корреспонденций, расщепление и поиск равновесной загрузки сети, выполняются итеративно для получения самосогласованного решения или, иными словами, пока не найдется неподвижная точка. Схематично 4-стадийная модель показана на рис. 1. Стоит также упомянуть, что первому блоку обычно предшествует некоторый этап предобработки графа улично-дорожной сети. На этом этапе добавляются фиктивные ребра в графе, соответствующие каждому из возможных направлений проезда развязок и перекрестков, а моделируемая область — город или городская агломерация — разделяется на районы.

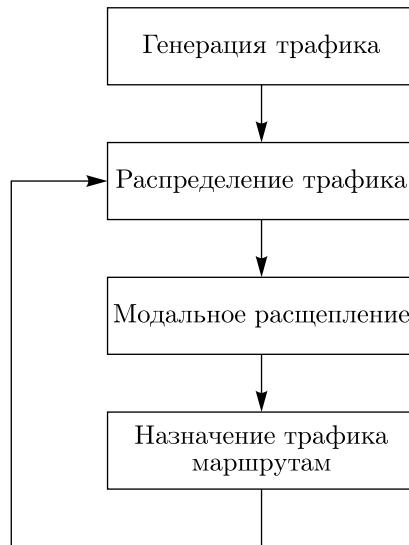


Рис. 1. Схематичное представление блоков 4-стадийной модели транспортного моделирования

Следуя [Gasnikova et al., 2023], введем основные обозначения. Будем рассматривать улично-дорожную сеть (УДС) в виде ориентированного графа $G = (V, E)$, где вершины V соответствуют перекресткам или центроидам [Sheffi, 1984], а ребра E соответствуют дорогам. Обозначим за $O \subseteq V$ множество районов-источников корреспонденций, а за $D \subseteq V$ — множество районов-стоков корреспонденций, при этом один и тот же район может являться как источником, так и стоком. В блоке генерации трафика для каждого $i \in O$ определяется объем отправлений l_i , а для каждого $j \in D$ — объемы прибытий w_j . Обычно такая информация получается из реальных замеров трафика на дорогах или на основании результатов проводимых социологических исследований.

В блоке распределения трафика на основании матрицы цен (затрат) определяется матрица корреспонденций $\{d_{ij}\}_{i \in O, j \in D}$, где d_{ij} обозначает суммарный трафик из i -го района в j -й район. Для расчета матриц корреспонденций используют, например, гравитационные или энтропийные модели [Вильсон, 1978], которые оценивают вероятность передвижения между районами на основе некоторой меры удаленности этих районов друг от друга. Обычно в качестве меры дальности в транспортных моделях используется обобщенная цена передвижения по оптимальному пути.

После оценки матрицы корреспонденций происходит расчет расщепления корреспонденций по типам передвижений (модам транспорта) с использованием той же самой матрицы цен, которая была задействована на предыдущем шаге. Для этого обычно используют модели дискретного выбора, такие как логит-модели. Наиболее используемыми на практике являются родственные логит-модели Nested Logit Model и Multinomial Logit Model, а также композитная модель Mixed Logit Model [Anderson, Palma, Thisse, 1992; Гасников и др., 2014].

Следующий этап предполагает определение общей загрузки сети. Для решения данной задачи применяются равновесные модели транспортных потоков, такие как модель Бэкмана или модель стабильной динамики. Распределение агентов по путям ведет к изменению издержек, затрат на перемещение, иными словами, изменяется матрица цен, которая использовалась во втором и третьем блоке 4-стадийной модели. Здесь и возникает итерационный процесс, который происходит, пока не будет выполнен некоторый критерий остановки. Обычно задача четвертого блока формулируется в виде некоторой задачи оптимизации, которая решается градиентными методами, такими как метод Франк–Вульфа [Frank, Wolfe, 1956] или универсальный метод подобных треугольников [Гасников, Нестеров, 2018]. Также активно исследуется и развивается

возможность объединения блока распределения трафика и назначения трафика маршрутам в единую задачу оптимизации. Это позволяет избежать итеративного выполнения блоков и тем самым ускорить поиск транспортного равновесия [Evans, 1976; Гасников, Гасникова, 2020; Гасникова и др., 2023].

В данной статье подробно рассматривается методология построения согласованных моделей расчета матриц корреспонденций, модельного расщепления и поиска равновесного распределения транспортных потоков, обобщающая идеи модели [Гасников и др., 2014]. В качестве одного из основных иллюстрирующих примеров будет рассмотрена энтропийная модель расчета матрицы корреспонденций. Эта модель была предложена Дж. Вильсоном [Вильсон, 1978]. В упомянутой монографии Вильсона также рассматриваются обобщенные модели распределения. Такие модели позволяют определить число корреспонденций из района i в район j , совершаемых пассажирами типа t с использованием транспортной моды k . Такое обобщение позволяет сделать модель более реалистичной. Также при использовании обобщенной модели в задаче поиска транспортного равновесия модальное расщепление уже не выделяется в отдельный блок, а происходит на этапе распределения трафика.

Отдельный интерес представляет интерпретация эволюционной модели расчета матрицы корреспонденций. Это позволяет дать новую интерпретацию двойственных множителей, возникающих при формулировке задачи оптимизации, соответствующей задаче поиска матрицы корреспонденций.

Далее в данной статье будут рассмотрены существующие работы, которые предлагают использование теории дискретного выбора для совместного распределения трафика по маршрутам и для расщепления по транспортным модам. Такой подход позволяет не выделять выбор моды транспорта в отдельный блок, а получить данную информацию на этапе расчета матрицы корреспонденций, объединив второй и третий блоки 4-стадийной модели. Далее будет сформулирована задача оптимизации, а также описана облачная модель и ее обобщение с учетом наличия в транспортной сети различных слоев спроса и типов транспорта. Целью такой модели является интерпретация поставленной задачи оптимизации с единой позиции поиска (стохастического) равновесия Нэша – Вардропа в расширенном графе.

