

Введение в математическое моделирование транспортных потоков

Под редакцией А. В. Гасникова

Издание второе, исправленное и дополненное

Предисловие руководителя Департамента транспорта и развития
дорожно-транспортной инфраструктуры г. Москвы
М. С. Ликсутова

*Рекомендовано Учебно-методическим объединением
высших учебных заведений Российской Федерации
по образованию в области прикладных математики и физики
в качестве учебного пособия для студентов вузов по направлению
«Прикладные математика и физика»*

Москва
Издательство МЦНМО
2013

УДК 519.1, 519.2, 519.6, 519.8(075)

ББК 22.1я73

В24

Оглавление

Авторский коллектив: А. В. Гасников, С. Л. Кленов, Е. А. Нурминский, Я. А. Холодов, Н. Б. Шамрай. Приложения: М. Л. Бланк; К. В. Воронцов, Ю. В. Чехович; Е. В. Гасникова; А. А. Замятин, В. А. Малышев; А. В. Колесников; Ю. Е. Нестеров, С. В. Шпирко; А. М. Райгородский. Практическое приложение: А. В. Прохоров, В. Л. Швецов. Вступительное слово: руководитель Департамента транспорта г. Москвы М. С. Ликсутов

Рецензенты:

Лаборатория волновых процессов механико-математического факультета МГУ им. М. В. Ломоносова (зав. лаб. проф. Н. Н. Смирнов, зам. декана мехмата МГУ);

к.ф.-м.н. В. И. Швецов (Институт системного анализа РАН)

Научный консультант академик А. А. Петров

В24 Введение в математическое моделирование транспортных потоков: Учебное пособие / Издание 2-е, испр. и доп. А. В. Гасников и др. Под ред. А. В. Гасникова. — М.: МЦНМО, 2013. — ??? с.

ISBN 978-5-4439-0040-7

Излагается математический аппарат и некоторые физические концепции, которые могут пригодиться при создании (модернизации) комплексной интеллектуальной транспортной системы (КИТС). Предназначено для студентов старших курсов и аспирантов физико-математических специальностей (МФТИ, МГУ, Независимого московского университета). Рекомендуется также научным работникам, интересующимся вопросами математического моделирования.

Первое издание вышло в 2010 году в издательстве МФТИ.

ББК 22.1я73

© Коллектив авторов, 2013

© МЦНМО, 2013

ISBN 978-5-4439-0040-7

Вступительное слово руководителя Департамента транспорта и развития дорожно-транспортной инфраструктуры г. Москвы М. С. Ликсутова	6
Предисловие к новому изданию	10
Предисловие	12
Введение	19

Глава 1. Исследование транспортных потоков с помощью теории экономического равновесия	25
Введение	25
1.1. Задача транспортного равновесия	26
1.1.1. Моделирование транспортных потоков как задача принятия решений	26
1.1.2. Формализация проблемы	27
1.1.3. Сведение к вариационному неравенству	30
1.1.4. Разрешимость задач транспортного равновесия	33
1.1.5. Симметричные задачи транспортного равновесия	36
1.2. Построение функций транспортных затрат	37
1.2.1. Аддитивные функции затрат	38
1.2.2. Неаддитивные функции затрат	40
1.2.3. Модель стационарной динамики	41
1.3. Соотношение между системным оптимумом и конкурентным равновесием	43
1.4. Численные методы решения задач транспортного равновесия	49
1.4.1. Проективные методы решения задачи транспортного равновесия	50
1.4.2. Декомпозиция проективных методов для поиска равновесных потоков	52
1.4.3. Проективный метод с генерацией маршрутов	53
1.4.4. Ступенчатая регулировка шага проективного метода	56
1.5. Построение матрицы корреспонденций	58
1.5.1. Гравитационная модель	59
1.5.2. Энтропийная модель	60
1.5.3. Связь между гравитационной и энтропийной моделями	64
1.6. Парадоксы транспортного равновесия	66
1.6.1. Парадокс Браесса	66
1.6.2. Транспортно-экологические парадоксы	68
1.7. Практическая работа	72

Литература	75
Глава 2. Математические модели транспортных потоков	79
2.1. Макроскопические модели	79
2.1.1. Модель Лайтхилла—Уизема—Ричардса (LWR)	79
2.1.2. Модель Танака	89
2.1.3. Модель Уизема	90
2.1.4. Модель Пейна и её обобщения	97
2.1.5. Кинетические модели	102
2.1.6. Практические приложения моделей	104
2.2. Микроскопические модели	106
2.2.1. Модель оптимальной скорости Ньюэлла	106
2.2.2. Модель следования за лидером «Дженерал Моторс»	111
2.2.3. Модель «разумного водителя» Трайбера	113
2.2.4. Модели клеточных автоматов	115
2.3. Модельные задачи	121
2.3.1. Эволюции глобального затора в транспортном потоке, описываемом моделями LWR и Уизема	121
2.3.2. Эволюции локального затора в транспортном потоке, описываемом моделями LWR и Уизема	135
2.3.3. Задача о светофоре (при каких условиях перед светофором не будет скапливаться очередь)	141
2.4. Теория Кернера—Конхойзера движущихся локальных кластеров в моделях класса «Дженерал Моторс»	143
2.4.1. Фундаментальные эмпирические свойства перехода от свободного транспортного потока к плотному и модели транспортного потока	144
2.4.2. Характеристические параметры широкого движущегося кластера	149
2.4.3. Линия J Кернера	150
Литература	153
Глава 3. Теория Кернера трех фаз в транспортном потоке — новый теоретический базис для интеллектуальных транспортных технологий	163
3.1. Три фазы транспортного потока	165
3.1.1. Предварительные сведения	165
3.1.2. Свободный транспортный поток — фаза F	166
3.1.3. Плотный транспортный поток	166
3.1.4. Определение фаз J и S в плотном транспортном потоке	167
3.1.5. Возникновение плотного потока (traffic breakdown) — фазовый переход $F \rightarrow S$	169
3.1.6. Бесконечное число пропускных способностей скоростной автомагистрали	174
3.1.7. Широкие движущиеся кластеры (локальные движущиеся заторы) — фаза J	176
3.1.8. Синхронизированный транспортный поток — фаза S	177

3.1.9. Фазовый переход $S \rightarrow J$	178
3.1.10. Неоднородные пространственно-временные структуры транспортного потока, состоящие из фаз S и J	179
3.2. Стохастические модели в рамках теории трех фаз Кернера	181
3.2.1. Стохастическая микроскопическая трехфазная модель транспортного потока	182
3.2.2. Моделирование свойств пространственно-временных структур в транспортном потоке вблизи въезда на скоростную автомагистраль	186
3.2.3. Трехфазная модель клеточных автоматов для транспортного потока (ККВ-модель)	189
3.2.4. Новая трехфазная модель клеточных автоматов для транспортного потока (ККШ-модель)	191
3.3. Применение теории трех фаз Кернера для интеллектуальных транспортных технологий	193
Литература	194

Приложения 200

<i>М.Л. Бланк.</i> Процессы с запретами в моделях транспортных потоков	200
<i>К.В. Воронцов, Ю.В. Чехович.</i> Интеллектуальный анализ данных в задачах моделирования транспортных потоков	226
<i>Е.В. Гасникова.</i> О возможной динамике в модели расчета матрицы корреспонденций	250
<i>А.А. Замятин, В.А. Малышев.</i> Введение в стохастические модели транспортных потоков	273
<i>А.В. Колесников.</i> Транспортная задача и концентрация	307
<i>Ю.Е. Нестеров, С.В. Шпирко.</i> Стохастическое транспортное равновесие	317
<i>А.М. Райгородский.</i> Модели случайных графов и их применения	328

Задачи 352

Задачи к главам пособия и приложениям	352
Задача Штейнера и задачи на графах транспортных сетей	377
Задачи от Яндекс.Пробки	390

Исследовательские вычислительные задачи, предлагавшиеся в 2011 г. 397

Практическое приложение 417

<i>А.В. Прохоров, В.Л. Швецов.</i> О практическом опыте моделирования транспортных потоков с помощью пакета программ PTV Vision ®	417
---	-----

Вступительное слово

Транспорт — одна из ключевых систем городского организма, которую по важности уместно сравнить с кровоснабжением. Именно транспорт позволяет городу в полной мере выполнять связующую, коммуникационную и обеспечивающую функции. Тема транспорта касается практически каждого городского жителя, и тем важнее становятся усилия по систематизации и распространению соответствующих знаний.

Для управления дорожным движением на транспортной сети городов повсеместно используются системы управления, алгоритмы работы которых основаны на моделях транспортных потоков. Требования к точности и сложности моделей чрезвычайно велики. Достаточно сказать, что на простейшем перекрестке может быть 12 направлений движения транспортных средств. Для участка улично-дорожной сети с 10 такими пересечениями речь идет уже о 120 направлениях и необходима минимизация задержек по каждому из этих направлений при условии, что интенсивность движения постоянно изменяется во времени и в пространстве.

Кроме того, без транспортного моделирования невозможно планирование строительства новых и модернизации существующих транспортных объектов, объектов жилищного и делового строительства, схем организации дорожного движения, действий при чрезвычайных ситуациях, решение целого ряда других практических задач.

Вниманию читателей предлагается издание, в котором собраны фундаментальные знания в области математического моделирования транспортных потоков, что позволяет рассматривать это издание как основу подготовки специалистов самой высокой квалификации для многих сфер городского хозяйства.

Одним из самых значительных применений транспортных моделей на перспективу не менее 10–15 лет будет проектирование интеллектуальных транспортных систем (ИТС), необходимость создания которых обусловлена коренными изменениями условий дорожного движения и задачами управления дорожным движением, вызванными, в свою очередь, интенсивным, взрывным ростом уровня автомобилизации.

С точки зрения управления движением при малой загрузке улично-дорожной сети (20–30% пропускной способности) движение фактически является свободным и управление сводится к локальному светофорному регулированию, которое вводится по критериям безопасности. Потребностей в применении каких-либо моделей и алгоритмов управления на их основе практически не возникает.

Интервал в 20–70% загрузки от пропускной способности улично-дорожной сети (УДС) — сфера традиционных автоматизированных систем

управления дорожным движением (АСУДД), когда ставится и решается задача увеличения пропускной способности за счет координированного управления светофорной сигнализацией.

Основное физическое явление, за счет которого достигается выигрыш при таком управлении — формирование и пропуск пачки (группы) транспортных средств. Если пачка «рассыпается» (при дистанции между светофорами 800–1000 м и более), выигрыш за счет координированного светофорного регулирования фактически не достигается. Отсюда вытекают требования к количеству светофорных объектов, к алгоритмам координации и, соответственно, к моделям.

При загрузке в 80% и более задача управления принципиально меняется. Любая перегрузка улично-дорожной сети сверх пропускной способности приводит к фатальным последствиям.

МКАД (Московская кольцевая автомобильная дорога), например, спроектирована на пропуск не менее 8–10 тысяч единиц транспорта в час. Но в настоящее время каждый день имеются участки, где пропускная способность падает до двух тысяч и менее единиц (ниспадающая ветвь фундаментальной диаграммы транспортного потока). При этом однажды возникшая пробка подвержена явлению гистерезиса, которое состоит в том, что как только трафик превышает пропускную способность полосы, движение входит в нестабильную зону функционирования в загруженном режиме и вернуться к свободному движению возможно лишь после того, как «спрос» станет явно ниже пропускной способности. Никакие алгоритмы координации (и соответствующие потоковые модели) в условиях перегрузки положительного результата не дают.

Эффективное управление дорожным движением в этих условиях должно обеспечивать загрузку транспортной сети на грани ее пропускной способности и поддерживать непрерывное равномерное движение — пусть даже на относительно небольших скоростях (оптимальные режимы движения при такой загрузке собственно и соответствуют синхронизированной фазе по Кернеру).

Иными словами, задача пропуска возможно большего количества транспортных средств меняется на задачу достижения транспортного баланса между реальной пропускной способностью УДС и спросом на объемы движения при максимальном использовании возможностей, предоставляемых геометрическими параметрами уличной сети.

Такая постановка задачи принципиально меняет построение системы управления, алгоритмы управления и, соответственно, модели, лежащие в их основе.

Прежде всего, необходимо понимать, что пропускная способность — величина переменная, зависящая от погодных условий, аварий, производства ремонтных работ и т.д. При загрузке до 60–70% имеется резерв,

который сглаживает изменения пропускной способности. При загрузке в 90% такого резерва нет и интеллектуальная транспортная система должна в реальном масштабе времени оценивать текущую пропускную способность и перераспределять потоки.

Далее, в условиях интенсивной автомобилизации достижение транспортного баланса невозможно без введения и поддержания механизмов ограничения спроса на дорожное движение за счет информирования участников движения о загрузке УДС и возможных маршрутах движения, без развития общественного транспорта, грамотной логистики, управления парковочным пространством, перераспределения транспортных потоков в зависимости от складывающихся условий и т. д. — вплоть до введения административных запретов.

Например, в Москве в настоящее время на улицах города в «часы пик» могут одновременно находиться в движении не более 400 тысяч автомобилей. Как показывают результаты наблюдения, город «встает», когда число выехавших на дороги автомобилей достигает 500 тысяч, а при экстремальных погодных условиях — и при меньшей численности. Данную цифру можно рассматривать как предельную пропускную способность улично-дорожной сети города.

Между тем по экспертным оценкам при благоприятных условиях движения в часы пик готовы выехать на улицы около четверти от общего числа транспортных средств. Учитывая, что в Москве с пригородами насчитывается более 5 млн. единиц транспорта, общий спрос на передвижение в часы пик можно оценить в 1 млн. единиц транспорта, что в два раза превышает пропускную способность улично-дорожной сети. Каждый второй водитель уже не садится за руль только из-за неудовлетворительных условий движения и парковки.

На перспективу до 2016 года прогнозируется, что численность автотранспортных средств в Московском регионе может увеличиться до 8 млн. единиц, т. е. спрос может четырехкратно превысить пропускную способность.

Примерно такая же ситуация, обусловленная интенсивной автомобилизацией в последние 10–15 лет, складывается в большинстве крупных и средних городов страны. В реально сложившихся условиях никакое наращивание дорожно-мостового строительства и принятие локальных мер не позволит удовлетворить «отложенный спрос».

Именно поэтому в современных условиях задача пропуска возможного большего числа транспортных средств по УДС меняется на задачу поддержания транспортного баланса между пропускной способностью существующей улично-дорожной сети и ее реальной загрузкой за счет перераспределения (а при необходимости — за счет введения ограничений на движение) транспортных потоков.

Несомненно, что инженерная реализация рассматриваемых моделей в виде реально работающих автоматизированных систем требует своих исследователей и разработчиков.

Но тот факт, что относительно небольшой набор математических идей «делает погоду» в разных областях транспортной науки, позволяет утверждать, что изучение собранных в монографии материалов может быть рекомендовано как ученым, так и инженерам, применяющим в сфере своей деятельности аппарат математического моделирования транспортных потоков.



Максим Станиславович Ликсутов
Руководитель Департамента транспорта и развития
дорожно-транспортной инфраструктуры города Москвы
12 апреля 2012 г.

Предисловие к новому изданию

В 2010 г. в издательстве МФТИ небольшим тиражом (250 экз.) вышла наша книга «Введение в математическое моделирование транспортных потоков». Книга оказалась довольно востребованной и практически сразу стала библиографической редкостью. Предложение о подготовке нового издания поступило к нам от И. В. Яценко и В. В. Фурина летом 2011 г. во время ЛШСМ-2011. На этой замечательной летней школе мы как раз рассказывали продвинутым школьникам ряд сюжетов из первого издания книги. В новом издании была учтена «обратная связь» от школьников ЛШСМ-2011 и от студентов, слушавших одноименный (с названием книги) курс лекций в весеннем семестре 2011 г. в Независимом московском университете.

По сравнению с прошлым изданием в новом издании были пополнены материалы главы 1, главы 3, приложения М. Л. Бланка, Е. В. Гасниковой, А. В. Колесникова, А. М. Райгородского и задачного раздела. В новое издание также были добавлены приложения:

- «Интеллектуальный анализ данных в задачах моделирования транспортных потоков», написанное известными специалистами в области интеллектуального анализа данных профессором К. В. Воронцовым (ВЦ РАН, МФТИ, «Яндекс») и к.ф.-м.н. Ю. В. Чеховичем (ВЦ РАН, «Форексис»);
- «Стохастическое транспортное равновесие», написанное известными специалистами в области численных методов выпуклой оптимизации профессором Ю. Е. Нестеровым (CORE/INMA (UCL), ПреМоЛаб МФТИ) и к.ф.-м.н. С. В. Шпирко (ПреМоЛаб МФТИ);
- «О практическом опыте моделирования транспортных потоков с помощью пакета программ PTV Vision», написанное А. В. Прохоровым («А+С Консалт») и В. Л. Швецовым («А+С Консалт») на основе собственного опыта выполнения различных транспортных проектов в ряде городов России.

Небольшие изменения были внесены и в другие части. За время, прошедшее с момента выхода первого издания, к нам обращались за консультациями по вопросам математического моделирования транспортных потоков разнообразные организации. Большинство задач, по которым требовалась консультация, имели интересную научную составляющую. Поэтому было решено включить в новое издание ряд таких исследовательских вычислительных задач, знакомство с которыми, на наш взгляд, значительно способствует лучшему усвоению основного материала.

Нельзя не отметить большой труд редакционного характера, затраченный при подготовке первого издания В. Н. Тарасовым, В. А. Дружининой, И. А. Волковой, О. П. Котовой, Л. В. Себовой. Всем им мы выражаем глубокую благодарность.

Мы также благодарим Юрия Николаевича Торхова, способствовавшего выходу настоящего издания.

Работа выполнена при поддержке гранта РФФИ и Лаборатории структурных методов анализа данных в предсказательном моделировании (ПреМоЛаб) МФТИ, грант правительства РФ дог. 11 11.G34.31.0073.

А. В. Гасников (avgasnikov@gmail.com)
доцент кафедры МОУ ФУПМ МФТИ
с.н.с. ПреМоЛаб МФТИ
5 декабря 2011 г.