Анализ литературы

В исследованиях, посвященных транспортному моделированию, уже рассматривались модели, в которых при описании дискретного выбора, совершаемого пассажирами, учитывались не только доступные пути, но и моды транспорта. Основной целью усложнения моделей является стремление приблизить модели к реальности и как можно более точно отразить эффекты, происходящие на практике в улично-дорожных сетях, а также то, как люди принимают решение о выборе пути перемещения, транспортной моды и цели поездки.

В статье [Fernández et al., 1994] авторы рассматривают возможности моделирования комбинированных мод транспорта. Использование комбинированных мод возможно, например, когда реализуется стратегия Park & Ride. К примеру, из окраинных районов города пассажиры доеzzают на личном автомобиле до ближайшей станции метро, оставляют автомобиль на парковке и до места работы добираются уже общественным транспортом. В настоящее время такой подход очень распространен в городском планировании, так как позволяет разгрузить центральные улицы города путем уменьшения числа личных автомобилей на дорогах. Авторы статьи выделяют две основные проблемы при моделировании: выбор между «чистой» транспортной модой и комбинированной, а также выбор трансферной вершины, то есть вершины в графе улично-дорожной сети, где будет происходить смена типа транспорта. В качестве решения авторы предлагают несколько различных подходов, которые основаны на применении логит-моделей. В первом подходе авторы моделируют сначала выбор транспортной моды, который производится совместно

среди «чистых» и комбинированных транспортных мод, а выбор конкретного маршрута в графе считается ассоциированным с выбором моды. То есть предполагается, что при выборе моды транспорта пользователь уже оценивает все маршруты, доступные любым типам транспорта, выбирает из них оптимальный, тем самым выбирая подходящий тип транспорта. Стоит отметить, что выбор трансферной вершины в такой модели является неявным, он «зашит» в выборе оптимального маршрута. Поэтому как альтернативный подход авторы предлагают объединить две логит-модели: уже описанную модель выбора транспортной моды и модель выбора трансферной вершины. Таким образом, зная объем корреспонденций между районами i и j , сперва будут подсчитаны доли тех, кто воспользуется комбинированной транспортной модой c , а затем доли тех из них, кто при этом поедет через трансферную вершину t .

Целью статьи [Abrahamsson, Lundqvist, 1999] является разработка комбинированной модели, которая бы объединяла в себе распределение трафика и модальное расщепление. Авторы подразумевают, что характеристики районов — количество выезжающих и въезжающих пассажиров — экзогенно заданы. Авторы предлагают несколько различных подходов: одновременная модель, традиционная иерархическая модель, где сначала происходит распределение трафика, а затем модальное расщепление и обратная иерархическая модель, где сначала происходит модальное расщепление, а затем распределение трафика. В одновременной модели авторы применяют некоторую модификацию гравитационной модели расчета матрицы корреспонденций, предложенной в монографии А. Дж. Вильсона [Вильсон, 1978]. А именно, авторы предлагают использовать гравитационную модель с двойными ограничениями (*doubly constrained gravity model*), тем самым они получают не одну матрицу корреспонденций, а отдельные матрицы для каждой доступной моды транспорта в сети. В традиционной иерархической модели авторы статьи объединяют гравитационную модель, которая используется для получения матрицы корреспонденций, с логит-моделью, которая определяет вероятности выбора определенной моды транспорта для заданной пары районов (i, j) . Такой подход наиболее часто встречается на практике. Еще одна альтернативная модель, которая рассматривается в данной статье, — обратная иерархическая модель. В данном подходе для каждого района отправления i сначала определяются доли использования каждой моды транспорта, а затем для заданной моды определяются районы-назначения j с использованием логит-модели. В статье также показаны соответствующие задачи оптимизации для каждого из описанных подходов.

В статье [De Cea et al., 2005] предлагается еще более детальная модель, в ней для каждой пары районов (i, j) рассматриваются различные классы пользователей k , цели поездки r и транспортные моды t . Для определения районов назначения j используется энтропийная модель, а затем для моделирования остальных уровней пользовательского выбора применяется иерархическая логит-модель. Также в данной статье авторы учитывают специфику личного и общественного транспорта. Так, например, для общественного транспорта величина потока задается некоторой фиксированной величиной для каждого ребра транспортного графа. Также по-разному задаются обобщенные цены для пользователей личных автомобилей и для пассажиров общественного транспорта. Так, для пользователей автомобилями в обобщенную цену будут входить время в пути, операционные расходы, платы за проезд и за парковку. Обобщенная цена для пассажиров общественного транспорта будет включать в себя время на то, чтобы дойти до остановки, время ожидания транспорта, время в пути, плату за билет, плату за пересадки и время, на них потраченное.

Постановка задачи оптимизации

Введем обозначения согласно [Гасникова и др., 2023]. Рассматриваем замкнутую транспортную систему, описываемую графом $G = (V, E)$, где V — множество вершин ($|V| = n$),

а E — множество ребер ($|E| = m$). Будем обозначать ребра графа через $e \in E$. Для стандартной транспортной системы можно ожидать, что $m \simeq 3n$. Для больших мегаполисов (таких как Москва) $n \simeq 10^5$. Транспортный граф G считается известным.

Будем считать, что имеются разные слои спроса, например, передвижения «дом – работа» или «работа – отдых», которые мы будем индексировать буквой r . Внутри каждого слоя спроса допустимы разные типы пользователей $t \in M(r)$ ($M(r)$ — множество типов пользователей, отвечающих слою спроса r), например, имеющие собственный транспорт, которые могут выбирать между личным и общественным транспортом, и класс пользователей, которые не имеют личный транспорт. Тип транспорта будем обозначать индексом k . Множество типов транспортных средств, доступных пользователям типа t , будем обозначать как $Z(t)$. Множество всех типов транспортных средств K можно разбить на классы транспортных средств $\{K_b\}$, использующих (внутри класса) одну и ту же транспортную сеть b ($b \in B$). Под транспортной сетью понимается набор маршрутов P^b (составленных из ребер G), доступных для перемещения по данной сети. Например, можно выделить обычную сеть дорог для личного транспорта, которой пользуются легковые автомобили и частично общественный транспорт. Можно выделить велодорожки, которыми могут пользоваться не только велосипедисты, но, например, и самокатные средства.

Источниками корреспонденций (*origin*) обозначим подмножество вершин $O \subseteq V$, а стоки корреспонденций (*destination*) — $D \subseteq V$. То есть вводится множество пар вида (источник, сток) корреспонденций $OD \subseteq V \otimes V$. Сами корреспонденции будем обозначать через d_{ij}^{rt} , где $(i, j) \in OD$. Как правило, $|OD| \ll n^2$ [Гасников и др., 2014]. Не ограничивая общности, будем далее считать, что $\sum_{r, t \in M(r); (i, j) \in OD} d_{ij}^{rt} = 1$. Множество пар OD считается известным. Корреспонденции не известны, однако известны (заданы) характеристики источников и стоков корреспонденций, соответственно, l_i^r и w_j . То есть известны величины [Вильсон, 1978] $\{l_i^r\}_{i \in O}$, $\{w_j\}_{j \in D}$:

$$\sum_{t \in M(r)} \sum_{j: (i, j) \in OD} d_{ij}^{rt} = l_i^r, \quad \sum_r \sum_{t \in M(r)} i: (i, j) \in OD d_{ij}^{rt} = w_j. \quad (1)$$

Заметим, что $\sum_r \sum_{i \in O} l_i^r = \sum_{j \in D} w_j = 1$. Условие (1) будем также для краткости записывать в виде $d \in (l, w)$.

Под энтропийной моделью расчета матриц корреспонденций $d^{rt}(\bar{T})$ понимается определенный способ вычисления набора корреспонденций $\{d_{ij}^{rt}\}_{(i, j) \in OD}$ по известным матрицам средних затрат $\bar{T} = \{\bar{T}_{ij}^t\}_{(i, j) \in OD}$. Этот способ заключается в решении задачи энтропийно-линейного программирования, которую можно понимать как энтропийно-регуляризованную транспортную задачу

$$\min_{d \in (l, w); d \geq 0} \left\{ \sum_r \beta^r \sum_{t \in M(r); (i, j) \in OD} d_{ij}^{rt} \bar{T}_{ij}^t + \sum_{r; t \in M(r); (i, j) \in OD} d_{ij}^{rt} \ln d_{ij}^{rt} \right\}, \quad (2)$$

где параметры $\beta^r > 0$ считаются известными [Вильсон, 1978; Гасников и др., 2013; Гасников, Гасникова, 2020]. Относительно выбора этих параметров см. [Гасников и др., 2014; Гасников, Гасникова, 2020; Иванова и др., 2018]. В действительности это так называемые структурные параметры, от подгонки которых существенно зависит качество модели. Если известны средние времена C^r в пути для разных слоев спроса r , то находить β^r можно из системы уравнений

$$\sum_{t \in M(r); (i, j) \in OD} d_{ij}^{rt}(\beta) \bar{T}_{ij}^{rk} \simeq C^r.$$

Заметим, что формулу (2) можно переписать, явно указав ограничения. Таким образом, финальная формулировка задачи оптимизации, с которой мы будем работать в следующем разделе,

выглядит следующим образом:

$$\max_{\lambda, \mu} \min_{d \geq 0} \left\{ \sum_r \beta^r \sum_{t \in M(r); (i, j) \in OD} d_{ij}^{rt} \bar{T}_{ij}^t + \sum_{r; t \in M(r); (i, j) \in OD} d_{ij}^{rt} \ln d_{ij}^{rt} + \right. \\ \left. + \sum_{i; r} \lambda_i^r \left(l_i^r - \sum_{j; t \in M(r)} d_{ij}^{rt} \right) + \sum_j \mu_j \left(\mu_j - \sum_{r; t \in M(r); i} d_{ij}^{rt} \right) \right\}, \quad (3)$$

где λ_i^r и μ_j — двойственные множители. Переходим теперь к облачной интерпретации сформулированной задачи.

Обобщение облачной модели

Понятие облачной модели для интерпретации энтропийной модели расчета матрицы корреспонденций впервые вводится в статье [Гасников и др., 2014]. В данной статье авторы рассматривают более простую модель, которая не выделяет различные типы пользователей и слои спроса. Энтропийная модель расчета матрицы корреспонденций приводит авторов статьи к следующей задаче:

$$\min_{\lambda, \mu} \max_{\sum_{i,j} d_{ij} = 1; d \geq 0} \left\{ -\beta \sum_{i,j} d_{ij} T_{ij} - \sum_{i,j} d_{ij} \ln d_{ij} + \sum_{i,j} \lambda_i \left(l_i - \sum_j d_{ij} \right) + \sum_{i,j} \mu_j \left(w_j - \sum_i d_{ij} \right) \right\}. \quad (4)$$

В данной задаче λ_i и μ_j — множители Лагранжа при балансовых ограничениях. В качестве обоснования энтропийного расчета матрицы корреспонденции в статье рассматривается облачная модель. Авторы статьи предлагают объединить все вершины — источники корреспонденций с некоторым фиктивным источником (облако 1), а вершины — стоки корреспонденций объединить с фиктивным стоком (облако 2), как это показано на рис. 2.

Рассмотрим один слой спроса «дом – работа». Зададим на ребрах из фиктивного источника к районам – источникам корреспонденций некоторую цену λ_i , которая может иметь смысл платы за проживание в районе i в день. Пусть на ребрах, идущих из района i в район j , будет некоторая стоимость перемещения T_{ij} между заданной парой районов. А ребра из районов – стоков корреспонденций к фиктивному стоку будут иметь некоторый отрицательный вес, который может быть проинтерпретирован как заработка плата, которую человек получает в день, работая в районе j .