Предисловие

Идея написания этого пособия принадлежит декану факультета управления и прикладной математики (ФУПМ) МФТИ профессору Александру Алексеевичу Шананину. Более двух лет назад он предложил начать читать на физтехе курс по выбору «Введение в математическое моделирование транспортных потоков», некоторые детали которого (глава 2) к тому моменту уже обсуждались в течение нескольких лет на семинаре «Квазилинейные уравнения и обратные задачи» в ВЦ РАН под его руководством¹⁾. Основная цель курса заключалась в том, чтобы познакомить заинтересованных студентов-старшекурсников и аспирантов физико-математических специальностей с математикой, необходимой для решения, например, таких задач:

- эволюция затора (как будет распространяться информация о заторе по транспортному потоку),
- задача о светофоре (при каких условиях перед светофором не будет скапливаться очередь),
- задача о выборе оптимальной топологии транспортной сети (где и какую дорогу «лучше» строить),
- расчет матрицы корреспонденций и распределения потоков,
- задача о надежности графа транспортной сети.

Курс содержал дополнительные главы следующих дисциплин:

- уравнений математической физики (обобщенные решения законов сохранения, групповой анализ, автомодельная редукция, принципы максимума для квазилинейных параболических уравнений);
- теории вероятностей и случайных процессов (аппарат производящих функций, системы массового обслуживания, концентрация меры, исследование асимптотик с помощью метода перевала);
- функционального анализа (сжимающие отображения, монотонные операторы, конусные методы);

¹⁾Здесь также хотелось бы обратить внимание на ту огромную роль, которую сыграл зав. кафедрой вычислительной математики МФТИ член-корреспондент РАН А.С. Холодов во внедрении этой тематики в образовательный процесс на физтехе. За несколько лет до того, как был поставлен упомянутый курс, Александр Сергеевич уже читал лекцию по гидродинамическим моделям транспортного потока в рамках своего семестрового курса для студентов ФУПМ МФТИ «Нелинейные вычислительные процессы». Энтузиазм Александра Сергеевича «зажег» тогда многих (и не только студентов). Отметим также, что с 2005 года В.И. Швецов ведёт курс «Математические модели транспортных потоков» для студентов ФУПМ МФТИ на базовой кафедре Института системного анализа РАН.

- теории динамических систем (методы функционалов Ляпунова) и эргодической теории (концентрация инвариантной меры, элементы статистической физики);
- кинетической теории (уравнения Колмогорова, социодинамика, динамика систем с мотивацией, самоорганизация);
- теории игр (эволюционные игры: равновесие Нэша как устойчивое положение равновесия динамики наилучших ответов);
- оптимизации в конечномерных и бесконечномерных пространствах (принцип Лагранжа, двойственность, отделимость, принцип Беллмана, элементы теории управления);
- дискретной математики (задачи на графах и эффективные (приближенные, вероятностные) алгоритмы их решения);
- численных методов выпуклой оптимизации (прямо-двойственные методы, стохастические субградиентные методы, субградиентные методы для задач огромной размерности и др.).

Настоящее пособие представляет собой попытку связно преподнести как материалы прочитанных курсов, так и в целом математический аппарат и некоторые «физические концепции», которые могут пригодиться при создании (модернизации) комплексной интеллектуальной транспортной системы (КИТС). О важности такой системы в «борьбе с пробками» (в Москве) было много сказано за последнее время.

По сути, речь идет о том, как оптимальным образом использовать имеющуюся информацию. Например, в Москве сейчас установлено (в основном на крупных перекрестках) в общей сложности более 500 видеокамер и несколько тысяч различных детекторов. Порядка 10^5 автомобилей¹⁾, курсирующих по Москве и области, оснащены GPS-навигаторами, что позволяет получать треки (пути следования) автомобилей с информацией о скоростях движения вдоль этих треков. Заметим, что всего в Москве ежедневно бывает более $4 \cdot 10^6$ автомобилей.

Создание КИТС на основе имеющихся данных предполагает выполнение следующих действий.

- Выработку адекватной (имеющимся данным²⁾) математической модели, описывающей транспортный поток. Например, можно в неплохом приближении уподоблять транспортный поток сжимаемой жидкости с мотивацией и использовать гидродинамические модели (или модели клеточных автоматов). Калибровка таких моделей на прямолинейных участках

¹⁾Если говорить о сечении по времени, то таких автомобилей на дорогах будет на порядок меньше.

²⁾Важно подобрать модель, адекватную имеющимся данным, дабы не «забывать микроскопом гвозди».

дороги (ребрах графа транспортной сети) довольно просто осуществляется исходя из имеющейся информации.

- Для постановки начально-краевых условий: описание характеристик источников и стоков автомобилей, узлов графа транспортной сети (перекрестки, въезды, съезды и т. п.) — также требуется работа с накопленными данными. В результате такой работы получается матрица корреспонденций, на основе этой матрицы рассчитываются распределения потоков, а затем и матрицы перемешивания в узлах графа транспортной сети. На наш взгляд, адекватная постановка начально-краевых условий — это одна из самых сложных текущих задач. И наметки приведенного здесь пути — далеко не единственный способ получения краевых условий.

- Откалиброванная модель может использоваться для локального (по времени) управления на основе текущей информации, например, светофорной сигнализацией (въездами на крупные магистрали). Возникающие здесь задачи связаны с управлением сложными (гибридными) динамическими системами в условиях неопределенности. Такое локальное управление позволит несколько разгрузить складывающуюся на дорогах в данный момент ситуацию (правильным образом распределяя ресурсы транспортной сети между ее пользователями). Например, в Калифорнии коллектив, работающий в Беркли и возглавляемый П. Варайя и А. Б. Куржанским, предложил несколько лет назад способ локального управления въездами на основные магистрали. Это привело к тому, что для среднестатистического водителя время в пути уменьшилось на 30%.

- Помимо задач локального управления имеются задачи долгосрочного управления. Где и какую дорогу следует построить при заданных бюджетных ограничениях? Каким образом (в каком размере) взимать плату за проезд с трасс в центре Москвы?¹⁾ На каких трассах стоит в первую очередь увеличивать число полос? Где стоит в первую очередь переделять развязки или делать новые (в частности, решать вопрос: а выгодно ли увеличивать степень непланарности графа транспортной сети)? Эти задачи, так же как и задачи предыдущего пункта, должны решаться для всей сети в целом (не локально по пространству!). Иначе говоря, сумматорные функционалы качества критериев должны в себя включать всех участников дорожного движения. Понятно, что для решения этих задач достаточно просматривать различные сценарии (в том числе предложенные руководством города) с помощью выработанной и откалиброванной модели на предмет их состоятельности, путем (разумного) перебора выбирать лучшие предложения по разгрузке ситуации на дорогах.

¹⁾Нужна некая «золотая середина»: с одной стороны, плата за проезд должна разгрузить эти трассы, с другой — желательно, чтобы пропускные ресурсы трасс использовались по максимуму.



Рис. 1. Транспортный затор на одной из улиц Москвы

Некоторые рассмотренные в пособии задачи имеют также и коммерческий выход. Например, актуальной в последнее время задачей¹⁾ является *задача маршрутизации*: выбор оптимального (кратчайшего) маршрута следования. Понятно, что если считать веса ребер графа транспортной сети известными и не меняющимися со временем, то эта задача довольно эффективно решается. Но на практике далеко не всегда все нужные веса ребер бывают известными. В зависимости от времени суток ситуация на дорогах может кардинально меняться, поэтому возникает необходимость прогнозирования загрузки элементов сети. Примером таких изменений служит образование заторов в «часы пик» — за короткий промежуток времени движение может быть практически парализовано даже на многополосных магистралях (рис. 1).

Несмотря на отмеченную актуальность приведенных выше задач, еще раз подчеркнем, что в пособии изложен в основном лишь математический аппарат и некоторые физические концепции, которые могут пригодиться для их решения. Важно также заметить, что формат пособия не предполагал включения технически сложных вещей, обремененных большим количеством деталей. Тем не менее по возможности мы старались хотя бы на концептуальном уровне разъяснять практически все основные нюансы. Как следствие, пособие вобрало в себя довольно много материала (который не удавалось рассказать студентам меньше чем за год), и в ходе его подготовки было использовано более 500 литературных источников, многие из которых впоследствии было решено привести в пособии. По-

¹⁾Ее решение может быть интересно, например, НИС ГЛОНАСС, ЗАО «Российские навигационные технологии», различным интернет-службам, следящим за пробками на дорогах и компаниям, производящим КПК-навигаторы (с выходом в интернет) для автомобилей.

следнее обстоятельство также не характерно для учебных пособий, но при выбранном уровне детализации и объеме излагаемого материала вполне уместно.

Конечно, представленный в книге материал далеко не полон. Причина проста — колоссальный объем накопившегося на данный момент материала, посвященного транспортной проблематике. Достаточно сказать, что сейчас в мире существуют десятки реферируемых научных журналов, в которых регулярно публикуются материалы на транспортную тематику. Упомянем лишь некоторые из них: «Transportation Research B», «Physical Review E», «Review of modern physics», «Transportation Science», не говоря уже об электронных ресурсах, таких, например, как <http://arxiv.org/>. Раз в два года проводится крупнейшая в транспортном сообществе конференция по математическому моделированию транспортных потоков и смежным вопросам: «Traffic and granular flow», труды которой публикует известное немецкое издательство Springer. Кстати, в 2011 г. эта конференция впервые прошла в Москве (см. <http://tgf11.ru>). Четвертый номер журнала «Труды МФТИ» за 2010 г. под редакцией вице-президента РАН академика В. В. Козлова всецело посвящен транспортной проблематике. Однако, несмотря на вышесказанное, мы все же постарались собрать наиболее базовые (математически) вещи и описать текущее состояние дел.

Об авторах. Во многом определяющим моментом в создании этого пособия стало участие в его написании ряда ведущих специалистов в своих областях. Так, глава 1 написана профессором Е. А. Нурминским и доцентом Н. Б. Шамрай (ИАПУ ДВО РАН) и посвящена применению теории бескоалиционного равновесия для расчета транспортной сети при условии стационарности потоков и моделям построения матрицы корреспонденций. Глава 2 написана А. В. Гасниковым при участии доцента С. Л. Кленова (кафедра общей физики МФТИ) и доцента Я. А. Холодова (кафедра вычислительной математики МФТИ). С. Л. Кленовым был написан раздел 2.4, а Я. А. Холодов принял участие в написании пунктов 2.1.4 и 2.2.4. Глава 3, посвященная теории трех фаз Кернера транспортного потока, всецело написана С. Л. Кленовым (коллегой Б. С. Кернера) и содержит как упомянутые выше физические концепции, так и примеры эмпирических (измеренных) пространственно-временных структур плотного потока на скоростных автомагистралях. Как показала обратная связь от студентов, слушавших упомянутый выше курс по выбору, востребованными оказались «стохастические» приложения доцента А. А. Замятина и профессора В. А. Мальшева (кафедра теории вероятностей мехмата МГУ) и профессора А. М. Райгородского (кафедра математической статистики и случайных процессов мехмата МГУ, кафедра анализа данных МФТИ). Важную роль в пособии играют эргодические приложения профессора М. Л. Бланка (лаборатория

Р. Л. Добрушина ИППИ РАН), аспирантки Е. В. Гасниковой (кафедра анализа систем и решений ФУПМ МФТИ) и приложение д.ф.-м.н. А. В. Колесникова (математический факультет ВШЭ), посвященное связи задачи Монжа—Канторовича о перемещении масс и явления концентрации меры. Эти три приложения, помимо того что представляют самостоятельную ценность, также завязывают (математически) между собой многие темы этого пособия. Другими словами, знакомство с ними желательно для формирования целостного восприятия.

В конце учебного пособия приведены задачи, часть из которых в разное время предлагалась студентам. При подготовке задач большую помощь оказали молодые ученые, работающие в близких направлениях. В пособии имеется целый раздел задач (написанный ассистентом кафедры МОУ ФУПМ МФТИ Е. Г. Молчановым), посвященный задачам на графах и, по сути, восполняющий нехватку в пособии темы «Транспортные потоки и Computer Science». В этом же разделе приводится довольно интересная задача, пришедшая из практических приложений от службы «Яндекс.Пробки».

Благодарности. В заключение хотелось бы выразить благодарность профессору А. А. Шананину, академику В. В. Козлову, академику А. А. Петрову, члену-корреспонденту А. С. Холодову. Общение и участие в мероприятиях, к которым они имели отношения, всегда доставляло большое удовольствие и иногда вдохновляло на улучшение материала данной книги. Ценную обратную связь при подготовке этого пособия получал от профессора А. П. Буслаева и доцента О. С. Розановой, а также от всех коллег, принимавших участие в его написании. Много полезных замечаний по всему тексту сделал самый активный слушатель курса — Ю. В. Дорн (студент 6-го курса ФАКИ МФТИ). Ряд ценных замечаний по темам, изложенным в пособии, сделали С. Я. Аввакумов, В. И. Аркин, Л. Г. Афанасьева, П. П. Бобрик, А. С. Бугаев, Е. В. Булинская, А. М. Валувев, Н. Д. Введенская, В. В. Веденяпин, И. Е. Виноградов, К. А. Волосов, А. Э. Воробьев, А. И. Голиков, А. Н. Дарьин, К. Дафермос, А. В. Дмитрук, В. А. Дородницын, В. А. Дружинина, В. Г. Жадан, А. В. Казейкина, Б. С. Кернер, А. В. Козлов, В. Ф. Колчин, Н. С. Кукушкин, А. Г. Куликовский, А. А. Куржанский, Г. Л. Литвинов, И. А. Лубашевский, А. Е. Макаров, В. П. Мартынов, И. С. Меньшиков, В. Д. Мильман, И. И. Морозов, Т. А. Нагапетян, А. И. Назаров, А. В. Назин, А. С. Немировский, В. И. Опойцев, В. П. Осипов, Я. С. Панасюк, Е. Ю. Панов, Д. И. Петрашко, Н. С. Петросян, С. А. Пирогов, В. М. Полтерович, Б. Т. Поляк, Ю. С. Попков, И. Г. Поспелов, В. Ю. Протасов, В. В. Пухначёв, В. Н. Разжевайкин, И. В. Рублев, А. Н. Рыбко, В. Ж. Сакбаев, А. Ю. Семёнов, Д. Серр, Н. Н. Смирнов, А. Н. Соболевский, П. А. Солоневич, В. Г. Спокойный,

Е. О. Степанов, Н. Н. Субботина, В. Н. Тарасов, С. П. Тарасов, Г. М. Хенкин, Б. Н. Четверушкин, Ю. В. Чехович, А. П. Чугайнова, С. В. Чуканов, Н. Г. Чурбанова, В. М. Шелкович, А. Шень, В. И. Швецов, М. В. Яшина. Ценным было общение с академиком А. Б. Куржанским и академиком В. П. Масловым.

Особо хотелось бы поблагодарить зав. кафедрой математических основ управления (МОУ) МФТИ доцента С. А. Гуза и зам. зав. кафедрой МОУ доцента О. С. Федько, создавших идеальные условия как для проведения занятий, так и для создания пособия, регулярно стимулировавших весь процесс написания и внимательно относившихся ко всем особенностям работы.

А. В. Гасников (avgasnikov@gmail.com)
доцент кафедры МОУ ФУПМ МФТИ
12 ноября 2010 г.

Введение

В 50-е годы прошлого века, в связи с исследованиями процессов, возникающих при взрыве бомбы, наблюдалось бурное развитие газовой динамики (обобщенные решения законов сохранения, устойчивые разностные схемы расчета решений). Тогда же появились первые макроскопические (гидродинамические) модели, в которых транспортный поток уподобляется потоку «мотивированной» сжимаемой жидкости (М. Лайтхилл и Дж. Уизем, П. Ричардс), и первые микроскопические модели (следования за лидером), в которых явно выписывается уравнение движения каждого автомобиля (А. Рёшель, Л. Пайпс и др.). В модели Лайтхилла—Уизема (—Ричардса) (1955) транспортный поток уподобляется потоку сжимаемой жидкости и описывается законом сохранения количества (погонной плотности) автомобилей. При этом в модели постулируется существование функциональной зависимости (уравнения состояния) между величиной потока автомобилей (= скорость × плотность) и плотностью. Эту зависимость часто называют *фундаментальной диаграммой* (как правило, вогнутая функция). Собственно, в эту зависимость и «зашиита» мотивация в простейших моделях.

В последующие годы класс микро- и макромоделей был значительно расширен. В современном макроскопическом подходе (А. Эу и М. Раскль, 2000) транспортный поток часто описывается нелинейной системой гиперболических уравнений (для плотности и скорости потока) с диффузией (Х. Пейн, Р. Кюне, Б. Кернер и П. Конхойзер). При этом уравнение состояния «зашиивается» во второе уравнение этой системы как стремление водителей двигаться с желаемой скоростью.

В современном микроскопическом подходе преобладают модели типа «разумного водителя», в которых ускорение автомобиля описывается некоторой функцией от скорости этого автомобиля, расстояния до впереди идущего автомобиля (лидера) и скорости относительно лидера (М. Трайбер, 1999). При этом в таких моделях и время может течь дискретно, и сама динамика движения автомобилей может быть стохастической (марковской). Как правило, тогда такие модели называют *моделями клеточных автоматов*. В приложении М. Л. Бланка продемонстрирован один из способов того, как с помощью простейших моделей клеточных автоматов можно получать (математически строго) правдоподобные макроскопические уравнения состояния транспортного потока (например, треугольную фундаментальную диаграмму).

Продолжая аналогию с газовой динамикой, И. Пригожин полвека назад (а затем С. Павери-Фонтана, Д. Хельбинг и др.) предложил описывать транспортный поток кинетическим уравнением (типа Больцмана с «ин-

тегралом взаимодействия автомобилей» вместо «интеграла столкновения частиц газа»). При таком подходе макроскопическая модель получается из кинетической подобно тому, как система уравнений Эйлера получается из уравнения Больцмана.

Отметим ввиду вышесказанного, что задача математически строгого обоснования кинетической модели, исходя из микроскопической, так же как и задача обоснования макроскопической модели, исходя из кинетической, является открытой. Более того, в режимах, соответствующих «фазовому переходу» в транспортном потоке, такое обоснование, по-видимому, принципиально невозможно: нельзя осуществить соответствующий скейлинг, нельзя перейти к динамике средних, нельзя пользоваться эргодичностью системы (инвариантная мера не единственна), неясно, как обрывать (замыкать) моментную цепочку зацепляющихся уравнений. В таких режимах можно лишь нестрого говорить о похожести моделей.

Несмотря на то что с момента появления первых фундаментальных работ прошло более полувека, по мнению ряда известных специалистов в области математического моделирования дорожного движения (К. Нагель, Х. Махмасани, М. Шрекенберг и др.), проблема образования заторных и заторных ситуаций еще до конца не изучена (и сродни проблеме описания турбулентных течений). Используя терминологию, предложенную Б. С. Кернером, можно сказать, что на данный момент нет общепринятого подхода, описывающего поведение движения автотранспорта в области синхронизированного потока. Иначе говоря, если автомобильный поток уподобляется жидкости, то наиболее сложная для моделирования ситуация — это замерзающая жидкость. Подтверждением вышесказанному может служить тот факт, что разные коллективы, занимающиеся моделированием транспортных потоков, как правило, используют разные модели: начиная от модели Лайтхилла—Уизема (А. А. Куржанский и др.), заканчивая моделями, в которых каждый водитель характеризуется своим вариационным принципом (И. А. Лубашевский и др.). Отметим здесь главу 3, в которой приводится «эмпирический базис», т. е. даются свойства реальных пространственно-временных структур, возникающих в плотном транспортном потоке вблизи «узкого места»,¹⁾ для анализа различных подходов к описанию транспортного потока. Важным атрибутом многих современных зарубежных работ, в которых предлагаются математические модели транспортного потока, является проверка предложенных моделей на возможность описания ими трех фаз Кернера транспортного потока, наблюдаемых в многочисленных эмпирических (измеренных) данных.

¹⁾Заметим, что, как правило, исследователи ограничиваются изучением транспортного потока на отдельном прямолинейном участке транспортной сети с простейшими начальными условиями, в то время как причиной заторов (согласно К. Даганзо) часто являются «узкие места» (перекрестки, вьезды).

Математическая теория управления транспортными потоками, как уже упоминалось выше, сейчас активно развивается в работах калифорнийской школы, возглавляемой П. Варайя и А. Б. Куржанским. Исходя из модели клеточных автоматов К. Даганзо (1994), которую можно представить как «схема Годунова + модель Лайтхилла—Уизема + треугольная фундаментальная диаграмма», предлагается способ оптимального управления светофорами и вьездами на магистралях в Калифорнии. Здесь стоит обратить внимание на соизмеримость грубости выбранной модели, качества имеющихся данных (см., например, <http://pems.dot.ca.gov/>) и простоты работы с этой моделью. Поясним основную идею того, как следует управлять. Из фундаментальной диаграммы следует, что одному и тому же значению потока автомобилей соответствуют разные (как правило, две) плотности и, как следствие, разные скорости. Очевидно, что более выгодным режимом является режим с большей скоростью. Задача управления (скажем, светофорами или вьездами на основные магистрали) заключается в том, чтобы большую часть времени среднестатистический водитель проводил именно в таких режимах.

Подробнее об изложенном выше можно прочитать в главах 2 и 3.

Из-за сильной неустойчивости (при достаточно больших плотностях) решений уравнений, описывающих транспортные потоки, задача получения достоверного прогноза загрузки транспортной сети по имеющимся данным на час вперед сродни задаче получения достоверного прогноза погоды на неделю вперед. При этом вычислительные мощности современных высокопроизводительных кластеров в терафлопс и выше позволяют просчитывать реальную ситуацию по Москве (в которой, напомним, порядка трех миллионов автомобилей) со значительным опережением реального времени. Другими словами, основной проблемой при моделировании транспортных потоков является не ограничение по вычислительным мощностям и ресурсам памяти, а большая чувствительность описываемой реальной транспортной системы к входным данным (характеристики источников и стоков автомобилей) и невозможность собрать достаточно полную информацию о входных данных.

Одним из возможных выходов из этого является рассмотрение в некотором смысле усредненных показателей транспортной системы — например, в смысле теории систем массового обслуживания. Обгоны на многополосной дороге, очереди перед светофорами и многое другое можно описывать таким образом — о чем говорится в приложении А. А. Замятина и В. А. Малышева. В основе моделей этого приложения лежат эргодические марковские процессы. При таком подходе исследователь следит лишь за трендом и «не обращает внимания» на высокочастотные случайные колебания (флуктуации), возможно большой амплитуды, вокруг этого тренда. В связи с упоминанием словосочетания *усредненные пока-*

затели отметим здесь также приложение А. М. Райгородского, в котором исследуются различные свойства случайных графов (транспортных графов, web-графов), например, такое важное свойство, как надежность графа транспортной сети к случайным отказам ребер (отказ ребра означает, что на ребре образовалась пробка). В этих приложениях наблюдается плотная концентрация исследуемых макровеличин (макропоказателей) в маленьких окрестностях своих математических ожиданий. Однако если в приложении Замятина—Малышева мера, которая концентрируется, порождается (как финальная = стационарная) эргодической марковской динамикой, то в приложении А. М. Райгородского она задается непосредственно, скажем, из соображений независимости и однородности (модель случайного графа Эрдеша—Реньи).

В зависимости от того, какая конкретная задача поставлена, следует отдавать предпочтение тому или иному подходу или даже какому-то их сочетанию.

Вернемся теперь к тому, как все-таки ставить начально-краевые условия для целостного описания транспортного потока на полном графе транспортной сети. Будем считать, что есть лишь информация о том, сколько людей живет в том или ином районе и сколько рабочих мест есть в том или ином районе. В главе 1 и приложении Е. В. Гасниковой приведены различные способы обоснования известной на практике энтропийной (гравитационной) модели А. Дж. Вильсона (1967) расчета матрицы корреспонденций (сколько людей, проживающих в районе i , работают в районе j). По сути, матрица корреспонденций определяется как наиболее вероятное макросостояние, в окрестности которого и будет плотная концентрация, стационарной меры «разумной» эргодической марковской динамики, порождающей изучаемую макросистему. Точнее говоря, эргодическая марковская динамика приводит на больших временах к стационарной (инвариантной) пуассоновской (сложной) мере на пространстве макросостояний. Эта мера экспоненциально быстро концентрируется с ростом числа агентов в окрестности наиболее вероятного макросостояния, которое и принимается за положение равновесия макросистемы. Задача поиска наиболее вероятного макросостояния асимптотически (по числу агентов) эквивалентна задаче максимизации энтропийного функционала на множестве (как правило, аффинной структуры), заданном ограничениями — законами сохранения. Приятной особенностью такого класса задач является явная (легко выписываемая) зависимость решения прямой задачи через двойственные переменные. Поскольку число ограничений, как правило, на много порядков меньше числа прямых переменных, то эффективные численные методы базируются на решении двойственной задачи: минимизации выпуклой функции. Отметим, что описанная здесь задача

энтропийно линейного программирования имеет много общего с обычной транспортной задачей.

Далее, исходя из известных потребностей (корреспонденций), водители начинают «нащупывать» некую равновесную конфигурацию потоков (конкурентное равновесие, равновесие Нэша—Вардропа, 1952). Понятно, что корреспонденция не определяет, вообще говоря, однозначно путь следования. Скажем, из МФТИ можно добираться до МГУ разными способами. И если ситуация равновесная, то никому не должно быть выгодно менять свой путь следования — стратегию (ситуация равновесия по Нэшу). Это означает, что времена движения по всем путям, которые хоть кто-нибудь выбрал, соответствующим данной корреспонденции, должны быть одинаковыми. О том, как происходит «нащупывание» равновесия, какие есть обобщения у этой модели и какие есть численные способы решения возникающих по ходу задач оптимизации, написано в главе 1 и в задаче, предложенной Е. В. Гасниковой и Ю. В. Дорном. К счастью, популярный сейчас формат данных о транспортной системе в виде GPS-треков автомобилей позволяет контролировать и тем самым постоянно «подкручивать» выводы рассмотренных моделей и некоторых их важных обобщений.

Имея информацию о том, как распределяются потоки, уже можно получать оценки матриц перемешивания в узлах графа транспортной сети, тем самым замыкая целостную модель. К сожалению, такой способ крайне чувствителен к точности (полноте) входных (обучающих) данных.

В приложении А. В. Колесникова рассматривается задача Монжа—Канторовича о перемещении масс¹⁾, эквивалентная, при весьма общих условиях, задаче Монжа. Оптимальный план перевозок (точнее, потенциал этого отображения) удовлетворяет уравнению в частных производных Монжа—Ампера и порождает метрику Канторовича (—Рубинштейна). С помощью этой метрики устанавливаются довольно тонкие функциональные неравенства о *концентрации меры* (М. Громов, М. Талагран, К. Мартон, М. Леду и др.). Сам термин *концентрации меры*, по-видимому, был впервые предложен В. Д. Мильманом, внесшим значительный вклад в эту область. Геометрически этот принцип можно довольно просто пояснить задачей Пуанкаре—Леви (см. с. 373, 374): площадь многомерной сферы (с выделенным северным и южным полюсом) практически полностью сосредоточена в маленькой полоске вокруг экватора. Этот принцип нашел широкое применение, например, в теории вероятностей (нелинейные законы больших чисел — концентрация значений липшицевых функций в окрестности медианы), асимптотической комбинаторике (например, при исследовании плотной концентрации различных функций, типа числа неза-

¹⁾Сильно связанная с транспортной задачей, о которой написано в главе 1, приложении Е. В. Гасниковой и которая фигурирует также в задачах, приведенных в конце пособия.

висимости, на случайных графах; см. приложение А. М. Райгородского). В понимании ряда материалов пособия явление концентрации меры играет немаловажную роль: концепция равновесия макросистем (модель расчета матрицы корреспонденций), исследования надежности графа транспортной сети (по модели Эрдёша—Реньи), оценка скорости сходимости к равновесию и др.

Глава 1

Исследование транспортных потоков с помощью теории экономического равновесия

Введение

Одним из наиболее агрегированных способов описания транспортных систем является экономический подход, суть которого заключается в соотношении интенсивности использования тех или иных технологий, ресурсов и пр. и итогового результата, выражающего конечный выпуск продуктов, оказание определенного объема услуг и т. п. Схематически этот взгляд на экономику отраслей, в том числе и на транспорт, представлен на рис. 1, где x — это усилия, предпринятые для перевозки, нагрузка на транспортную систему, y — объем грузов или количество людей, перевезенные системой. Помимо этого у системы — черного ящика присутствует социально-экономическая оценка q технологического процесса (x, y).

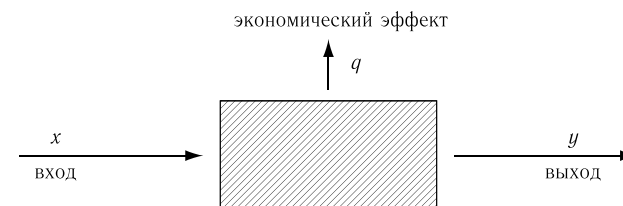


Рис. 1. Представление экономиста о транспорте и не только

Допустимые сочетания затрат x и выпусков y образуют технологическое множество, описанием которого экономист не занимается, в его задачу входит формирование понятия эффективного функционирования системы и отбор и анализ эффективных вариантов.

В данной главе описан один из подходов к моделированию и исследованию транспортных потоков, основанный на теории конкурентного бескоалиционного равновесия, которая позволяет описать достаточно адекватный механизм функционирования автомобильных улично-дорожных сетей (УДС). Будут рассмотрены основные элементы транспортной системы, включающие в себя УДС, факторы, определяющие потребность в перевозках, критерии эффективности транспортной системы и принципы ее функционирования.

Рассматриваемые модели применяются для получения прогнозных оценок загрузки элементов транспортной сети. Подобные задачи интересны в частности тем, что являются одним из инструментов для объективной оценки эффективности проектов по модификации УДС с точки зрения загрузки наиболее проблемных участков дорог и уменьшения общих затрат на передвижение пользователей сети.

1.1. Задача транспортного равновесия

1.1.1. Моделирование транспортных потоков как задача принятия решений

Для определения объемов загрузки УДС в первую очередь необходимо выявить правила, по которым водители выбирают тот или иной маршрут следования. Поведенческие принципы пользователей транспортной сети окончательно были сформулированы в работе [64], где постулировались следующие две возможные ситуации.

1) Пользователи сети независимо друг от друга выбирают маршруты следования, соответствующие их минимальным транспортным расходам.

2) Пользователи сети выбирают маршруты следования исходя из минимизации общих транспортных расходов в сети.

С тех пор в транспортной науке приведенные поведенческие принципы получили названия соответственно *первого* и *второго принципов Вардропа*.

Распределение транспортных потоков согласно первому принципу Вардропа соответствует конкурентному бескоалиционному равновесию, предполагающему совершенный эгоизм участников дорожного движения — каждый стремится достигнуть конечного пункта своей поездки как можно выгоднее для себя и из имеющихся возможных вариантов следования выбирает тот маршрут, по которому будет нести минимальные затраты (временные, финансовые, моральные и т.п.) на проезд. Поэтому данный принцип также называют *пользовательской оптимизацией* (user optimization).

Стоит отдельно отметить, что первый поведенческий принцип предполагает определенные допущения. Во-первых, это совершенная информированность участников движения о ситуации на дорогах — каждый знает затраты на передвижения по тем или иным маршрутам. В настоящее время такое предположение не выглядит изрядной идеализацией, поскольку интенсивно развиваются и внедряются в практическое использование автоматизированные навигаторы и интеллектуальные транспортные системы, идет активное оповещение о ситуациях на дорогах через интернет, радио и другие средства информации. Во-вторых, предполагается

ничтожно малое влияние отдельного участника движения на затраты по всем маршрутам. Хотя такое предположение и заведомо неверно для крупногабаритных транспортных средств, для легковых автомобилей оно представляется достаточно разумным, исключая случаи аварийных ситуаций или неопытных водителей за рулем.

Второй принцип Вардропа предполагает централизованное управление движением в сети. Соответствующее ему распределение транспортных потоков называют *системным оптимумом*, а сам принцип — *системной оптимизацией* (system optimization). Примером пользователей, передвигающихся согласно второму принципу, служат водители маршрутизированного транспорта.

Несмотря на то что приведенные поведенческие принципы широко цитируются как принципы Вардропа, на самом деле чуть ранее их сформулировали Ф. Найт [48] и А. Пигу [60], утверждая, что все участники движения, направляющиеся из одного узла сети в другой, распределяются по различным маршрутам таким образом, чтобы удельные (в расчете на один автомобиль) затраты на проезд были одни и те же для всех.

В ситуации массовой автомобилизации, имеющей место практически во всех странах, подавляющее большинство участников дорожного движения любого города составляют легковые автомобили, совершающие преимущественно маятниковые поездки: место проживания — место работы и обратно. Именно такие поездки создают пиковые нагрузки на УДС, вызывают основные потери времени и других ресурсов, повышают аварийность и усложняют социально-экономическую ситуацию. Вместе с тем маятниковые поездки трудовой миграции имеют ряд особенностей, делающих их удобными для моделирования. В первую очередь в силу их повторяемости характеристики таких поездок можно считать стационарными, а самих водителей — имеющих полную информацию о возможных издержках на различных маршрутах. Более того, разумно предполагать совершенный эгоизм участников этих поездок и стремление нести минимальные потери при проезде. Очевидно также и отсутствие возможностей организовать коалиционное поведение, за исключением разовых акций, которые в общем-то не связаны с ежедневными регулярными поездками. Такое поведение явно соответствует первому принципу Вардропа, поэтому при исследовании загрузки УДС рассмотрим потоки, порождаемые именно легковым частным автотранспортом в утренне-вечерние часы пик.

1.1.2. Формализация проблемы

Исходя из приведенных соображений, построим экономико-математическую модель распределения транспортных потоков в УДС, соответствующую первому поведенческому принципу Вардропа.

Транспортную сеть опишем в виде ориентированного графа $\Gamma(V, E)$, где V — множество вершин, E — множество дуг сети. Каждая дуга соответствует реальному участку автодороги без перекрестков. Каждая вершина представляет узел, разделяющий участки дорог. Направление дуги определяет ход следования автотранспорта. Магистралы с двусторонним движением соответственно имеют парные противоположно ориентированные дуги.

При исследовании потокообразующих факторов в множестве вершин V выделим два подмножества. Первое, $S \subseteq V$, содержит пункты, порождающие потоки; элементы множества S назовем *источниками*. Второе, $D \subseteq V$, содержит пункты, поглощающие потоки; элементы множества D назовем *стоками*. Применительно к задаче моделирования потоков, порождаемых ежедневной трудовой миграцией для утренних часов пик, источниками являются спальные районы и пригороды, стоками — деловые районы города. Множество всех потокообразующих пар представим в виде декартова произведения

$$W = \{\omega = (i, j) : i \in S, j \in D\}.$$

Путем (маршрутом) в сети Γ , соединяющим вершины i и j , назовем последовательность дуг

$$e_1 = (i \rightarrow k_1), e_2 = (k_1 \rightarrow k_2), \dots, e_m = (k_{m-1} \rightarrow k_m), e_{m+1} = (k_m \rightarrow j),$$

где $e_l \in E$ при всех $l = 1, \dots, m + 1$. В маршрутах предполагается отсутствие петель и циклов. Обозначим через P_ω множество альтернативных маршрутов, связывающих пару $\omega \in W$. Совокупность всех путей в сети Γ обозначим через $P = \bigcup_{\omega \in W} P_\omega$.

Перемещаясь от источников к стокам, пользователи сети выбирают тот или иной маршрут следования. Обозначим через x_p величину потока, идущего по пути $p \in P$; тогда загрузку всей сети задает вектор $x = (x_p : p \in P)$.

Преодоление каждого из путей $p \in P$ сопровождается некоторыми затратами (время, топливо, деньги, амортизация автомобиля, износ дороги и т. п.). Количественная характеристика таких затрат зависит от интенсивности и плотности движения в сети. Как правило, в моделях рассматриваются временные или финансовые затраты. Обозначим через G_p удельные затраты пользователей на проезд по пути p . Поскольку на затраты по одному маршруту может влиять загрузка других путей (естественным примером тому служат пересечения главных и второстепенных дорог, дублирующие дороги и т. д.), в общем случае G_p представляют собой функции от загрузки всей сети, то есть $G_p = G_p(x)$.