Авторы статьи рассматривают естественную в данном контексте логит-динамику, которая приводит к формулировке задачи, похожей на задачу (4):

$$\max_{\sum_{i,j} d_{ij} = 1; d \geq 0} \left\{ -\beta \sum_{i,j} d_{ij} T_{ij} - \sum_{i,j} d_{ij} \ln d_{ij} - \beta \sum_i \lambda_i \left(\sum_j d_{ij} \right) - \beta \sum_j \mu_j \left(\sum_i d_{ij} \right) \right\}. \quad (5)$$

Таким образом, авторы приходят к формулировке задачи, эквивалентной задаче (4), с той лишь разницей, что в задаче (5) нет необходимости оптимизировать по $2n$ двойственным множителям. К тому же в полученной в результате облачной модели задаче λ_i и μ_j имеют некоторую естественную интерпретацию, такую как плата за проживание в районе i и заработка плата в районе j соответственно. Так как обычно при выборе места жительства и места работы люди руководствуются более сложной системой предпочтений, которая основывается не только на заработной плате и плате за жилье, то более правильно интерпретировать параметры λ_i и μ_j как

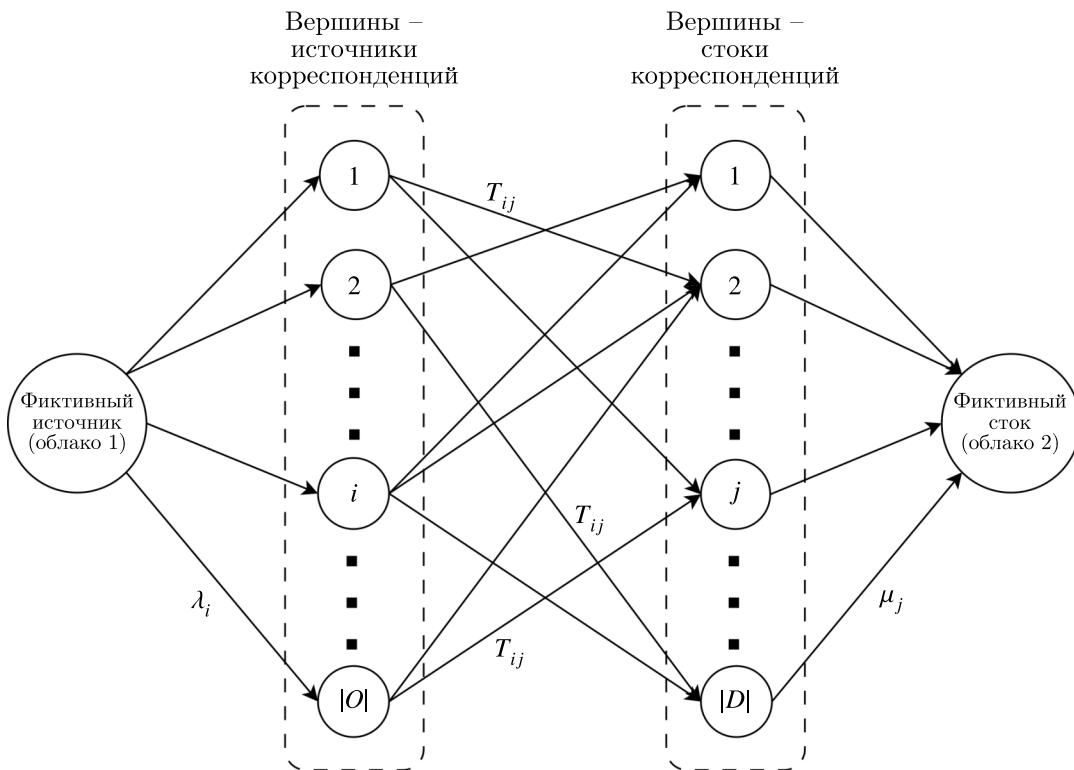


Рис. 2. Схема облачной модели расчета матрицы корреспонденций, рассматриваемой в статье [Гасников и др., 2014]

некоторые потенциалы притяжения/отталкивания, которые, помимо денежных характеристик, будут включать в себя и другие факторы, которые сложно описать количественно.

Концептуальная схема моделирования, описанная в облачной модели с дальнейшим переходом к логит-динамике, может быть сформулирована следующим образом.

1. Вводится расширенный граф, описывающий последовательное принятие решения агентом.
2. Если стоимости некоторых ребер или потоки на них известны (или оцениваются другими моделями), то соответствующие значения подставляются напрямую.
3. В расширенном графе рассчитывается (стохастическое) равновесие Нэша – Вардропа с помощью решения (сглаженной) задачи оптимизации.

В [Гасников и др., 2014] авторы ввели в расширенный граф только фиктивные ребра, соответствующие выбору источника и стока корреспонденции. В общем же случае можно вводить много различных этапов выбора, как это будет показано далее.

Энтропийный функционал можно рассматривать и как прокс-функцию, сглаживающую исходный целевой функционал, и как способ введения случайной ошибки при выборе оптимального маршрута.

Множители Лагранжа для ограничений на общее число агентов могут быть интерпретированы как стоимости использования соответствующих вершин. В более общем случае многих этапов выбора они также могут быть интерпретированы как «стоимости», отражающие предпочтительность того или иного выбора, доступность транспортных мод для различных районов и так далее.

Продемонстрируем применение этого подхода и обоснование с его помощью более общей задачи (3), которая учитывает различные слои спроса и типы транспортных средств.

Будем полагать, что $|R|$ — число различных слоев спроса, а $|T|$ — количество доступных в сети типов транспортных средств. Вместо одного фиктивного источника введем r фиктивных источников (см. рис. 3), каждый из которых обозначает слой спроса, так как считаем, что слои спроса в сети фиксированы. Иными словами, люди фиксировано относятся к одному из слоев спросов, например, «дом – работа», «дом – школа» или «дом – поездка за продуктами».