Во введенных обозначениях первый принцип Вардропы можно формализовать следующим образом. Водители выбирают путь с наименьшими

транспортными расходами, поэтому для каждой пары ω выполнены условия: если по пути $p \in P_\omega$ идет ненулевой поток x_p^\dagger , то затраты по этому пути минимальны, то есть

$$\text{если } x_p^\dagger > 0, \text{ то } G_p(x^\dagger) = \min_{q \in P_\omega} G_q(x^\dagger) = u_\omega(x^\dagger), \quad (1)$$

где $u_\omega(x^\dagger)$ — минимальные транспортные затраты по маршрутам, соединяющим пару $\omega \in W$, при загрузке сети, определяемой вектором x^\dagger .

Соотношения (1), определенные для каждой пары $\omega \in W$, задают так называемые *условия равновесия транспортных потоков*. Потоки $x^\dagger = (x_p^\dagger : p \in P)$, удовлетворяющие условию (1), называются *равновесными*. Для полноты картины необходимо ввести ограничения на допустимость потоков.

Каждой паре источник-сток $\omega = (i, j) \in W$ соответствует свой спрос на перевозку ρ_ω — общий объем пользователей, которые из пункта i должны прибыть в пункт j . Набор $(\rho_\omega : \omega \in W)$ называется *матрицей корреспонденций*. В общем случае предполагается, что корреспонденции являются функциями от минимальных затрат на передвижения в сети, то есть $\rho_\omega = \rho(u_\omega^\dagger)$, где $u_\omega^\dagger = u_\omega(x^\dagger)$.

Традиционно для транспортных задач потоковые переменные должны быть неотрицательными и удовлетворять балансовым ограничениям. Поэтому допустимой областью для рассматриваемых потоков x является множество

$$X(x^\dagger) = \left\{ x \geq 0 : \sum_{p \in P_\omega} x_p = \rho_\omega(u_\omega(x^\dagger)), \omega \in W \right\} \quad (2)$$

(запись $x \geq 0$ означает, что все компоненты $x_p \geq 0$).

Как видно из определения (2), допустимая область $X(x^\dagger)$ является «подвижной» и непосредственно зависит от равновесного распределения потоков x^\dagger . Однако если объемы корреспонденций известны и имеют стационарные значения $\rho_\omega(u_\omega(x^\dagger)) \equiv \rho_\omega$, что вполне характерно для трудовых миграций в УДС, то допустимое множество имеет фиксированную структуру:

$$X = \left\{ x \geq 0 : \sum_{p \in P_\omega} x_p = \rho_\omega, \omega \in W \right\}. \quad (3)$$

Проблема поиска равновесных потоков $x^\dagger \in X(x^\dagger)$ называется *задачей транспортного (потокового) равновесия с эластичным спросом*. При заданных корреспонденциях проблема поиска равновесных потоков $x^\dagger \in X$ называется *задачей транспортного (потокового) равновесия с фиксированным спросом*.

Допустимое множество X , определенное в (3), обладает замечательными свойствами. Во-первых, это полиэдральное¹⁾ и ограниченное множество. Во-вторых, X естественным образом можно представить в виде декартова произведения непересекающихся обобщенных симплексов, т. е. $X = \prod_{\omega \in W} X_\omega$, где

$$X_\omega = \left\{ x_p \geq 0: p \in P_\omega, \sum_{p \in P_\omega} x_p = \rho_\omega \right\}. \quad (4)$$

Эти свойства X качественно влияют на построение теоретического и алгоритмического аппаратов задач транспортного равновесия.

Основной подход к решению и исследованию задач транспортного равновесия состоит в сведении исходной проблемы к эквивалентной оптимизационной задаче или вариационному неравенству. Прямое влияние на то, какая из эквивалентных форм будет рассматриваться, оказывают свойства функций транспортных затрат $G_p(x)$. Для компактности последующего изложения объединим компоненты $G_p(x)$ в вектор-функцию $G(x) = (G_p(x): p \in P)$.

Отметим, что все результаты настоящей работы получены в конечномерном евклидовом пространстве \mathbb{R}^n со скалярным произведением xu и нормой $\|x\| = \sqrt{xx}$, $x, y \in \mathbb{R}^n$. Элементами пространства являются вектор-столбцы, однако знак транспонирования и дополнительные скобки при скалярном умножении будем опускать, чтобы не загромождать запись формул.

1.1.3. Сведение к вариационному неравенству

Предположим, что имеет место непрерывная монотонная зависимость транспортных издержек от объемов загрузки УДС. Традиционно в этом случае поиск равновесных потоков сводится к решению вариационного неравенства (в частном случае к оптимизационной задаче) [23, 39, 50, 54, 59]. Обоснование для такого сведения дает следующий результат.

Теорема 1. *Вектор $x^\dagger \in X(x^\dagger)$ удовлетворяет условию равновесия (1) тогда и только тогда, когда является решением квазивариационного неравенства*

$$G(x^\dagger)(x - x^\dagger) \geq 0 \quad \forall x \in X(x^\dagger). \quad (5)$$

Доказательство. Пусть вектор $x^\dagger = (x_p^\dagger: p \in P) \in X(x^\dagger)$ является решением квазивариационного неравенства (5). Покажем, что в x^\dagger выполнено условие (1). Предположим противное, а именно, что для пары ω

¹⁾ Полиэдральным (многогранным) множеством называется множество решений конечной системы неравенств $Ax \leq b$, это замкнутое выпуклое множество.

существует такой путь $\bar{p} \in P_\omega$, что $x_{\bar{p}}^\dagger > 0$ и $G_{\bar{p}}(x^\dagger) > G_{\bar{q}}(x^\dagger)$ для некоторого $\bar{q} \in P_\omega$, $\bar{q} \neq \bar{p}$. Рассмотрим такой вектор $x^\varepsilon = (x_p^\varepsilon: p \in P)$, что

$$x_p^\varepsilon = \begin{cases} x_p^\dagger, & p \neq \bar{p}, p \neq \bar{q}, \\ x_{\bar{p}}^\dagger - \varepsilon, & p = \bar{p}, \\ x_{\bar{q}}^\dagger + \varepsilon, & p = \bar{q}, \end{cases}$$

где $\varepsilon > 0$ достаточно мало и не нарушает условия неотрицательности $x^\varepsilon \geq 0$. Нетрудно видеть, что $x^\varepsilon \in X(x^\dagger)$, при этом

$$G(x^\dagger)(x^\varepsilon - x^\dagger) = G_{\bar{p}}(x^\dagger)(x_{\bar{p}}^\varepsilon - x_{\bar{p}}^\dagger) + G_{\bar{q}}(x^\dagger)(x_{\bar{q}}^\varepsilon - x_{\bar{q}}^\dagger) = \varepsilon(G_{\bar{q}}(x^\dagger) - G_{\bar{p}}(x^\dagger)) < 0,$$

что противоречит тому, что x^\dagger — решение неравенства (5). Следовательно, в точке x^\dagger условие (1) всегда выполнено.

Покажем обратное, а именно, если $x^\dagger \in X(x^\dagger)$ удовлетворяет условию (1), то x^\dagger — решение квазивариационного неравенства (5). Нетрудно видеть, что для всех $p \in P_\omega$, $\omega \in W$ и $x \in X(x^\dagger)$ выполнены соотношения

$$G_p(x^\dagger) - u_\omega(x^\dagger) \geq 0, \quad (G_p(x^\dagger) - u_\omega(x^\dagger))x_p^\dagger = 0, \quad (G_p(x^\dagger) - u_\omega(x^\dagger))x_p \geq 0,$$

следовательно, имеет место оценка

$$\begin{aligned} 0 &\leq \sum_{\omega \in W} \sum_{p \in P_\omega} (G_p(x^\dagger) - u_\omega(x^\dagger))(x_p - x_p^\dagger) = \\ &= \sum_{\omega \in W} \sum_{p \in P_\omega} G_p(x^\dagger)(x_p - x_p^\dagger) - \sum_{\omega \in W} \sum_{p \in P_\omega} u_\omega(x^\dagger)(x_p - x_p^\dagger) = \\ &= G(x^\dagger)(x - x^\dagger) - \sum_{\omega \in W} u_\omega(x^\dagger) \left(\sum_{p \in P_\omega} x_p - \sum_{p \in P_\omega} x_p^\dagger \right) = \\ &= G(x^\dagger)(x - x^\dagger) - \sum_{\omega \in W} u_\omega(x^\dagger)(\rho_\omega(x^\dagger) - \rho_\omega(x^\dagger)) = G(x^\dagger)(x - x^\dagger), \end{aligned}$$

то есть x^\dagger — решение квазивариационного неравенства (5). \square

Для задач транспортного равновесия с фиксированным спросом квазивариационное неравенство заменяется на классическое вариационное неравенство вида:

$$G(x^\dagger)(x - x^\dagger) \geq 0 \quad \forall x \in X. \quad (6)$$

Очевидно, что вариационное неравенство (6) представляет частный случай квазивариационного неравенства (5). Теория и методы решения обоих классов неравенств к настоящему времени уже достаточно хорошо разработаны. Познакомиться с достижениями в этой области можно, например, по работам [3, 5, 41, 49–51, 54].

Далее, для того чтобы изучать свойства задач транспортного равновесия в рамках единого математического аппарата, примем дополнительные соглашения и сведем решение задачи транспортного равновесия с эластичным спросом к решению именно вариационного неравенства.

Предположим, что для каждого маршрута $p \in P$ транспортные затраты $G_p(x)$ строго положительны, а для всех пар $\omega \in W$ функция спроса $\rho_\omega(u_\omega)$ принимает только неотрицательные значения.

Объединим величины u_ω в вектор $u = (u_\omega: \omega \in W)$, функции $\rho_\omega(u_\omega)$ — в вектор $\rho(u) = (\rho_\omega(u_\omega): \omega \in W)$. Рассмотрим векторы

$$z = \begin{pmatrix} x \\ u \end{pmatrix}, \quad F(z) = \begin{pmatrix} G(x) - \Xi u \\ \Xi^T x - \rho(u) \end{pmatrix},$$

где $\Xi = (\xi_{p\omega}: p \in P, \omega \in W)$ — матрица инцидентности путей и пар источник-сток:

$$\xi_{p\omega} = \begin{cases} 1, & \text{если путь } p \text{ соединяет пару } \omega, \\ 0 & \text{в противном случае.} \end{cases}$$

Допустимое множество для вектора z представляет собой неотрицательный ортант $Z = \{z: z \geq 0\}$.

Утверждение 1. Вектор $z^\dagger = (x^\dagger, u^\dagger) \geq 0$ является решением вариационного неравенства

$$F(z^\dagger)(z - z^\dagger) \geq 0 \quad \forall z \in Z \quad (7)$$

тогда и только тогда, когда x^\dagger — решение квазивариационного неравенства (5).

Доказательство. Пусть $x^\dagger \in X(x^\dagger)$ — решение квазивариационного неравенства (5). Тогда для любых $x \geq 0$ и $u \geq 0$ выполнены условия

$$\begin{aligned} (G(x^\dagger) - \Xi u^\dagger)x^\dagger &= 0, & (G(x^\dagger) - \Xi u^\dagger)x &\geq 0, \\ (\Xi^T x^\dagger - \rho(u^\dagger))u^\dagger &= 0, & (\Xi^T x^\dagger - \rho(u^\dagger))u &= 0. \end{aligned}$$

Откуда следует, что

$$0 \leq (G(x^\dagger) - \Xi u^\dagger)(x - x^\dagger) + (\Xi^T x^\dagger - \rho(u^\dagger))(u - u^\dagger) = F(z^\dagger)(z - z^\dagger).$$

Покажем обратное. Пусть $z^\dagger = (x^\dagger, u^\dagger) \geq 0$ — решение вариационного неравенства (7), то есть для любых $z \geq 0$ выполнено

$$F(z^\dagger)z^\dagger \leq F(z^\dagger)z.$$

Рассмотрим точки $z^\sigma = \sigma z^\dagger \geq 0$ для всех $\sigma \geq 0$. Имеем

$$F(z^\dagger)z^\dagger \begin{cases} \leq 0 & \text{при } \sigma = 0, \\ \geq \frac{F(z^\dagger)z^\dagger}{\sigma} \rightarrow 0 & \text{при } \sigma \rightarrow +\infty. \end{cases}$$

Следовательно, $F(z^\dagger)z^\dagger = 0$ и $F(z^\dagger)z \geq 0$.

Если предположить существование такого индекса l , что соответствующий ему элемент вектора $F(z^\dagger)$ отрицательный, $F_l(z^\dagger) < 0$, то, выбирая $z_l \rightarrow +\infty$, получаем нарушение неравенства $F(z^\dagger)z \geq 0$. Отсюда $F(z^\dagger) \geq 0$. Таким образом, точка $z^\dagger = (x^\dagger, u^\dagger) \geq 0$ удовлетворяет системе

$$F(z^\dagger) \geq 0, \quad z^\dagger \geq 0, \quad F(z^\dagger)z^\dagger = 0, \quad (8)$$

известной в литературе как *нелинейная задача дополненности* (см., например, [23, 41]).

Перепишем условия (8) в виде

$$G(x^\dagger) - \Xi u^\dagger \geq 0, \quad x^\dagger \geq 0, \quad (G(x^\dagger) - \Xi u^\dagger)x^\dagger = 0, \quad (9)$$

$$\Xi^T x^\dagger - \rho(u^\dagger) \geq 0, \quad u^\dagger \geq 0, \quad (\Xi^T x^\dagger - \rho(u^\dagger))u^\dagger = 0. \quad (10)$$

Система (9) показывает, что вектор u^\dagger соответствует минимальным транспортным затратам в сети при загрузке, определяемой потоками x^\dagger . При условии положительности транспортных затрат из (10) следует, что $\Xi^T x^\dagger - \rho(u^\dagger) = 0$, тогда неравенство (7) можно переписать в виде

$$G(x^\dagger)(x - x^\dagger) \geq u^\dagger(\Xi^T x - \rho(u^\dagger)) = 0 \quad \forall x \in X(u^\dagger),$$

что окончательно доказывает утверждение. \square

Очевидно, что допустимая область Z вариационного неравенства (7) является полиэдральным множеством, однако в отличие от допустимой области X вариационного неравенства (6) множество Z неограничено. Это критическим образом влияет на условия разрешимости транспортных задач с эластичным и фиксированным спросом.

1.1.4. Разрешимость задач транспортного равновесия

Критерии существования равновесных транспортных потоков формулируются на базе теории разрешимости вариационных неравенств. Рассмотрим вариационное неравенство (6) как общую форму задачи, специально не оговаривая свойства вектор-функции G и допустимого множества X . Для изложения основных результатов данного раздела понадобятся следующие определения.

Определение 1. Проекцией точки $y \in \mathbb{R}^n$ на множество $X \subset \mathbb{R}^n$ называется точка $\pi_X(y) = \operatorname{argmin}\{\|y - x\|: x \in X\}$.

Определение 2. Точка x^* называется неподвижной точкой отображения $T: X \rightarrow X$, если $x^* = T(x^*)$.

Критерием проверки, является ли вектор p проекцией точки $y \in \mathbb{R}^n$ на множество X , служит выполнение условия

$$(p - y)(x - p) \geq 0 \quad \forall x \in X. \quad (11)$$

Решение вариационного неравенства (6) тесно связано с поиском неподвижных точек проективного отображения

$$H(x) = \pi_X(x - \lambda G(x)),$$

где $\lambda > 0$ — некоторое фиксированное число.

Утверждение 2. Множество решений $X^\dagger \subseteq X$ вариационного неравенства (6) совпадает с множеством неподвижных точек отображения $H(x)$, то есть $X^\dagger = \{x^\dagger \in X: x^\dagger = H(x^\dagger)\}$.

Доказательство. Пусть $x^\dagger \in X^\dagger$ и $\lambda > 0$, тогда выполнено неравенство

$$(x^\dagger - (x^\dagger - \lambda G(x^\dagger)))(x - x^\dagger) \geq 0 \quad \forall x \in X,$$

следовательно, в силу свойства (11) имеем

$$x^\dagger = \pi_X(x^\dagger - \lambda G(x^\dagger)) = H(x^\dagger).$$

Пусть $x^\dagger = H(x^\dagger)$, тогда в силу свойства (11) для любого $x \in X$ выполнено условие

$$0 \leq (x - x^\dagger)(x^\dagger - (x^\dagger - \lambda G(x^\dagger))) = G(x^\dagger)(x - x^\dagger),$$

то есть $x^\dagger \in X^\dagger$. \square

Теорема 2. Пусть вектор-функция G непрерывна, множество X непусто, выпукло и замкнуто. Если X ограничено, то вариационное неравенство (6) разрешимо.

Доказательство. Для выпуклого множества X отображение $H(x): X \rightarrow X$ является непрерывным и однозначным. Множество X по условию теоремы компактно, следовательно, по теореме Брауэра (см., например, [2, 14, 24]) у $H(x)$ существует неподвижная точка $x^\dagger = H(x^\dagger)$, которая в силу утверждения 2 одновременно является решением вариационного неравенства (6). \square

Из теоремы 2 следует, что если затраты на передвижение $G_p(x)$ являются непрерывными функциями от потоков $x \in X$, то транспортная задача с фиксированным спросом всегда разрешима.

В случаях неограниченного допустимого множества X вводят дополнительные предположения о свойствах задачи, например, ограниченность потенциального множества решений, коэрцитивность, сильную монотонность и прочие. Общая идея выявления таких свойств состоит в следующем. Выберем такой радиус $R > 0$, что пересечение замкнутого шара $B = \{x: \|x\| \leq R\}$ с выпуклым замкнутым множеством X непусто, положим $X_R = X \cap B \neq \emptyset$. По теореме 2 существует такая точка $x_R^\dagger \in X_R$, что

$$G(x_R^\dagger)(x - x_R^\dagger) \geq 0 \quad \forall x \in X_R. \quad (12)$$

Теорема 3. Пусть вектор-функция G непрерывна, множество X непусто, выпукло и замкнуто. Если существует такой радиус $R > 0$, что $X_R \neq \emptyset$ и решение $x_R^\dagger \in X_R$ вариационного неравенства (12) удовлетворяет условию $\|x_R^\dagger\| < R$, то вариационное неравенство (6) разрешимо.

Доказательство. Для произвольного $x \in X$ выберем такое $\lambda \in (0, 1]$, что точка $\bar{x} = x_R^\dagger + \lambda(x - x_R^\dagger) \in X_R$. Имеем

$$0 \leq G(x_R^\dagger)(\bar{x} - x_R^\dagger) = G(x_R^\dagger)(x_R^\dagger + \lambda(x - x_R^\dagger) - x_R^\dagger) = \lambda G(x_R^\dagger)(x - x_R^\dagger),$$

то есть x_R^\dagger одновременно является решением вариационного неравенства (6). \square

Из теоремы 3 можно получить ряд следствий (см., например, [23]).

Следствие 1. Пусть вектор-функция G непрерывна, множество X непусто, выпукло и замкнуто. Если вектор-функция $G(x)$ коэрцитивна относительно X , то есть для некоторого $\bar{x} \in X$ выполнено

$$\lim_{\|x\| \rightarrow \infty, x \in X} \frac{G(x)(x - \bar{x})}{\|x\|} \rightarrow \infty, \quad (13)$$

то вариационное неравенство (6) разрешимо.

Доказательство. Условие коэрцитивности (13) позволяет для каждого фиксированного $C > 0$ подобрать достаточно большое $R_C > 0$ такое, что

$$G(x)(x - \bar{x}) \geq C\|x\| \quad \forall x \in X, \quad \|x\| = R_C,$$

для какого-то $\bar{x} \in X_{R_C}$, не зависящего от C и R_C .

В силу теоремы 2 разрешимо вариационное неравенство

$$G(x_{R_C}^\dagger)(x - x_{R_C}^\dagger) \geq 0 \quad \forall x \in X_{R_C}.$$

Если $\|x_C^\dagger\| < R_C$, то по теореме 3 точка $x_{R_C}^\dagger$ является решением исходного вариационного неравенства (6).

Если $\|x_{R_C}^\dagger\| = R_C$, то получаем

$$G(x_{R_C}^\dagger)(\bar{x} - x_{R_C}^\dagger) \leq -C\|x_{R_C}^\dagger\| = -CR < 0,$$

что противоречит определению $x_{R_C}^\dagger$. \square

Таким образом, для разрешимости задачи транспортного равновесия с эластичным спросом нужны более сильные, чем непрерывность, предположения о свойствах вектор-функции $F(z)$.

Вопрос единственности равновесного распределения транспортных потоков решается за счет свойств строгой монотонности функции транспортных издержек.

Определение 3. Вектор-функция $G: X \rightarrow \mathbb{R}^n$ называется строго монотонной на X , если для любых таких $x, y \in X$, что $x \neq y$, выполнено неравенство $(G(x) - G(y))(x - y) > 0$.

Теорема 4. Если вектор-функция $G(x)$ строго монотонна, то вариационное неравенство (6) может иметь не более одного решения.

Доказательство. Предположим противное, а именно, что существуют два различных решения $x^1, x^2 \in X$, $x^1 \neq x^2$, задачи (6). Очевидно, при этом выполнены неравенства

$$G(x^1)(x^2 - x^1) \geq 0, \quad G(x^2)(x^1 - x^2) \geq 0,$$

складывая которые, получаем

$$(G(x^1) - G(x^2))(x^2 - x^1) \geq 0,$$

что противоречит свойству строгой монотонности $G(x)$. \square

С точки зрения задачи транспортного равновесия утверждение теоремы 4 означает, что если транспортные затраты возрастают с увеличением загрузки сети — а это весьма естественное предположение, — то существует единственное равновесное распределение транспортных потоков.

1.1.5. Симметричные задачи транспортного равновесия

Задачу транспортного равновесия будем называть *симметричной*, если для вектор-функции транспортных издержек $G(x)$ матрица Якоби $\nabla G(x) = \left(\frac{\partial G_p(x)}{\partial x_q} : p, q \in P \right)$ симметрична для всех $x \in X$.

Свойство симметричности матрицы $\nabla G(x)$ является достаточным условием для того, чтобы гарантировать существование дифференцируемой функции $f: X \rightarrow \mathbb{R}$ такой, что $\nabla f(x) = G(x)$ для всех $x \in X$. Вектор-функция G в таком случае называется *потенциальной*, а вариационное неравенство (6) можно переписать в виде

$$\nabla f(x^\dagger)(x - x^\dagger) \geq 0 \quad \forall x \in X. \quad (14)$$

Из теории оптимизации известно, что условие (14) представляет необходимый критерий оптимальности в задаче

$$f(x) \rightarrow \min, \quad x \in X, \quad (15)$$

где X — выпуклое замкнутое множество.

В самом деле, пусть $x^\dagger = \operatorname{argmin}\{f(x) : x \in X\}$. Рассмотрим точку $x_\lambda = x^\dagger + \lambda(x - x^\dagger) \in X$, где $\lambda \in (0, 1)$ достаточно мало. Имеет место следующая оценка:

$$0 \leq \frac{f(x_\lambda) - f(x^\dagger)}{\lambda} = \frac{f(x^\dagger) + \lambda \nabla f(x^\dagger)(x - x^\dagger) + o(\lambda) - f(x^\dagger)}{\lambda} = \nabla f(x^\dagger)(x - x^\dagger) + \frac{o(\lambda)}{\lambda}.$$

Переходя к пределу при $\lambda \rightarrow 0$, получаем (14).

Считается, что решить оптимизационную задачу намного проще, чем вариационное неравенство [56]. Теория методов оптимизации богата разнообразными алгоритмами. Кроме того, существует множество программных пакетов для решения этого класса задач, чего нельзя сказать о вариационных неравенствах. Однако основная трудность состоит в построении функции $f(x)$. Для потенциальных отображений такую функцию можно построить, проведя следующие рассуждения.

Рассмотрим кривую L , зафиксируем на ней точку x^0 и вычислим интеграл $G(x)$ вдоль этой кривой до некоторой точки $x \in L$.

Пусть кривая L задана параметрически: $L = \{x(t) : t \in [0, 1]\}$, где $x(t)$ — гладкая вектор-функция, при этом $x(0) = x^0$, $x(1) = x$. Имеем

$$\begin{aligned} \mathcal{I} &= \oint_{x^0}^x G(x(t)) d(x(t)) = \int_0^1 G(x(t)) x'_t(t) dt = \int_0^1 \nabla f(x(t)) x'_t(t) dt = \\ &= \int_0^1 df(x(t)) = f(x(t)) \Big|_0^1 = f(x(1)) - f(x(0)) = f(x) - f(x^0). \end{aligned}$$

Видим, что значение интеграла \mathcal{I} не зависит от параметрического задания кривой L . Рассмотрим простейший пример такого задания: $x(t) = x^0 + t(x - x^0)$, тогда при $G(x) = \nabla f(x)$ вариационное неравенство (6) эквивалентно следующей оптимизационной задаче:

$$f(x) = f(x^0) + \int_0^1 G(x^0 + t(x - x^0))(x - x^0) dt \rightarrow \min, \quad x \in X. \quad (16)$$

Таким образом, решение симметричной задачи транспортного равновесия эквивалентно решению некоторой оптимизационной задачи. Именно этому классу задач посвящена большая часть работ по исследованию транспортного равновесия [22, 28, 29, 35, 37, 45, 46].

1.2. Построение функций транспортных затрат

Сложность численного решения задачи транспортного равновесия во многом зависит от аналитического задания функций $G_p(x)$. Интуитивно вполне очевидно, что на транспортные затраты при проезде из источника в сток в первую очередь влияют издержки на дугах, составляющих маршрут следования. В литературе, посвященной изучению проблем моделирования транспортных потоков, рассматриваются разные формы такой зависимости.

Обозначим через y_e величину потока по дуге $e \in E$. Зная распределение потоков по путям, можно рассчитать загрузку каждой дуги по следующей

формуле:

$$y_e = \sum_{p \in P} \theta_{ep} x_p, \quad (17)$$

где

$$\theta_{ep} = \begin{cases} 1, & \text{если путь } p \text{ проходит через дугу } e; \\ 0 & \text{в противном случае.} \end{cases}$$

Определим $\Theta = (\theta_{ep}: e \in E, p \in P)$ — матрицу инцидентности дуг и путей, $y = (y_e: e \in E)$ — вектор, описывающий загрузку дуг сети Γ . В матричной форме взаимосвязь потоков по путям и дугам описывается уравнением $y = \Theta x$.

В ряде случаев рассматриваются транспортные задачи в терминах только потоковых переменных по дугам. Отметим, что в множестве X , определенном в (3), от потоковых переменных по путям x можно легко перейти к вектору y ; обратный переход неоднозначен.

Затраты на прохождение единицы потока по дуге e (удельные затраты) обозначим через τ_e . В общем случае значение τ_e зависит не только от величины потока y_e , но и от потоков по другим дугам сети. Характерным примером тому служат нерегулируемые перекрестки, где порядок движения определяется приоритетом дорог, регулируемые перекрестки с дополнительной стрелкой сигнала светофора — движение в так называемом режиме «просачивания» и т. п. Поэтому правильно предположить, что $\tau_e = \tau_e(y)$. Сформируем вектор-функцию $\tau(y) = (\tau_e(y): e \in E)$.

1.2.1. Аддитивные функции затрат

Самым распространенным и простым предположением о свойствах функций транспортных затрат является аддитивная зависимость $G(x)$ от $\tau(y)$, означающая, что транспортные затраты на прохождение каждого пути $p \in P$ складываются только из затрат на проезд по дугам, составляющим этот путь [38, 39, 54]:

$$G_p(x) = \sum_{e \in E} \theta_{ep} \tau_e(y). \quad (18)$$

В результате получаем, что вектор-функция $G(x)$ вариационного неравенства (6) имеет вид

$$G(x) = \Theta^T \tau(y), \quad y = \Theta x. \quad (19)$$

Рассмотрим частный случай, когда затраты на проезд по дуге $\tau_e(y)$ зависят только от объема идущего по ней потока y_e , то есть $\tau_e(y) \equiv \tau_e(y_e)$. В этом случае для любых $p, q \in P, p \neq q$, имеем

$$\frac{\partial G_p}{\partial x_q} = \sum_{e \in E} \theta_{ep} \frac{\partial \tau_e}{\partial y_e} \frac{\partial y_e}{\partial x_q} = \sum_{e \in E} \theta_{ep} \theta_{eq} \frac{\partial \tau_e}{\partial y_e} = \frac{\partial G_q}{\partial x_p}.$$

Следовательно, матрица Якоби $\nabla G(x)$ симметрична для любых $x \in X$, то есть вектор-функция $G(x)$ потенциальна и равновесные транспортные потоки можно определить как решение оптимизационной задачи (16). Учитывая соотношения (19), вид целевой функции $f(x)$ определяется как:

$$\begin{aligned} f(x) &= \int_0^1 \sum_{p \in P} G_p(x^0 + t(x - x^0))(x_p - x_p^0) dt = \\ &= \int_0^1 \sum_{p \in P} \left(\sum_{e \in E} \theta_{ep} \tau_e(y_e^0 + t(y_e - y_e^0)) \right) (x_p - x_p^0) dt = \\ &= \int_0^1 \sum_{p \in P} \sum_{e \in E} \theta_{ep} (x_p - x_p^0) \tau_e(y_e^0 + t(y_e - y_e^0)) dt = \\ &= \int_0^1 \sum_{e \in E} \tau_e(y_e^0 + t(y_e - y_e^0)) \sum_{p \in P} \theta_{ep} (x_p - x_p^0) dt = \\ &= \int_0^1 \sum_{e \in E} \tau_e(y_e^0 + t(y_e - y_e^0)) (y_e - y_e^0) dt = \\ &= \int_0^1 \sum_{e \in E} \tau_e(y_e^0 + t(y_e - y_e^0)) d(y_e^0 + t(y_e - y_e^0)) = \sum_{e \in E} \int_{y_e^0}^{y_e} \tau_e(z) dz. \end{aligned}$$

Таким образом, при $\tau_e(y) \equiv \tau_e(y_e)$ задача (16) переписывается в виде:

$$\sum_{e \in E} \int_0^{y_e} \tau_e(z) dz \rightarrow \min, \quad y = \Theta x, \quad x \in X. \quad (20)$$

От потоковых переменных по путям x в модели (20) можно избавиться, если ввести следующие обозначения. В общем потоке y_e по каждой дуге $e \in E$ отдельно выделим поток y_e^s , порождаемый источником $s \in S$ и идущий по e . Вектор $y^s = (y_e^s: e \in E)$ определяет загрузку дуг транспортной сети, порожденную источником s . Очевидно, должны быть выполнены условия:

$$y = \sum_{s \in S} y^s, \quad y^s \geq 0, \quad s \in S.$$

Через ρ_{sv} обозначим объем потока, который из источника $s \in S$ должен достичь вершины $v \in V$. При заданной матрице корреспонденций ($\rho_w: w \in W = S \times D$) величины ρ_{sv} определяются по правилу:

$$\rho_{sv} = \begin{cases} \rho_w, & \text{если } v \neq s, w = (s, v) \in W, \\ 0, & \text{если } v \neq s, w = (s, v) \notin W, \\ -\rho_w, & \text{если } v = s. \end{cases}$$

Балансовые ограничения при переходе к новым переменным y_e^s запишутся в виде

$$\sum_{e \in E_v^+} y_e^s - \sum_{e \in E_v^-} y_e^s = \rho_{sv}, \quad (s, v) \in S \times V, \quad (21)$$

где

$$E_v^+ = \{e \in E: \text{дуга } e \text{ входит в вершину } v\}, \\ E_v^- = \{e \in E: \text{дуга } e \text{ выходит из вершины } v\}.$$

Определим множество $Y_s = \{y^s \geq 0: \text{выполнены условия (21)}\}$. В результате симметричная задача транспортного равновесия (20) переписывается в виде:

$$\sum_{e \in E} \int_0^{y_e} \tau_e(z) dz \rightarrow \min, \quad y = \sum_{s \in S} y^s, \quad y^s \in Y_s. \quad (22)$$

Одной из широко используемых форм функции затрат $\tau_e(y)$ является так называемая BPR-функция (Bureau of Public Road), описывающая временные затраты на проезд следующим образом:

$$\tau_e(y) = \tau_e^0 \left(1 + \mu \left(\frac{y_e}{c_e} \right)^n \right),$$

где τ_e^0 — время проезда по свободной дуге e , c_e — пропускная способность дуги e , μ и n — некоторые положительные константы. При использовании BPR-функции задача транспортного равновесия сводится к оптимизационной задаче (20).

1.2.2. Неаддитивные функции затрат

В общем случае, построение функции затрат $\tau_e(y)$ является задачей, требующей отдельных исследований. Здесь окажутся полезными как натурные замеры потоков и соответствующих им задержек в реальных УДС (улично-дорожных сетях), так и результаты компьютерного моделирования, например, при помощи специальных программ для агентного моделирования, так активно развивающиеся в последние годы.

Существуют ситуации, когда предположение об аддитивности функций $G_p(x)$ не подходит для описания транспортных затрат. Стремление к более адекватному моделированию автомобильных потоков привело к новым формам аналитического описания затрат [32, 43, 52]. Неаддитивные транспортные затраты возникают, например, в случаях, когда при моделировании одновременно учитываются и временные, и финансовые расходы. Так, в работе [43] предложена функция, характеризующая финансовые затраты, на которые, в свою очередь, влияют временные затраты:

$$G_p(x) = \Phi_p \left(\sum_{e \in E} \theta_{ep} \tau_e(y) \right) + \Psi_p(x) + \eta \sum_{e \in E} \theta_{ep} \tau_e(y),$$

где $\tau_e(\cdot)$ — время, потраченное на прохождение дуги e , $\Phi_p(\cdot)$ — функция, преобразующая временные затраты для пути p в финансовые затраты, $\Psi_p(\cdot)$ — финансовые затраты, характеризующие маршрут p , которые могут меняться в зависимости от загрузки сети, $\eta > 0$ — эксплуатационные расходы в единицу времени. В работе [32] предложен более общий вид неаддитивной функции затрат:

$$G_p(x) = U_\omega \left(\sum_{e \in E} \theta_{ep} \tau_e(y) + g_p(\Psi_p) \right), \quad p \in P_\omega,$$

где Ψ_p — фиксированные финансовые затраты, характеризующие маршрут p , $g_p(\cdot)$ — функция, преобразующая финансовые затраты во временные, $U_\omega(\cdot)$ — функция потерь (отрицательной полезности) для пары $\omega \in W$.

С одной стороны, неаддитивные затраты более реалистично могут описать функционирование транспортной системы, с другой — вариационное неравенство (6) (а тем более (7)) при сложных функциях $G_p(x)$ весьма трудоемко для анализа и решения.

1.2.3. Модель стационарной динамики

С формальной точки зрения, если объемы потоков не ограничены сверху пропускной способностью транспортной сети, то монотонные функции затрат допускают сколь угодно большие значения потоков, что едва ли согласуется с реальностью и справедливо критикуется. Однако полученный опыт моделирования транспортных потоков в реальной УДС [18, 19, 25, 26] показывает, что именно монотонное возрастание затрат является сдерживающим фактором для получения нереально больших потоков на дугах.

Одна из попыток избавиться от представления τ_e от упомянутого недостатка описана в работе [57]. Здесь для каждой дуги $e \in E$ транспортной сети предлагается ввести два вполне измеримых показателя: пропускную способность c_e и время проезда по свободной дуге (минимальное время проезда) τ_e^0 . Далее естественным образом предполагается, что в транспортной сети поток по дуге не может превышать ее пропускную способность, а потраченное на проезд время не может быть меньше, чем минимальное. Согласно [57], ситуация потокового равновесия в транспортной сети теперь определяется как загрузка ее дуг $y^\dagger = (y_e^\dagger: e \in E)$ и временные затраты на дугах $\tau^\dagger = (\tau_e^\dagger: e \in E)$, которые удовлетворяют ограничениям:

$$0 \leq y_e^\dagger \leq c_e, \quad \tau_e^\dagger \geq \tau_e^0, \quad e \in E, \quad (23)$$

при этом выполнены условия пользовательской оптимальности (1):

$$\tau_e^\dagger \begin{cases} = \tau_e^0, & \text{если } y_e^\dagger < c_e, \\ \geq \tau_e^0, & \text{если } y_e^\dagger = c_e. \end{cases} \quad (24)$$

Условие (24) показывает, что временные затраты зависят от потока, но эта зависимость не является монотонной. В случае полной загрузки сети можно гарантировать лишь только то, что временные затраты на дугах будут не меньше минимального времени проезда. Пара $(y^\dagger, \tau^\dagger)$, удовлетворяющая условиям (23), (24), называется стационарным динамическим решением задачи транспортного равновесия.