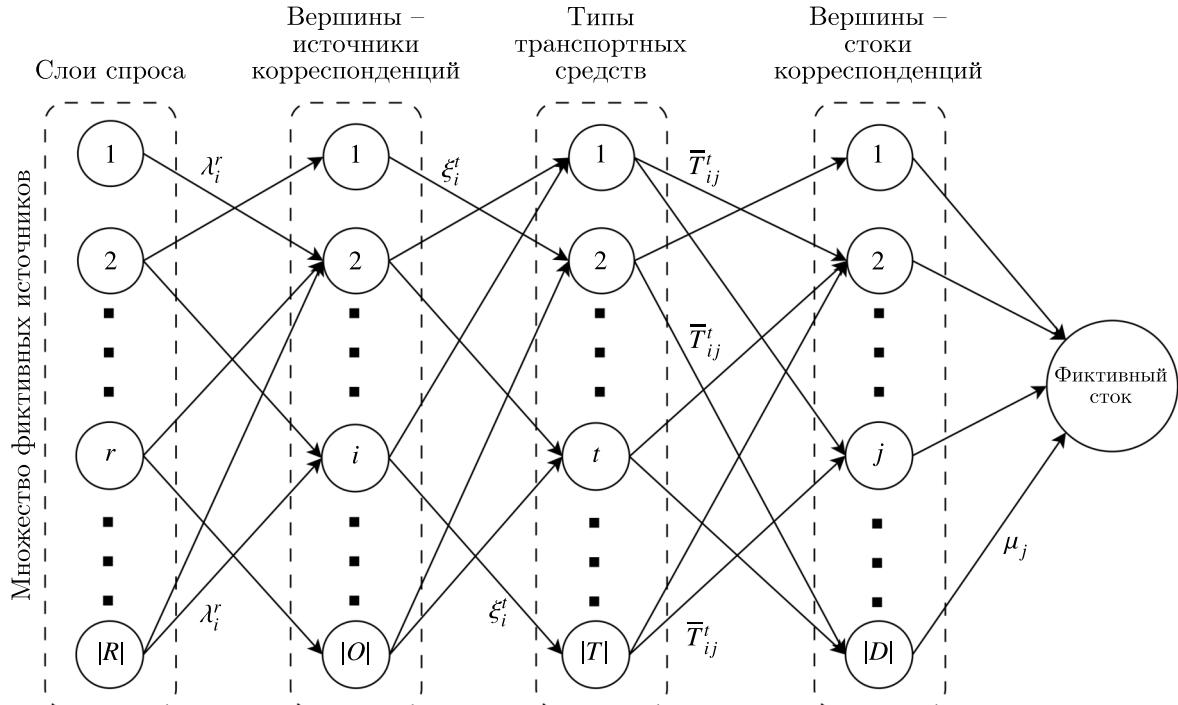


Рис. 3. Схема обобщенной облачной модели расчета матрицы корреспонденций

Так как теперь мы рассматриваем различные слои спроса, то есть цели поездок, то параметры, которые в базовой модели интерпретировались как плата за проживание в некотором районе, совершенно не обязаны интерпретироваться аналогично в обобщенной модели. Логичнее понимать λ_i^r и μ_j как некоторые параметры, обозначающие, например, вместимость парковочных мест в районах отправления и прибытия и цены, ассоциированные с количеством этих парковочных мест. В случае общественного транспорта это может быть вместимость остановочного пункта в районе прибытия. Стоит явно отметить, что в данной модели мы не разделяем потоки прибытий по слоям спроса, полагая, что вне зависимости от цели прибытия все участники движения будут пользоваться общими парковочными местами. Положим, что вес ребер между вершинами – источниками корреспонденций и типами транспортных средств нулевой. Иными словами, в момент принятия решения, на каком транспорте отправиться в район j , пассажир еще не несет никаких затрат. Затраты на непосредственное передвижение из района i в район j составляют \bar{T}_{ij}^r и зависят также от моды транспорта, с помощью которой совершается поездка. Отметим также, что для каждого слоя спроса будет своя ценность времени β^r . Например, люди, отправляющиеся на работу, готовы будут заплатить за проезд больше, чтобы добраться до пункта прибытия вовремя, так как цена за опоздание гораздо выше. А те, кто едет, например, за продуктами могут позволить себе проехать дальше, так как формально не ограничены строгими временными рамками.

Применяя к этим рассуждениям логит-динамику, в частности, используя мультиномиальную логит-модель репрезентативного потребителя [Anderson, Palma, Thisse, 1992], получаем следующую задачу:

$$\min_{d \geq 0} \left\{ \sum_r \beta^r \sum_{t \in M(r); (i, j) \in OD} d_{ij}^{rt} \bar{T}_{ij}^t + \sum_{i, t \in M(r)} \xi_i^t \sum_{j, r} d_{ij}^{rt} + \sum_r \gamma^r \sum_{t \in M(r); (i, j) \in OD} d_{ij}^{rt} \ln d_{ij}^{rt} + \right. \\ \left. + \sum_{i; r} \lambda_i^r \sum_{j; t \in M(r)} d_{ij}^{rt} + \sum_j \mu_j \sum_{r; t \in M(r); i} d_{ij}^{rt} \right\}, \quad (6)$$

которая по своей форме сильно похожа на задачу (3), сформулированную нами в предыдущем разделе статьи. Отличие, как и для базовой модели, состоит в том, что теперь мы не оптимизируем двойственные переменные при балансовых ограничениях, а могут быть интегрированы в модель напрямую при наличии наблюдений. Параметры λ_i^r и μ_j получили интерпретацию, которую также можно считать некоторым потенциалом притяжения/отталкивания районов i, j , но теперь при его рассмотрении учитывается слой спроса. Параметр ξ_i^t отражает доступность типа транспорта t для района i . В данную модель также укладывается и схема, предложенная в [Kubentayeva et al., 2023]. Так, например, для слоя спроса «дом – поездка за продуктами» не имеет смысла ехать в промышленный район, который востребован скорее для рабочих корреспонденций, но в котором отсутствуют крупные магазины и торговые центры. Заметим также, что при энтропии появился некоторый коэффициент γ^r . В теории репрезентативного потребителя [Anderson, Palma, Thisse, 1992] такой параметр отвечает за поведение, направленное на поиск разнообразия (variety-seeking behaviour). Иначе его можно интерпретировать как некоторую чувствительность к ошибкам. Так, люди, спешащие на работу, скорее будут ехать привычным маршрутом, боясь отклониться и увеличить свое время в пути. В то время как люди, отправляющиеся за покупками, могут легко отклоняться от маршрута и ехать по не самому оптимальному пути.