На взгляд авторов, независимость времени проезда по дуге от загрузки вплоть до достижения предельного значения выглядит весьма идеализированным, поскольку реальная практика вождения показывает, что с увеличением числа автомобилей на дороге скорость движения все-таки уменьшается. Однако этот подход сопровождается весьма интересными теоретическими результатами, о которых вкратце стоит упомянуть.

Для каждой потокообразующей пары $w \in W$ длину кратчайшего (по времени) пути при временных затратах на дугах, определяемых вектором τ , задает вогнутая кусочно-линейная функция

$$u_w(\tau) = \min_{q \in P_w} \left\{ G_q = \sum_{e \in E} \theta_{eq} \tau_e \right\}.$$

В качестве стационарного динамического решения задачи транспортного равновесия авторы [57] предлагают брать такое решение $(y^\dagger, \tau^\dagger)$ негладкой оптимизационной задачи

$$\sum_{w \in W} d_w u_w(\tau) - \sum_{e \in E} y_e \tau_e \rightarrow \max, \quad \tau_e \geq \tau_e^0, \quad e \in E, \quad (25)$$

на котором бы выполнялись равенства $\eta_e^\dagger = c_e - y_e^\dagger$, где η_e^\dagger — оптимальные значения двойственных переменных задачи (25).

Интересным результатом является тот факт, что сложную во всех отношениях задачу (25) можно заменить на двойственную к ней, которая в свою очередь является задачей линейного программирования, интерпретируемой как задача минимизации издержек в многопродуктовой транспортной задаче:

$$\tau^0 y \rightarrow \min, \quad y = \sum_{s \in S} y^s \leq c, \quad y^s \in Y_s, \quad (26)$$

где $\tau^0 = (\tau_e^0: e \in E)$ — вектор минимальных временных затрат в сети, $c = (c_e: e \in E)$ — вектор пропускной способности сети, переменные y^s и множество Y_s определены в разделе 1.2.1.

Практика применимости задачи (26) к УДС Владивостока описана в работе [19].

1.3. Соотношение между системным оптимумом и конкурентным равновесием

Очевидно, что общие затраты при системной оптимизации не могут превышать общих затрат при пользовательской оптимизации. Поэтому разность между совокупными транспортными затратами, которые несут пользователи сети, перемещаясь согласно либо только первому, либо только второму поведенческим принципам Вардропы, можно рассматривать как цену анархии, и существуют примеры, когда эта цена составляет существенную долю от общих расходов.

На принципиальную разницу между конкурентным транспортным равновесием и системным оптимумом одним из первых обратил внимание А. Пигу [60]. Он рассмотрел простейшую транспортную сеть, состоящую из двух дуг, соединяющих два пункта, скажем, спальный район A и бизнес-зону B (см. рис. 2).

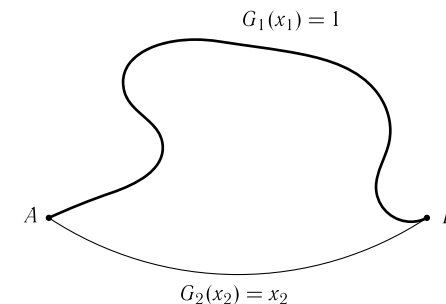


Рис. 2. Пример транспортной сети Пигу

Жители пункта A вольны выбирать, по какой из двух дорог им лучше добираться до работы. Обозначим через x_1 и x_2 доли общего объема трудового потока, едущего по первой и второй дорогам соответственно. Дороги в рассматриваемой сети неравноценны. Первая представляет магистральное шоссе, которое способно принять весь поток автомобилей из пункта A в пункт B без всякого замедления движения. Однако эта дорога достаточно длинная и проезд по ней требует определенного времени G_1 , которое будем считать равным, например, одному часу, то есть $G_1(x_1) = 1$. По второй дороге путь существенно короче, но это дорога узкая и движение сильно замедляется при наличии на ней потока автомобилей. Чтобы подчеркнуть суть примера, будем считать, что время проезда по второй дороге G_2 линейно зависит от потока x_2 по этой дороге и задается соотношением $G_2(x_2) = x_2$. Тогда в соответствии с первым принципом Вардропы (Пигу—Найта—Вардропы) равновесному состоянию будет со-

ответствовать такое распределение потоков $(x_1^\dagger, x_2^\dagger)$, что

$$G_1(x_1^\dagger) = G_2(x_2^\dagger), \quad x_1^\dagger + x_2^\dagger = 1, \quad x_1^\dagger, x_2^\dagger \geq 0,$$

откуда немедленно следует, что $x_1^\dagger = 0$, $x_2^\dagger = 1$, при этом системные затраты $c(x_1^\dagger, x_2^\dagger) = 1 \cdot x_1^\dagger + x_2^\dagger \cdot x_2^\dagger = 1$.

Распределение потоков в соответствии со вторым принципом Вардропы (системный оптимум) определяется как решение оптимизационной задачи:

$$x_1 + x_2^2 \rightarrow \min: \quad x_1 + x_2 = 1, \quad x_1, x_2 \geq 0, \quad (27)$$

минимум которой достигается в точке $x_1^* = x_2^* = 0,5$, минимальные затраты $c(x_1^*, x_2^*) = 0,75$, что на 25% уменьшает системные издержки в сети.

Приведенный пример Пигу показывает, что суммарные затраты в конкурентном равновесии могут составлять 4/3 от суммарных затрат системного оптимума. Оказывается, это соотношение представляет собой наилучшую оценку сверху для конкурентного потокового равновесия с аффинными функциями затрат и не зависит от топологии сети. Для подробного изложения этого результата установим некоторые полезные соотношения, характеризующие равновесные и оптимальные потоки.

Из условия равновесия (1), очевидно, следует, что если $x_p^\dagger > 0$ и $x_q^\dagger > 0$ для путей $p, q \in P_\omega$, то $G_p(x^\dagger) = G_q(x^\dagger) = u_\omega^\dagger$. Поскольку вклад в суммарные системные затраты (обозначим их $c(x^\dagger)$) при равновесном распределении x^\dagger вносят только ненулевые потоки $x_p^\dagger > 0$, а для них все удельные затраты в пределах одной пары ω одинаковы и равны u_ω^\dagger , то значение $c(x^\dagger)$ можно рассчитать следующим образом:

$$c(x^\dagger) = \sum_{p \in P} G_p(x^\dagger) x_p^\dagger = \sum_{\omega \in W} \sum_{p \in P_\omega} G_p(x^\dagger) x_p^\dagger = \sum_{\omega \in W} u_\omega^\dagger \sum_{p \in P_\omega} x_p^\dagger = \sum_{\omega \in W} u_\omega^\dagger \rho_\omega. \quad (28)$$

Распределение потоков по второму принципу Вардропы x^* и системный оптимум $c(x^*)$ соответствуют решению оптимизационной задачи:

$$c(x) = \sum_{p \in P} G_p(x) x_p \rightarrow \min, \quad x \in X. \quad (29)$$

Положим $c_p(x) = G_p(x) x_p$ и будем предполагать, что для всех $p \in P$ функции $c_p(x)$ являются выпуклыми и непрерывно дифференцируемыми. Справедлива следующая теорема.

Теорема 5. Пусть x^* — решения задачи (29), то есть оптимальное распределение потоков в сети. Тогда для всякой пары $\omega \in W$ верно следующее: если $x_p^* > 0$, $p \in P_\omega$, то $\frac{\partial c(x^*)}{\partial x_p} \leq \frac{\partial c(x^*)}{\partial x_q}$ для всех $q \in P_\omega$.

Доказательство. Предположим противное, а именно, что для пары ω существует такой путь $\bar{p} \in P_\omega$, что $x_{\bar{p}}^* > 0$ и $\frac{\partial c(x^*)}{\partial x_{\bar{p}}} > \frac{\partial c(x^*)}{\partial x_{\bar{q}}}$ для некоторого $\bar{q} \in P_\omega$, $\bar{q} \neq \bar{p}$. Рассмотрим такой вектор $x^\varepsilon = (x_p^\varepsilon: p \in P)$, что

$$x_p^\varepsilon = \begin{cases} x_p^*, & p \neq \bar{p}, p \neq \bar{q}, \\ x_{\bar{p}}^* - \varepsilon, & p = \bar{p}, \\ x_{\bar{q}}^* + \varepsilon, & p = \bar{q}, \end{cases}$$

где $\varepsilon > 0$ достаточно мало и не нарушает условия неотрицательности $x^\varepsilon \geq 0$. Нетрудно видеть, что $x^\varepsilon \in X$, при этом в силу выпуклости функции $c(x)$ имеем оценку

$$c(x^\varepsilon) - c(x^*) \leq \nabla c(x^*)(x^\varepsilon - x^*) = \varepsilon \left(\frac{\partial c(x^*)}{\partial x_{\bar{q}}} - \frac{\partial c(x^*)}{\partial x_{\bar{p}}} \right) < 0,$$

что противоречит оптимальности x^* . \square

В частном случае, когда $G_p(x) \equiv G_p(x_p)$, из теоремы 5 непосредственно следует, что в оптимальном распределении потоков x^* для всякой пары $\omega \in W$ верно следующее: если $x_p^* > 0$, $p \in P_\omega$, то $\frac{\partial c_p(x^*)}{\partial x_p} \leq \frac{\partial c_q(x^*)}{\partial x_q}$ для всех $q \in P_\omega$. При этом, как и в случае равновесных потоков, для оптимальных потоков справедливы равенства $\frac{\partial c_p(x^*)}{\partial x_p} = \frac{\partial c_q(x^*)}{\partial x_q}$ для тех $p, q \in P_\omega$, для которых $x_p^* > 0$, $x_q^* > 0$. Эти условия известны также как условия Гиббса (см., например, [11]).

Рассмотрим случай, когда транспортные затраты на прохождение каждого пути $p \in P$ складываются только из затрат на проезд по дугам, составляющим этот путь, то есть $G_p(x)$ определяются по формуле (18), при этом затраты по дугам $\tau_e(y)$ описываются аффинными функциями $\tau_e(y) = a_e y_e + b_e$, где a_e и b_e — неотрицательные коэффициенты для всех $e \in E$. При этом функция системных затрат и ее частные производные определяются как

$$c(x) = \sum_{p \in P} \sum_{e \in E} \theta_{ep} \tau_e(y) x_p = \sum_{e \in E} \tau_e(y) y_e = \sum_{e \in E} (a_e y_e^2 + b_e y_e),$$

$$\frac{\partial c(x)}{\partial x_p} = \sum_{e \in E} \theta_{ep} (2a_e y_e + b_e).$$

Обозначим через $y^\dagger = (y_e^\dagger: e \in E)$ и $y^* = (y_e^*: e \in E)$ загрузку дуг сети, порожденную потоками x^\dagger и x^* соответственно. Опираясь на приведенные

выше результаты для равновесных x^\dagger и оптимальных x^* потоков, можем утверждать, что выполнены следующие условия:

равновесие: если $x_p^\dagger > 0$, $p \in P_w$, то для любого $q \in P_w$ выполнено

$$\sum_{e \in E} \theta_{ep} (a_e y_e^\dagger + b_e) \leq \sum_{e \in E} \theta_{eq} (a_e y_e^\dagger + b_e); \quad (30)$$

оптимальность: если $x_p^* > 0$, $p \in P_w$, то для любого $q \in P_w$ выполнено

$$\sum_{e \in E} \theta_{ep} (2a_e y_e^* + b_e) \leq \sum_{e \in E} \theta_{eq} (2a_e y_e^* + b_e). \quad (31)$$

Любопытно, что при линейных функциях задержек $\tau_e(y) = a_e y_e$ из неравенств (30) и (31) следует совпадение равновесных и оптимальных потоков.

Следуя работе [61], через тройку $[\Gamma, \rho, G(x)]$ обозначим транспортную модель, определенную на сети Γ , с матрицей корреспонденций $\rho = (\rho_w: w \in W)$ и затратами $G(x) = (G_p(x): p \in P)$. Везде далее будем полагать

$$G_p(x) = \sum_{e \in E} \theta_{ep} \tau_e(y) = \sum_{e \in E} \theta_{ep} (a_e y_e + b_e). \quad (32)$$

Имеет место следующий результат.

Лемма 1. Пусть x^\dagger — решение задачи транспортного равновесия для модели $[\Gamma, \rho, G(x)]$. Тогда вектор $\frac{1}{2}x^\dagger$ является решением оптимизационной задачи для модели $[\Gamma, \frac{1}{2}\rho, G(x)]$.

Доказательство. Если x^\dagger является допустимым решением равновесной модели $[\Gamma, \rho, G(x)]$, то, очевидно, $\frac{1}{2}x^\dagger$ — допустимое решение оптимизационной модели $[\Gamma, \frac{1}{2}\rho, G(x)]$, при этом неравенства (31) для $\frac{1}{2}x^\dagger$ переходят в (30). \square

Более того, для каждого такого маршрута $p \in P_w$, что $x_p^\dagger > 0$, выполнено

$$\frac{\partial c\left(\frac{1}{2}x^\dagger\right)}{\partial x_p} = \sum_{e \in E} \theta_{ep} \{a_e y_e^\dagger + b_e\} = u_w(x^\dagger).$$

Лемма 2. Пусть x^* — оптимальное распределение потоков, отвечающее транспортной модели $[\Gamma, \rho, G(x)]$. Тогда для любого допустимого потока x^δ в модели $[\Gamma, (1 + \delta)\rho, G(x)]$ справедлива оценка

$$c(x^\delta) \geq c(x^*) + \delta \sum_{w \in W} v_w(x^*) \rho_w, \quad (33)$$

$$\text{где } \delta \geq 0, v_w(x^*) = \min_{p \in P_w} \frac{\partial c(x^*)}{\partial x_p}.$$

Доказательство. Рассмотрим допустимые относительно модели $[\Gamma, (1 + \delta)\rho, G(x)]$ потоки x^δ . При затратах $G_p(x)$, определенных в (32), где все коэффициенты $a_e \geq 0$, функция $c(x)$ выпукла, отсюда

$$\begin{aligned} c(x^\delta) &\geq c(x^*) + \frac{\partial c(x^*)}{\partial x_p} (x^\delta - x^*) = c(x^*) + \sum_{p \in P} \frac{\partial c(x^*)}{\partial x_p} (x_p^\delta - x_p^*) = \\ &= c(x^*) + \sum_{w \in W} \sum_{p \in P_w} \frac{\partial c(x^*)}{\partial x_p} (x_p^\delta - x_p^*) = \\ &= c(x^*) + \sum_{w \in W} \left(\sum_{p \in P_w} \frac{\partial c(x^*)}{\partial x_p} x_p^\delta - \sum_{p \in P_w} \frac{\partial c(x^*)}{\partial x_p} x_p^* \right). \end{aligned}$$

Поскольку для таких p , что $x_p^* > 0$, производная $\frac{\partial c(x^*)}{\partial x_p}$ принимает минимальное значение:

$$\frac{\partial c(x^*)}{\partial x_p} = \min_{q \in P_w} \frac{\partial c(x^*)}{\partial x_q} = v_w(x^*),$$

то

$$\sum_{p \in P_w} \frac{\partial c(x^*)}{\partial x_p} x_p^* = \sum_{p \in P_w} v_w(x^*) x_p^*.$$

Следовательно, продолжая оценку снизу для $c(x^\delta)$, получаем

$$\begin{aligned} c(x^\delta) &\geq c(x^*) + \sum_{w \in W} \left(\sum_{p \in P_w} v_w(x^*) x_p^\delta - \sum_{p \in P_w} v_w(x^*) x_p^* \right) = \\ &= c(x^*) + \sum_{w \in W} v_w(x^*) \left(\sum_{p \in P_w} x_p^\delta - \sum_{p \in P_w} x_p^* \right) = \\ &= c(x^*) + \sum_{w \in W} v_w(x^*) ((1 + \delta)\rho_w - \rho_w) = c(x^*) + \delta \sum_{w \in W} v_w(x^*) \rho_w. \quad \square \end{aligned}$$

Итоговый результат текущего раздела устанавливает следующая теорема.

Теорема 6. Для транспортной модели $[\Gamma, \rho, G(x)]$ с аффинными функциями задержек (32) для оптимального x^* и равновесного x^\dagger распределений потоков выполняется соотношение

$$\frac{c(x^\dagger)}{c(x^*)} \leq \frac{4}{3}.$$

Доказательство. Согласно (28) системные затраты для равновесного распределения x^\dagger рассчитываются как

$$c(x^\dagger) = \sum_{w \in W} u_w(x^\dagger) \rho_w,$$

по лемме 1 поток $\frac{1}{2}x^\dagger$ оптимален для транспортной модели $[\Gamma, \frac{1}{2}\rho, G(x)]$, при этом $v_w(\frac{1}{2}x^\dagger) = u_w(x^\dagger)$.

Положим $\delta = 1$ в оценке (33). Тогда для произвольного потока x , допустимого в модели $[\Gamma, 2 \cdot \frac{1}{2}\rho, G(x)] = [\Gamma, \rho, G(x)]$, имеем

$$\begin{aligned} c(x) &\geq c\left(\frac{1}{2}x^\dagger\right) + \sum_{w \in W} \frac{1}{2}v_w\left(\frac{1}{2}x^\dagger\right)\rho_w = \\ &= c\left(\frac{1}{2}x^\dagger\right) + \frac{1}{2} \sum_{w \in W} u_w(x^\dagger)\rho_w = c\left(\frac{1}{2}x^\dagger\right) + \frac{1}{2}c(x^\dagger). \end{aligned} \quad (34)$$

Осталось получить оценку снизу $c\left(\frac{1}{2}x^\dagger\right)$ в терминах $c(x^\dagger)$, что легко сделать, учитывая вид функций задержки:

$$c\left(\frac{1}{2}x^\dagger\right) = \sum_{e \in E} \frac{1}{2}y_e^\dagger \left(\frac{1}{2}a_e y_e^\dagger + b_e\right) \geq \frac{1}{4} \sum_{e \in E} y_e^\dagger (a_e y_e^\dagger + b_e) = \frac{1}{4}c(x^\dagger),$$

где для промежуточных вычислений использовались потоки по дугам $(y_e^\dagger, e \in E)$, индуцированные равновесными потоками по маршрутам x^\dagger . Очевидно, что при этом потоки $\frac{1}{2}x^\dagger$ будут индуцировать загрузку дуг $\left(\frac{1}{2}y_e^\dagger, e \in E\right)$.

В результате, продолжая оценку (34), получим

$$c(x) \geq \frac{1}{4}c(x^\dagger) + \frac{1}{2}c(x^\dagger) = \frac{3}{4}c(x^\dagger).$$

Вычисляя в последнем неравенстве минимум левой части по всем x , допустимым в модели $[\Gamma, \rho, G(x)]$, получаем

$$\frac{c(x^\dagger)}{c(x^*)} \leq \frac{4}{3}. \quad \square$$

Относительно общего случая нелинейных функций транспортных затрат нетрудно убедиться в том, что цена анархии может быть сколь угодно большой.

1.4. Численные методы решения задач транспортного равновесия

Эквивалентность задачи транспортного равновесия вариационному неравенству, а в частном случае оптимизационной задаче, позволяет адаптировать численные методы решения последних для поиска равновесных потоков. В данном разделе рассматриваются подходы к решению задачи транспортного равновесия с фиксированным спросом.

В зависимости от того, в каком пространстве переменных рассматривается исходная задача, выделяют два основных подхода к построению алгоритмических схем. Если равновесие моделируется только через потоковые переменные по дугам y_e , то применяют так называемые *дуговые алгоритмы* (arc-based algorithms). Если основными переменными задачи являются потоки по путям x_p и, соответственно, итерирование ведется по допустимым маршрутам, то такие алгоритмы называются *маршрутными* (path-based algorithms) [34, 59].

Основным преимуществом поиска равновесия через переменные x_p является возможность «убить двух зайцев одним выстрелом»: зная распределение потоков по маршрутам и используя соотношение (17), всегда можно определить загрузку дуг транспортной сети. Обратное преобразование, очевидно, неоднозначно. Информация о распределении потоков по путям сама по себе является важной при моделировании других задач, например, проблем загрязнения окружающей среды, оценки матрицы корреспонденций, планирования транспортных развязок и модернизации улично-дорожной сети, эффективного регулирования движения и т. п.

Несомненным плюсом в пользу работы в пространстве потоковых переменных по путям с алгоритмической точки зрения является возможность естественной проверки выполнения условия равновесия (1) и поиска распределения потоков, удовлетворяющих заданной точности. Информация только о потоках по дугам такой возможности не дает. Кроме того, структура допустимого множества X , определенного в (3), представляет собой декартово произведение непересекающихся симплексов X_w , и такое свойство может породить целый класс методов, использующих принципы декомпозиции и идеи параллелизации итерационных схем.

Последним аргументом в пользу исследования задачи транспортного равновесия именно в терминах потоковых переменных по путям является тот простой факт, что при общем задании функции издержек $G_p(x)$, не обязательно складывающихся из затрат на передвижение по дугам, переформулировка условия равновесия (1) в терминах переменных y_e невозможна.

Основной недостаток работы с потоковыми переменными по путям — это необходимость априорного задания множества всех допустимых маршрутов.

рутов P . Такая задача является очень трудоемкой, особенно для реальных транспортных сетей. Как вариант, для каждой потокообразующей пары можно ограничиться рассмотрением k кратчайших маршрутов, заведомо исключить неперспективные пути, но от этого проблема проще не становится. На практике используется не так много вариантов движения, поэтому нет необходимости знать все элементы множества P . Более того, существует стандартная техника, часто называемая методом генерации столбцов, когда входные данные непосредственно строятся в процессе решения задачи. Применение такой техники к проблеме поиска транспортного равновесия позволит строить множество допустимых и перспективных для использования маршрутов непосредственно в процессе решения задачи.

Таким образом, из приведенных аргументов видно, что потоковое равновесие предпочтительней искать как решение вариационного неравенства (6) в пространстве потоковых переменных по путям.

1.4.1. Проективные методы решения задачи транспортного равновесия

Среди существующих методов решения вариационных неравенств отдельно можно выделить проективные алгоритмы, отличающиеся простотой своих итерационных схем и гибкостью к различного рода модификациям.

В основу проективных методов положена связь между множеством решений вариационного неравенства и неподвижными точками проективного отображения, установленная в утверждении 2.

Далее понадобятся следующие определения.

Определение 4. Отображение $G: X \rightarrow \mathbb{R}^n$ на множестве X называется:

- липшицевым, если существует константа $L > 0$ такая, что $\|G(x) - G(y)\| \leq L\|x - y\|$ для всех $x, y \in X$;
- сильно монотонным с константой $\tau > 0$, если $(G(x) - G(y))(x - y) \geq \tau\|x - y\|^2$ для всех $x, y \in X$;
- обратно сильно монотонным (ко-коэрцитивным) с константой $\tau > 0$, если $(G(x) - G(y))(x - y) \geq \tau\|G(x) - G(y)\|^2$ для всех $x, y \in X$;
- монотонным, если $(G(x) - G(y))(x - y) \geq 0$ для всех $x, y \in X$;
- псевдомонотонным, если из неравенства $G(x)(y - x) \geq 0$ следует $G(y)(y - x) \geq 0$ для всех $x, y \in X$;
- строго псевдомонотонным, если из неравенства $G(x)(y - x) \geq 0$ следует $G(y)(y - x) > 0$ для всех $x, y \in X, x \neq y$.

В самой простой форме проективный метод строит последовательность $\{x^k\} \in X$, генерируемую рекуррентным соотношением

$$x^{k+1} = \pi_X(x^k - \lambda_k G(x^k)), \quad \lambda_k > 0, \quad k = 0, 1, 2, \dots \quad (35)$$

На текущий момент известно, что сходимость процесса (35) гарантируется при выполнении одного из следующих условий (см. [50] и библиографию в ней):

- 1) отображение G сильно монотонно с константой τ и липшицево с константой $L, \lambda_k \in (0, 2\tau/L^2)$;
- 2) отображение G ко-коэрцитивно с константой $\mu, \lambda_k \in (0, 2\mu)$;
- 3) для любых $x \in X \setminus X^*$ и $x^* \in X^*$ выполнено

$$G(x)(x - x^*) > 0, \quad (36)$$

где X^* — множество решений вариационного неравенства, $\lambda_k = \frac{\alpha_k}{\|G(x^k)\|}, \sum_{k=0}^{\infty} \alpha_k = \infty, \sum_{k=0}^{\infty} \alpha_k^2 < \infty$.

Нетрудно показать, что любое сильно монотонное липшицево отображение является ко-коэрцитивным и влечет выполнение неравенства (36). Обратное, очевидно, неверно, поэтому второе и третье условия являются менее ограничительными, чем первое, однако остаются достаточно сильными предположениями, что существенно сужает круг задач, для которых метод (35) гарантированно сходится. Кроме того, правила выбора шагового множителя λ_k являются весьма общими для всех трех случаев, что на практике приводит к медленной скорости сходимости всего процесса.

Одной из попыток улучшить ситуацию было предложение о замене процесса (35) на так называемый экстраградиентный метод [1, 13]:

$$u^k = \pi_X(x^k - \lambda_k G(x^k)), \quad x^{k+1} = \pi_X(x^k - \lambda_k G(u^k)), \quad \lambda_k > 0, \quad k = 0, 1, 2, \dots \quad (37)$$

Первую проекцию u^k можно трактовать как предиктор, вторую x^{k+1} — как корректор.

Сходимость экстраградиентного метода гарантируется при выполнении одного из следующих условий [1, 13, 31, 53]:

- 1) G монотонно и липшицево с константой $L, \lambda_k \in (0, 1/L)$;
- 2) G псевдомонотонно, локально липшицево, динамическая регуловка шага $\lambda_k \in (0, \min\{\bar{\lambda}, \gamma\|x^k - u^k\|/\|G(x^k) - G(u^k)\|\})$, $\gamma \in (0, 1)$.

Для ускорения сходимости экстраградиентного метода величину шага λ_k у предиктора и корректора можно сделать разной, при этом отображение G может удовлетворять всего лишь свойству псевдомонотонности [44, 62, 65]. Хороший обзор, посвященный исследованиям методов проективного типа, приведен в [66].

Нетрудно видеть, что условия сходимости экстраградиентного метода менее ограничительные по сравнению с проективным. Однако плата за

такое послабление — это многократное решение задачи проекции на этапе определения шагового множителя.

Основной трудностью при реализации процессов (35) и (37) является вычисление проекции на множество X . В общем случае требуется решение вспомогательной оптимизационной задачи, однако при поиске транспортного равновесия как решения вариационного неравенства (6) операцию проектирования $\pi_X(\cdot)$ можно свести к более простым вычислениям $\pi_{X_\omega}(\cdot)$.

1.4.2. Декомпозиция проективных методов для поиска равновесных потоков

Как уже было отмечено, множества X_ω , определенные в (4), не пересекаются по переменным, поэтому вектор $x \in X$ можно разделить на подвектора $x^\omega = (x_p: p \in P_\omega) \in X_\omega$ — потоки по путям, соединяющим пару $\omega \in W$. Аналогично выделим вектор-функции $G^\omega(x) = (G_p(x): p \in P_\omega)$ — издержки по путям, соединяющим пару $\omega \in W$, при загрузке сети потоками x . Таким образом $x = (x^\omega: \omega \in W)$ и $G(x) = (G^\omega(x): \omega \in W)$.

Отмеченная специфика множества X , определенного в (3), естественным образом позволяет свести операцию $\pi_X(\cdot)$ к вычислениям $\pi_{X_\omega}(\cdot)$ для каждого $\omega \in W$. Поэтому для поиска транспортного равновесия общая схема проективного метода (35) трансформируется в следующий процесс:

$$\begin{aligned} x^k &= ((x^\omega)^k: \omega \in W), & (x^\omega)^{k+1} &= \pi_{X_\omega}((x^\omega)^k - \lambda_k G^\omega(x^k)), \\ x^{k+1} &= ((x^\omega)^{k+1}: \omega \in W), & \lambda_k &> 0, \quad k = 0, 1, 2, \dots \end{aligned} \quad (38)$$

Аналогично можно преобразовать и процесс (37).

Для вычисления $\pi_{X_\omega}(\cdot)$ существует замечательный алгоритм поиска проекции точки $z \in \mathbb{R}^n$ на симплекс $\Delta = \{x \in \mathbb{R}_+^n: ex = \rho\}$, где e — единичный вектор размерности n . Общая схема алгоритма следующая [47]:

ШАГ 0. Вычислить $x^0 = z + ((\rho - ez)/n)e$, положить $k = 0$.

ШАГ 1. Если $x^k \geq 0$, то проекция z на множество Δ найдена, алгоритм заканчивает работу. В противном случае выполняется шаг 2.

ШАГ 2. Вычислить x^{k+1} по правилу:

$$x_i^{k+1} = \begin{cases} 0, & \text{если } i \notin I, \\ x_i^k + \frac{1}{|I|} \left(\rho - \sum_{j \in I} x_j^k \right), & \text{если } i \in I. \end{cases}$$

где $I = \{i: x_i^k > 0\}$, положить $k = k + 1$ и перейти на шаг 1.

Приведенный алгоритм за не более чем n шагов приведет к искомой точке $\pi_\Delta(z)$.

Отметим, что подвекторы $(x^\omega)^k$ определяются независимо друг от друга, поэтому при программной реализации процесс (38) легко поддается

параллелизации, что несомненно улучшает его вычислительные свойства и делает проективные методы привлекательными для использования.

1.4.3. Проективный метод с генерацией маршрутов

Специфика допустимого множества X задачи транспортного равновесия позволяет без труда вычислять операцию проектирования. Однако проективные алгоритмы предполагают априорное задание полного множества допустимых маршрутов P , чего хотелось бы избежать по следующим трем причинам. Во-первых, трудоемкость алгоритмов построения всех путей, соединяющих пару вершин, растет экспоненциально с увеличением размерности графа. Поэтому для реальных сетей большой размерности, в которых рассматриваются сотни или тысячи потокообразующих пар, задача построения полного множества маршрутов может потребовать больших вычислительных ресурсов, превосходящих даже возможности суперкомпьютеров. Во-вторых, мощность множества P определяет размерность решаемой задачи, поэтому чем больше элементов в P , тем труднее с вычислительной точки зрения будет проходить поиск равновесного распределения потоков в сети. И наконец, в-третьих, скорее всего, большая часть путей из множества P не будет использоваться при переносе заданного трафика, поэтому их включение в P бессмысленно.

В [46] для решения симметричной задачи транспортного равновесия предложен метод, который наряду с решением оптимизационной задачи (20) последовательно строит множество допустимых маршрутов в сети. Такой подход можно обобщить и для несимметричного случая.

Обозначим через $x^\dagger(\mathcal{P})$ равновесное распределение потоков на множестве путей \mathcal{P} , через $\mathbf{0}_{|\mathcal{P}|}$ — нулевой вектор размерности, равной количеству элементов в множестве \mathcal{P} , через $VI(\mathcal{P})$ — вариационное неравенство вида:

$$\sum_{p \in \mathcal{P}} G_p(x^\dagger(\mathcal{P}))(x_p(\mathcal{P}) - x_p^\dagger(\mathcal{P})) \geq 0,$$

$$x(\mathcal{P}) = (x_p: p \in \mathcal{P}) \in X(\mathcal{P}) = \prod_{\omega \in W} X_\omega(\mathcal{P}_\omega),$$

где $X_\omega(\mathcal{P}_\omega) = \left\{ x_p \geq 0: p \in \mathcal{P}_\omega, \sum_{p \in \mathcal{P}_\omega} x_p = \rho_\omega \right\}$ — допустимое множество потоков для пары ω на множестве путей \mathcal{P}_ω и $\mathcal{P} = \bigcup_{\omega \in W} \mathcal{P}_\omega$. В силу теоремы 1

вектор $x^\dagger(\mathcal{P}) \in X(\mathcal{P})$ является решением вариационного неравенства $VI(\mathcal{P})$.

Пусть для каждой пары $\omega \in W$ задано некоторое непустое подмножество допустимых маршрутов $P_\omega^0 \subseteq P_\omega$. Тогда текущее множество путей в сети определяется как $P^0 = \bigcup_{\omega \in W} P_\omega^0$.

Предположим, что для P^0 в процессе решения вариационного неравенства $VI(P^0)$ найдены равновесные потоки $x^\dagger(P^0)$. При этом по всем путям $p \in P \setminus P^0 = \bar{P}^0$ движения нет, поэтому $x(\bar{P}^0) = \mathbf{0}_{|\bar{P}^0|}$. Вектор $x^0 = (x^\dagger(P^0), \mathbf{0}_{|\bar{P}^0|})$ содержится в X , т. е. является допустимым распределением потоков в сети. Осталось проверить, является ли x^0 равновесным.

Потоки x^0 характеризуют загрузку сети и приводят к определенным задержкам на ее дугах. Если для некоторой пары $\omega \in W$ при текущем состоянии сети найден маршрут $q_\omega \notin P_\omega^0$ такой, что

$$G_{q_\omega}(x^0) < u_\omega(x^0), \quad (39)$$

то, очевидно, распределение x^0 не является равновесным и необходимо перераспределить потоки уже с учетом q_ω . Если ни для одной из потокообразующих пар такого пути не существует, то потоки x^0 удовлетворяют условиям равновесия (1).

Обнаружить наименее затратный путь q_ω можно при помощи специальных алгоритмов. Возможно, что такие пути существуют для нескольких потокообразующих пар, поэтому при перераспределении потоков предпочтительно учитывать все такие пути. Новую загрузку сети можно получить в результате решения вариационного неравенства $VI(P^1)$, где $P^1 = P^0 \cup_{\omega \in W} q_\omega$. Снова получаем допустимое распределение потоков $x^1 = (x^\dagger(P^1), \mathbf{0}_{|\bar{P}^1|}) \in X$, которое необходимо проверить на соответствие условиям равновесия (1). Процесс будет повторяться до тех пор, пока существуют пути q_ω , удовлетворяющие (39).

Приведенные выше рассуждения и декомпозированный проективный метод (38) приводят к следующему алгоритму поиска потокового равновесия в транспортных сетях.

Алгоритм 1

Шаг 0. Сформировать множество $P^0 = \bigcup_{\omega \in W} P_\omega^0$, где $P_\omega^0 \subseteq P_\omega$ и $P_\omega^0 \neq \emptyset$ для всех $\omega \in W$. Положить $k = 0$.

Шаг 1. Найти решение $x^\dagger(P^k) \in X(P^k)$ вариационного неравенства $VI(P^k)$ проективным методом (38).

Шаг 2. Построить множество $Q^k = \{q_\omega: \omega \in W, G_{q_\omega}(x^k) < u_\omega(x^k)\}$. Если $Q^k = \emptyset$, то текущее распределение потоков x^k является равновесным, алгоритм заканчивает работу. В противном случае сформировать множество $P^{k+1} = P^k \cup Q^k$, положить $k = k + 1$ и повторить шаг 1.

Так как $|P^{k+1}| > |P^k|$ и множество всех путей в графе $\Gamma(V, E)$ конечно, то приведенный алгоритм 1 за конечное число шагов сойдется к равновесному распределению потоков $x^\dagger \in X$.

Заметим, что с каждой следующей итерацией алгоритма 1 размерность решаемого вариационного неравенства $VI(P^k)$ увеличивается, что

может весьма усложнить численные расчеты. Для преодоления «проклятия размерности» предлагается следующая модификация алгоритма 1. При $Q^k \neq \emptyset$ на шаге 2 множество P^{k+1} строится по правилу

$$P^{k+1} = [P^k]_+ \cup Q^k, \quad [P^k]_+ = \{p \in P^k: x_p^\dagger > 0\}, \quad (40)$$

т. е. к началу следующей итерации допустимое множество маршрутов P^{k+1} формируется только из путей множества P^k , которые участвуют в переносе трафика, и найденных кратчайших путей, не учтенных в P^k .