Решение оптимизационной задачи

Рассмотрим теперь методы, которыми можно решать полученную задачу (6). В частности, особый интерес представляет рассмотрение возможности объединения этой задачи со следующим блоком назначения трафика маршрутам. В статье [Гасникова и др., 2023] говорится о том, что наличие различных коэффициентов β^r в постановке задачи (2) препятствует объединению задач оптимизации, соответствующих блоку распределения трафика, совмещенного с модальным расщеплением, и блоку назначения трафика маршрутам, в единую седловую (выпукло-вогнутую) задачу оптимизации.

Введем для начала обозначения, которые фигурируют в постановке задачи блока назначения трафика маршрутам.

Обозначим через (зависимость $b(k)$ определяет тип сети по типу транспортного средства: $b = b(k)$ тогда и только тогда, когда $k \in K_b$)

$$\tau_e^k(f_{e,b(k)}, f_e^k) = F(\tau_e(f_{e,b(k)}), \bar{\tau}_e^k(f_e^k))$$

функцию затрат (например, временных) на проезд по ребру (участку дороги) e на типе транспортных средств k на сети $b(k)$, если суммарный поток на этом участке сети равен $f_{e,b(k)}$, а поток транспортных средств типа k равен $f_e^k \left(f_{e,b(k)} = \sum_{k \in K_b} f_e^k \right)$. Функция $F(\cdot)$ – некоторая агрегирующая

функция. Далее в качестве такой функции будем рассматривать сумму, чтобы иметь возможность свести задачу к задаче оптимизации. Функции $\tau_e(f_{e,b(k)})$ и $\tilde{\tau}_e^k(f_e^k)$ считаются заданными, например, таким образом [Гасников и др., 2013; Patriksson, 2015]:

$$\tau_e(f) = \bar{t}_e \left(1 + \zeta \left(\frac{f}{\bar{f}_e} \right) \right)^{1/\mu}, \quad (7)$$

где \bar{t}_e — время прохождения ребра e , когда участок свободный (определяется разрешенной скоростью на данном участке), а \bar{f}_e — пропускная способность ребра e (определяется полностью: [пропускная способность] \leq [число полос] * [2000 авт./ч] и характеристиками перекрестков). Считается, что эти характеристики известны [Stabler, Bar-Gera, Sall, 2018]. Параметр $\mu = 0,25$ — BPR-функции [Patriksson, 2015], но допускается и $\mu \rightarrow 0+$ — модель стабильной динамики [Nesterov, De Palma, 2003; Гасников, Гасникова, 2020; Котлярова и др., 2022]. Параметр $\zeta > 0$ также считается заданным. Относительно $\tilde{\tau}_e^k(f)$ часто считают, что $\tilde{\tau}_e^k(f) \equiv c_e^k$ [Швецов, 2003].

Например, если транспортное средство k весит больше 3,5 тонн, то по некоторым ребрам e проезд может быть запрещен, и тогда можно полагать $c_e^k = \infty$.

Полезно также ввести t_e^k — (временные) затраты на прохождения ребра e на транспортном средстве типа k . Обозначение t ранее уже было использовано под тип пользователя, однако это не должно вызывать недоразумение, поскольку далее по смыслу будет понятно, какое t имеется в виду в том или ином случае. Согласно написанному выше

$$t_e^k = F(t_{e,b(k)}, \bar{t}_e^k) = F(\tau_e(f_{e,b(k)}), \tilde{\tau}_e^k(f_e^k)) = \tau_e^k(f_{e,b(k)}, f_e^k). \quad (8)$$

По этим затратам $t^k = \{t_e^k\}_{e \in E}$ можно определить затраты на перемещение из источника i в сток j по кратчайшему пути:

$$T_{ij}^k(t^k) = \min_{p \in P_{ij}^{b(k)}} \sum_{e \in E} \delta_{ep} t_e^k, \quad (9)$$

где p — путь (без самопересечений-циклов) на графе (набор ребер), $P_{ij}^{b(k)}$ — множество всевозможных путей на графе транспортной сети $b(k)$, стартующих из источника i и заканчивающихся в стоке j , $\delta_{ep} = 1$, если ребро e принадлежит пути p , и $\delta_{ep} = 0$ — если иначе.

В ряде выкладок далее также будет полезен вектор $x^{rtk} = \{x_p^{rtk}\}_{p \in P^{b(k)}}$ — вектор распределения потоков по путям, где $P^b = \bigcup_{(i,j) \in OD} P_{ij}^b$. Заметим, что

$$f_e^k = \sum_{r; t \in M(r); p \in P^b(k)} \delta_{ep} x_p^{rtk}. \quad (10)$$

Введем функцию (опуская возможные различные индексы и волны)

$$\sigma_e(f_e) = \int_0^{f_e} \tau_e(z) dz$$

и сопряженную к ней функцию по Фенхелю – Рокафеллару:

$$\sigma_e^*(t_e) = \max_{f_e \geq 0} \{t_e f_e - \sigma_e(f_e)\}.$$

Это определение делается для того, чтобы «обратить» формулу $t_e = \tau_e(f_e)$. А именно, $t_e = \tau_e(f_e)$ тогда и только тогда, когда $f_e = \frac{d\sigma_e^*(t_e)}{dt_e}$.

Опуская далее подробный вывод прямой задачи, а затем получение двойственной к ней задачи оптимизации, а также доказательство потенциальности функционала, которые описаны в статье [Гасникова и др., 2023], выпишем сразу задачу, которая соответствует блоку назначения трафика маршрутам:

$$\max_{t=\{t_{b(k)}, \vec{t}^k\}_{k \in K} \geq 0} \left\{ \sum_{r; t \in M(r); (i, j) \in OD} d_{ij}^{rt} \bar{T}_{ij}^t - \sum_{b \in B; e \in E} \sigma_e^*(t_{e,b(k)}) - \sum_{k \in K; e \in E} \tilde{\sigma}_e^{*,k}(\vec{t}_e^k) \right\}, \quad (11)$$

где

$$\bar{T}_{ij}^t = \min \left\{ \left\{ T_{ij}^k(t_{b(k)} + \vec{t}^k) \right\}_{k \in Z(t)} \right\}. \quad (12)$$

Заметим, что задачи (6) и (11) имеют общую часть, зависящую от переменных d и t . Следуя статье [Гасникова и др., 2023], обозначим эту часть как $G = \widehat{G}$, а композитные члены в функционалах задач (6) и (11) переобозначим как g и h соответственно. Для согласованности с моделью, рассматриваемой в статье [Гасникова и др., 2023], будем полагать в задаче (6) параметры $\xi_i^t = 0$. Тогда краткий вид этих задач будет следующим:

$$\begin{aligned} & \min_{d \geq 0} G(d, \bar{T}(t)) + g(d), \\ & \max_{t \geq 0} \widehat{G}(d, \bar{T}(t)) - h(t). \end{aligned}$$

Если $G = \widehat{G}$, то эти задачи можно объединить в единую седловую задачу. В постановке задачи (2) такое условие не выполнялось, так как присутствовали параметры β^r , которые в общем случае различны для каждого слоя спроса. При этом мы не можем просто разделить все части задачи на β^r , так как тогда не ясен физический смысл коэффициента, возникающего при энтропийном слагаемом. Однако в эквивалентной постановке задачи (6), полученной из обоснования при помощи облачной модели, возникает коэффициент γ^r при энтропии, который имеет естественную интерпретацию. Таким образом, мы можем говорить, что в текущей постановке $G = \widehat{G}$ при соответствующей нормировке на β^r , а значит, задачи соответствующих блоков можно объединить в единую седловую задачу:

$$\min_{d \geq 0} \max_{t \geq 0} G(d, \bar{T}(t)) + g(d) - h(t),$$

которую в свою очередь можно переписать как (теорема фон Неймана–Сиона–Какутани)

$$\max_{t \geq 0} \min_{d \geq 0} G(d, \bar{T}(t)) + g(d) - h(t).$$

Запишем последнее выражение в развернутой форме:

$$\begin{aligned} & \max_{t \geq 0} \min_{d \geq 0} \left\{ \sum_{r; t \in M(r); (i, j) \in OD} d_{ij}^{rt} \bar{T}_{ij}^t + \sum_r \frac{\gamma^r}{\beta^r} \sum_{t \in M(r); (i, j) \in OD} d_{ij}^{rt} \ln d_{ij}^{rt} + \sum_{i; r} \frac{\lambda_i^r}{\beta^r} \sum_{j; t \in M(r)} d_{ij}^{rt} + \right. \\ & \quad \left. + \sum_{j; r} \frac{\mu_j}{\beta^r} \sum_{t \in M(r); i} d_{ij}^{rt} - \sum_{b \in B; e \in E} \sigma_e^*(t_{e,b(k)}) - \sum_{k \in K; e \in E} \tilde{\sigma}_e^{*,k}(\vec{t}_e^k) \right\}. \quad (13) \end{aligned}$$

В свою очередь, уже для этой задачи можно построить двойственную по части переменных d , в результате чего получится вогнутая задача с переменными t и d .

Таким образом, опуская подробные выкладки, мы рассмотрели также схему объединения задач двух блоков в единую задачу оптимизации.

Заключение

В данной статье была рассмотрена облачная интерпретация энтропийной модели расчета матрицы корреспонденций при наличии в сети различных слоев спроса и типов транспорта. Облачная модель позволила не только дополнительно обосновать, но и получить эквивалентную форму записи задачи, в которой двойственные множители и параметры имеют дополнительный физический смысл. Полученная эквивалентная форма задачи позволила также перейти от итеративного отрешивания (прогонки) двух блоков к их объединению в единую задачу оптимизации.

Список литературы (References)

- Вильсон А.* Энтропийные методы моделирования сложных систем. — М.: Мир, 1978. — 248 с.
Wilson A. G. Entropy in urban and regional modelling. — Pion Limited, 1970. (Russ. ed.: *Wilson A.* Entropijnye metody modelirovaniya slozhnyh sistem. — Moskva: Mir, 1978. — 248 p.)
- Гасников А. В., Гасникова Е. В.* Модели равновесного распределения потоков в больших сетях. — М.: МФТИ, 2020. — 204 с.
Gasnikov A. V., Gasnikova E. V. Modeli ravnovesnogo raspredeleniya potokov v bol'shikh setyakh [Models of equilibrium distribution of flows in large networks]. — Moscow: MIPT, 2020. — 204 p. (in Russian).
- Гасников А. В., Дорн Ю. В., Нестеров Ю. Е., Шпирко С. В.* О трехстадийной версии модели стационарной динамики транспортных потоков // Математическое моделирование. — 2014. — Т. 26. — С. 34–70.
Gasnikov A., Dorn Yu., Nesterov Yu., Shpirko S. O tryohstadiinoi versii modeli stacionarnoi dinamiki [On the three-stage version of stable dynamic model] // Mat. Model. — 2014. — Vol. 26. — P. 34–70 (in Russian).
- Гасников А. В., Кленов С. Л., Нурминский Е. А., Холодов Я. А., Шамрай Н. Б.* Введение в математическое моделирование транспортных потоков / под ред. А. В. Гасникова с приложениями М. Л. Бланка, К. В. Воронцова и Ю. В. Чеховича, Е. В. Гасниковой, А. А. Замятиной и В. А. Малышева, А. В. Колесникова, Ю. Е. Нестерова и С. В. Шпирко, А. М. Райгородского, с предисловием руководителя департамента транспорта г. Москвы М. С. Ликсутова. — М.: МЦНМО, 2013. — 2-е изд. — 427 с.
Gasnikov A. V., Klenov S. L., Nurminskii E. A., Kholodov Ya. A., Shamrai N. B. Vvedenie v matematicheskoe modelirovaniye transportnykh potokov [Introduction to mathematical modeling of transport flows]. — Moscow: MCNMO, 2013. — 427 p. (in Russian).
- Гасников А. В., Нестеров Ю. Е.* Универсальный метод для задач стохастической композитной оптимизации // Comput. Math. Math. Phys. — 2018. — Т. 58, № 1. — С. 48–64.
Gasnikov A. V., Nesterov Yu. E. Universal'nyi metod dlya zadach stochasticheskoy kompozitnoy optimizacii [Universal method for stochastic composite optimization problems] // Comput. Math. Math. Phys. — 2018. — Vol. 58, No. 1. — P. 48–64 (in Russian).
- Гасникова Е. В., Гасников А. В., Ярмошик Д. В., Кубентаева М. Б., Персиянов М. И., Подлипнова И. В., Комлярова Е. В., Склонин И. А., Подобная Е. Д., Матюхин В. В.* О многостадийной модели равновесного распределения транспортных потоков по путям и достаточных условиях, когда поиск равновесия сводится к решению задачи оптимизации // arXiv preprint. — 2023. — arXiv:2305.09069
Gasnikova E., Gasnikov A., Yarmoshik D., Kubentaeva M., Persianov M., Podlipnova I., Kotlyarova E., Sklonin I., Podobnaya E., Matyukhin V. O mnogostadiinoi modeli ravnovesnogo raspredeleniya potokov po putyam i dostatochnykh usloviyach, kogda poisk ravnovesiya svoditsya k resheniyu zadachi optimizacii [Sufficient conditions for multi-stages traffic assignment model to be the convex optimization problem] // arXiv preprint. — 2023. — arXiv:2305.09069
- Иванова А. С., Омельченко С. С., Комлярова Е. В., Матюхин В. В.* Калибровка параметров модели расчета матрицы корреспонденций для г. Москвы // Компьютерные исследования и моделирование. — 2020. — Т. 12, № 5. — С. 961–978.
Ivanova A. S., Omelchenko S. S., Kotliarova E. V., Matyukhin V. V. Kalibrovka parametrov modeli rascheta matricy correspondencii [Calibration of model parameters for calculating correspondence matrix for Moscow] // Computer research and modeling. — 2020. — Vol. 12, No. 5. — P. 961–978 (in Russian).
- Комлярова Е. В., Кривошеев К. Ю., Гасникова Е. В., Шароватова Ю. И., Шурупов А. В.* Обоснование связи модели Бэкмана с вырождающимися функциями затрат с моделью стабильной динамики // Компьютерные исследования и моделирование. — 2022. — Т. 14, № 2. — С. 335–342.

- Kotlyarova E. V., Krivosheev K. Yu., Gasnikova E. V., Sharovatova Yu. I., Shurupov A. V. Obosnovanie svyazi modeli Beckmana s vyrozhdayushchimisya funkciyami zatraat s model'yu stabil'noi dinamiki [Proof of the connection between the Beckman model with degenerate cost functions and the model of stable dynamics] // Computer research and modeling. — 2020. — Vol. 14, No. 2. — P. 335–342 (in Russian).
- Швецов В. И. Математическое моделирование транспортных потоков // Автоматика и телемеханика. — 2003. — № 11. — С. 3–46.
- Shvetsov V.I. Mathematical modeling of traffic flows // Automation and Remote Control. — 2003. — Vol. 64. — P. 1651–1689. — DOI: 10.1023/A:1027348026919 (Original Russian paper: Shvetsov V.I. Matematicheskoe modelirovaniye transportnyh potokov // Avtomatika i telemekhanika. — 2003. — No. 11. — P. 3–46.)
- Abrahamsson T., Lundqvist L. Formulation and estimation of combined network equilibrium models with applications to Stockholm // Transportation Science. — 1999. — Vol. 33, No. 1. — P. 80–100.
- Anderson S. P., Palma A. D., Thisse J.-F. Discrete choice theory of product differentiation. — MIT Press, 1992. — 454 p.
- De Cea J., Fernández J. E., Dekock V., Soto A. Solving network equilibrium problems on multimodal urban transportation networks with multiple user classes // Transport Reviews. — 2005. — Vol. 25, No. 3. — P. 293–317.
- Evans S. P. Derivation and analysis of some models for combining trip distribution and assignment // Transportation research. — 1976. — Vol. 1, No. 1. — P. 37–57.
- Frank M., Wolfe P. An algorithm for quadratic programming // Naval research logistics quarterly. — 1956. — Vol. 3, No. 1–2. — P. 95–110.
- Gasnikova E., Gasnikov A., Kholodov Y., Zukhba A. An evolutionary view on equilibrium models of transport flows // Mathematics. — 2023. — Vol. 11, No. 4. — P. 858.
- Fernández E., De Cea J., Florian M., Cabrera E. Network equilibrium models with combined modes // Transportation Science. — 1994. — Vol. 28, No. 3. — P. 182–192.
- Kubentayeva M. et al. Primal-dual gradient methods for searching network equilibria in combined models with nested choice structure and capacity constraints // arXiv preprint. — 2023. — arXiv:2307.00427
- Nesterov Yu., De Palma A. Stationary dynamic solutions in congested transportation networks: summary and perspectives // Networks and spatial economics. — 2003. — Vol. 3, No. 3. — P. 371–395.
- Ortúzar J. de D., Willumsen L. G. Modelling transport. — John Wiley & Sons, 2011. — 751 p.
- Patriksson M. The traffic assignment problem: models and methods. — Courier Dover Publications, 2015.
- Pigou A. C. The economics of welfare. — Macmillan, 1920.
- Sheffi Y. Urban transportation networks: equilibrium analysis with mathematical programming methods. — Prentice-Hall, N.J.: Englewood Cliffs, 1984. — 399 p.
- Stabler B., Bar-Gera H., Sall E. Transportation networks for research core team // Transportation Networks for Research. — 2018. — [Electronic resource]. — <https://github.com/bstabler/TransportationNetworks> (accessed: 16.09.2023).
- Wardrop J. Some theoretical aspects of road traffic research // Proceedings of the Institute of Civil Engineers Part II. — 1952. — Vol. 1. — P. 325–378.