Несмотря на то что применение правила (40) исключает последовательное накопление допустимых маршрутов и есть вероятность появления в P^k ранее рассмотренного пути, алгоритм 1 остается конечным, что гарантирует следующий результат [26].

Рассмотрим оценочную функцию $\varphi(x) = \max_{\eta \in X} G(x)(x - \eta) \geq 0$ для вариационного неравенства (6). Известно [41, 49, 50], что точка $x^\dagger \in X$ является решением $VI(P)$ тогда и только тогда, когда $x^\dagger = \operatorname{argmin}\{\varphi(x): x \in X\}$ и $\varphi(x^\dagger) = 0$.

Утверждение 3. Если $x^k \in X$ не является равновесной загрузкой сети и множество P^{k+1} формируется по правилу (40), то $\varphi(x^{k+1}) < \varphi(x^k)$.

Доказательство. Пусть для множества маршрутов P^k определено равновесное распределение потоков $x^*(P^k)$ и множество $Q^k \neq \emptyset$. Следуя алгоритму 1, строим новое множество P^{k+1} по правилу (40) и для него находим равновесное распределение потоков $x^*(P^{k+1})$.

Рассмотрим оценочную функцию

$$\varphi_{P^{k+1}}(x) = \max_{\eta \in X(P^{k+1})} G(x)(x - \eta),$$

определенную для вариационного неравенства $VI(P^{k+1})$. Следовательно,

$$x^*(P^{k+1}) = \operatorname{argmin}\{\varphi_{P^{k+1}}(x): x \in X(P^{k+1})\}, \quad \varphi_{P^{k+1}}(x^*(P^{k+1})) = 0.$$

Построим вектор $z = (z_p: p \in P^{k+1})$ по правилу: $z_p = x_p^*$ при $p \in [P^k]_+$ и $z_p = 0$ при $p \in Q^k$. Очевидно, что $z \in X(P^{k+1})$, поэтому имеет место соотношение

$$0 = \varphi_{P^{k+1}}(x^*(P^{k+1})) \leq \varphi_{P^{k+1}}(z).$$

Оценочную функцию $\varphi_{P^{k+1}}(z)$ можно переписать в виде

$$\begin{aligned} \varphi_{P^{k+1}}(z) &= \sum_{\omega \in W} \sum_{p \in P_\omega^{k+1}} z_p G_p(z) - \sum_{\omega \in W} G_{q_\omega}(x^*(P^k)) d_\omega = \\ &= \sum_{\omega \in W} \sum_{p \in P_\omega^{k+1}} (G_p(z) - G_{q_\omega}(x^*(P^k))) z_p > 0. \end{aligned}$$

Следовательно, $\varphi_{P^{k+1}}(x^*(P^{k+1})) < \varphi_{P^{k+1}}(z)$.

Рассмотрим два вектора

$$x^k = (x^*(P^k), \mathbf{0}_{|\bar{P}^k|}) \in X, \quad x^{k+1} = (x^*(P^{k+1}), \mathbf{0}_{|\bar{P}^{k+1}|}) \in X.$$

Имеет место оценка

$$\begin{aligned} \varphi(x^{k+1}) &= \sum_{p \in P^{k+1}} G_p(x^{k+1})x_p^{k+1} - \sum_{w \in W} u_w(x^{k+1})d_w < \\ &< \sum_{p \in [P^k]_+} G_p(x^k)x_p^k = \sum_{p \in P} G_p(x^k)x_p^k = \varphi(x^k), \end{aligned}$$

что доказывает сходимость алгоритма. \square

Из утверждения 3 следует, что если $k \neq l$, то $P^k \neq P^l$, поэтому модификация (40) не нарушает конечности алгоритма 1.

1.4.4. Ступенчатая регулировка шага проективного метода

Проективные методы относятся к классу так называемых фейеровских алгоритмов [8, 12], получивших широкое распространение при решении задач допустимости, оптимизации, дополненности, равновесия и др. Обобщение свойства фейеровости было предложено в [16], что позволило направить порождаемый фейеровским оператором процесс к некоторому заданному подмножеству данного множества.

С целью ускорения сходимости процесса (38) к равновесному распределению потоков предлагается использовать теорию фейеровских процессов с адаптивным шагом [17, 58]. Через $U(y, \varepsilon) = \{x: \|x - y\| < \varepsilon\}$ обозначим ε -окрестность точки y . Будем использовать следующие определения.

Определение 5. Отображение $R: X \rightarrow \mathbb{R}^n$ называется локально сильным аттрактантом подмножества $Z \subset X$, если для любой точки $x' \in X \setminus Z$ найдутся такие $\varepsilon > 0$ и $\gamma > 0$, что для любых $x \in U(x', \varepsilon)$ и $z \in Z$ выполнено $R(x)(z - x) \geq \gamma$.

Определение 6. Оператор F называется локально сильно фейеровским относительно множества X , если $F(x) = x$ для любых $x \in X$ и для любого $y' \notin X$ существуют окрестность $U(y', \varepsilon)$ и число $\alpha \in [0, 1)$ такие, что для всех $y \in U(y', \varepsilon)$ и $x \in X$ выполнено $\|F(y) - x\| \leq \alpha\|y - x\|$.

Обозначим через $D(k, m) = \text{conv}\{d^k, d^{k+1}, \dots, d^m\}$ выпуклую оболочку векторов $\{d^k, d^{k+1}, \dots, d^m\}$, через $\mathcal{N}_X(z) = \{v: v(x - z) \geq 0, x \in X\}$ — нормальный конус к множеству X в точке z .

Рассматривается итерационный процесс

$$\begin{aligned} x^{k+1} &= x^k + \lambda_k d^k, \\ d^k &= \frac{F(x^k + \lambda_k R(x^k)) - x^k}{\lambda_k}, \quad k = 0, 1, 2, \dots, \end{aligned} \quad (41)$$

где размер шагового множителя λ_k выбирается на основе накопленной информации за предыдущие итерации. Для заданной последовательности $\theta_m \rightarrow +0$ при $m \rightarrow \infty$ определим последовательность индексов $\{k_m\}$ и шаговые множители λ_k по следующим правилам:

1) Для $m = 0$ определим $k_m = 0$ и зафиксируем произвольное положительное λ_0 . Выберем $\alpha \in (0, 1)$.

2) Для данных m и k_m определим такой индекс k_{m+1} , что $0 \notin D(k_m, k) + \theta_m B$, $k_m \leq k < k_{m+1}$, $0 \in D(k_m, k_{m+1}) + \theta_m B$ при $\lambda_k = \lambda_{k_m}$ для $k_m \leq k < k_{m+1}$.

3) Положим $\lambda_{k_{m+1}} = \alpha \lambda_{k_m}$.

4) Увеличиваем номер итерации $m = m + 1$ и повторяем (41) для текущего значения λ_k .

Обоснование сходимости описанного процесса к некоторому стационарному множеству $Z^\dagger = \{z^\dagger \in X: 0 \in \mathcal{N}_X(z^\dagger) + R(x^\dagger)\}$ дает следующая теорема.

Теорема 7 (см. [17]). Пусть оператор F — локально сильно фейеровский относительно X , отображение $R(x)$ — локально сильный аттрактант множества Z^\dagger . Если последовательность $\{x^k\}$, порожденная процессом (41), ограничена, то все ее предельные точки принадлежат Z^\dagger .

Перепишем рекуррентные соотношения (38) в эквивалентной форме:

$$\begin{aligned} x^k &= ((x^w)^k: w \in W), \quad d^k = ((d^w)^k: w \in W), \\ (d^w)^k &= \frac{\pi_{X_w}((x^w)^k - \lambda_k G^w(x^k)) - (x^w)^k}{\lambda_k}, \\ x^{k+1} &= x^k + \lambda_k d^k, \quad k = 0, 1, 2, \dots, \end{aligned} \quad (42)$$

где размер шага λ_k регулируется по правилам 1–4.

Операция проектирования $\pi_X(\cdot)$ является ярким представителем локально сильно фейеровских относительно X операторов. Геометрическая интерпретация решения вариационного неравенства (6) говорит о том, что $G(x^\dagger) \in \mathcal{N}_X(x^\dagger)$. Так как последовательность $\{x^k\}$, генерируемая процессом (42), принадлежит компактному множеству X , то из теоремы 7 непосредственно вытекает следующий результат.

Теорема 8. Если $(-G(x))$ — локально сильный аттрактант множества решений X^\dagger вариационного неравенства (6), то все предельные точки последовательности $\{x^k\}$, порожденной процессом (42) и правилами 1–4, принадлежат X^\dagger .

Существует связь между свойствами аттрактантности и псевдомонотонности.

Утверждение 4. Если отображение $G(x)$ вариационного неравенства (6) строго псевдомонотонно на X , то $(-G(x))$ является локально сильным аттрактантом единственного решения $x^\dagger \in X^\dagger$.

Доказательство. Произвольно выберем точку $x' \in X$ и определим такую ее окрестность $U(x', \varepsilon)$, что $x^* \notin U(x', \varepsilon)$. В силу псевдомонотонности и непрерывности G имеет место оценка

$$\inf_{x \in U(x', \varepsilon)} (-G(x)(x^* - x)) \geq \inf_{x \in U} (-G(x)(x^* - x)) = \gamma \geq 0.$$

Значение $\gamma = 0$ противоречит строгой псевдомонотонности G , поэтому $\gamma > 0$ и для любых $x \in U(x', \varepsilon)$ выполнено

$$-G(x)(x^* - x) \geq \gamma > 0,$$

отсюда $(-G)$ — локально сильный аттрактант X^* . \square

Нетрудно показать, что если G — строго псевдомонотонно, то вариационное неравенство (6) имеет единственное решение. Действительно, пусть $x^1, x^2 \in X^\dagger$, $x^1 \neq x^2$. Имеют место неравенства:

$$G(x^1)(x^2 - x^1) \geq 0, \quad G(x^2)(x^2 - x^1) \leq 0,$$

что противоречит строгой псевдомонотонности G , следовательно, не может существовать более одного решения. Учитывая это свойство и утверждение 4, приходим к следующему выводу.

Следствие 2. Если $G(x)$ — строго псевдомонотонно на X , то последовательность $\{x^k\}$, порожденная процессом (42) и правилами 1–4, имеет единственную предельную точку $x^\dagger \in X^\dagger$.

Таким образом, полученные на основе теории фейеровских процессов условия сходимости проективного метода (35) для решения задач транспортного равновесия с фиксированным спросом менее ограничительные, чем ранее приведенные.

1.5. Построение матрицы корреспонденций

В задаче транспортного равновесия с фиксированным спросом каждая корреспонденция ρ_ω , $\omega = (i, j)$, рассматривается как средний поток пользователей, который из источника $i \in S$ должен прибыть в сток $j \in D$. В данном разделе вместо ρ_ω будем использовать обозначение ρ_{ij} , чтобы выделять характеристики источников i и стоков j .

Существуют разные методики для вычисления элементов матрицы $\rho = (\rho_{ij} : i \in S, j \in D)$, в том числе с применением математических моделей. Рассмотрим наиболее часто используемые, а именно гравитационную и энтропийную модели построения матрицы корреспонденций. Описание указанных моделей для транспортных сетей можно найти, например, в работах [6, 7, 9, 20, 21, 28, 40].

1.5.1. Гравитационная модель

Идею построения гравитационной модели дал всемирный закон тяготения, утверждающий, что все тела притягиваются друг к другу с силой, прямо пропорциональной произведению масс этих тел и обратно пропорциональной квадрату расстояния между ними. Применительно к транспортной системе в качестве тел выступают пункты, порождающие/поглощающие потоки, за массу тела принимается суммарный объем выезжающего/въезжающего потока, физическое расстояние можно заменить на любые другие затраты, связанные с передвижением. В самой простой форме гравитационная модель имеет вид

$$\rho_{ij} = \kappa \frac{s_i d_j}{c_{ij}^2}, \quad i \in S, \quad j \in D, \quad (43)$$

где s_i — общий объем выезжающих из пункта $i \in S$, d_j — общий объем въезжающих в пункт $j \in D$, c_{ij} — удельные затраты на передвижение из i в j , $\kappa > 0$ — калибровочный коэффициент.

Система (43) обладает существенным недостатком. Нетрудно видеть, что при увеличении объемов s_i и d_j , например, в два раза модель (43) приведет к увеличению корреспонденции ρ_{ij} в четыре раза, что совершенно нелогично. Поэтому вместо классической гравитационной модели (43) на практике используют ее модификацию, в которой к условию (43) добавляют дополнительные условия, например, балансовые ограничения на выезд и въезд. Кроме того, квадрат расстояния (затрат) c_{ij}^2 заменяют на так называемую функцию тяготения $f(c_{ij})$, характеризующую предпочтения индивидуумов при выборе пары источник-сток (i, j) для передвижения. В результате модифицированная гравитационная модель имеет вид

$$\rho_{ij} = \frac{s_i d_j}{f(c_{ij})}, \quad \sum_{j=1}^n \rho_{ij} = s_i, \quad \sum_{i=1}^m \rho_{ij} = d_j, \quad \rho_{ij} \geq 0, \quad i \in S, \quad j \in D,$$

или, что то же самое:

$$\rho_{ij} = \alpha_i \beta_j s_i d_j f(c_{ij}), \quad i \in S, \quad j \in D, \quad (44)$$

где калибровочные коэффициенты α_i и β_j определяются из системы

$$\alpha_i = \left[\sum_{j \in D} \beta_j d_j f(c_{ij}) \right]^{-1}, \quad \beta_j = \left[\sum_{i \in S} \alpha_i s_i f(c_{ij}) \right]^{-1}. \quad (45)$$

Очевидно, что система будет совместной только тогда, когда суммарные объемы по выезду и въезду равны: $\sum_{i \in S} s_i = \sum_{j \in D} d_j$.

Выбор функции тяготения f осуществляется либо в процессе калибровки модели на основе сопоставления расчетных данных по модели и эмпи-

рических наблюдений, либо на основе некоторых соображений о предпочтениях при выборе пары источник-сток. Одна из аппроксимаций функции имеет следующий вид: $f(c_{ij}) = \exp(-\gamma c_{ij}^\theta)$, где при расчете корреспонденций трудовых миграций полагают $\gamma \approx 0,065$, $\theta \approx 1$ (см., например, [28] и ссылки там).

Важно отметить, что величины α_i и β_j зависят от всего набора s_i и d_j , а следовательно, и объемы корреспонденций ρ_{ij} зависят от загрузки всей системы.

Численные значения α_i и β_j определяют специальной итеративной процедурой. В отечественной литературе такая процедура известна как метод балансировки Шацкого—Шелейховского [27, 30]. В зарубежной литературе метод балансировки имеет свою независимую историю развития. Например, в работе [33] описана следующая процедура: начиная с матрицы

$$\rho_{ij}^0 = s_i d_j f(c_{ij}) \left[\sum_{l \in D} d_l f(c_{il}) \right]^{-1},$$

каждая итерация метода состоит из последовательности операций:

$$\rho_{ij}^k = \begin{cases} \rho_{ij}^k d_j \left[\sum_{i \in S} \rho_{ij}^k \right]^{-1}, & \text{если } \sum_{i \in S} \rho_{ij}^k > d_j, \\ \rho_{ij}^k & \text{в противном случае,} \end{cases}$$

$$q_i = s_i - \sum_{j \in D} \rho_{ij}^k, \quad r_j = d_j - \sum_{i \in S} \rho_{ij}^k; \quad (46)$$

$$\rho_{ij}^{k+1} = \rho_{ij}^k + q_i r_j f(c_{ij}) \left[\sum_{l \in D} r_l f(c_{il}) \right]^{-1}.$$

Вычислительные эксперименты по расчету матрицы корреспонденций на примере УДС Владивостока [18, 19] (строилась матрица размерности 638×638) показали высокую скорость сходимости процесса (46) к искомой матрице корреспонденций — сбалансированная матрица была получена всего за 4 итерации.

1.5.2. Энтропийная модель

Как и в случае гравитационного подхода, идею построения энтропийной модели подсказала физика, а именно второй закон термодинамики, утверждающий, что любая замкнутая физическая система стремится достичь устойчивого равновесного состояния, которое характеризуется максимумом энтропии этой системы. Впервые концепция энтропии для определения матрицы корреспонденций была использована в работе [63].

Транспортную систему как систему передвижения индивидуумов по УДС города объединяет с физической наличие очень большого числа

неуправляемых элементов. При определенных допущениях, например таких, как неизменность затрат на проезд по маршрутам, неизменность топологии УДС (исключаются реконструкция, введение новых, закрытие старых дорог) и т. п., транспортную систему можно считать замкнутой. Таким образом, проблему определения корреспонденций ρ_{ij} можно ставить как задачу максимизации энтропии в транспортной системе.

Пусть задано фиксированное пространственное распределение населения по зонам, порождающим потоки, — как и ранее, назовем такие зоны *источниками* и объединим их в множество S — и по зонам, поглощающим потоки, — назовем их *стоками* и объединим в множество D . Источниками, например, могут служить районы жилых массивов, стоками — места приложения труда. Индивидуумы в транспортной системе перемещаются от источников к стокам. Предположим, что каждый индивидуум имеет уникальный идентификатор, например номер паспорта. Состояние транспортной системы определяется распределением «помеченных» индивидуумов между парами источник-сток.

При определении объемов корреспонденций значимым является только общее количество индивидуумов без детализации по составу их идентификаторов. Поэтому каждой паре источник-сток соответствует величина корреспонденции ρ_{ij} — количество индивидуумов, выезжающих из источника $i \in S$ и прибывающих в сток $j \in D$. Очевидно, что существует множество состояний, приводящих к одной и той же матрице корреспонденций $\rho = (\rho_{ij} : i \in S, j \in D)$. Следуя принципу максимизации энтропии, будем искать значения ρ_{ij} , доставляющие максимум функции $P(\rho)$, определяющей вероятность реализации состояния системы, соответствующего матрице корреспонденций ρ .

Обозначим через $\nu(\rho)$ вероятность каждой реализации матрицы ρ , через $Q(\rho)$ — количество состояний системы, соответствующих ρ . Тогда

$$P(\rho) = \nu(\rho) Q(\rho). \quad (47)$$

Пусть в системе имеется n источников и m стоков. Обозначим через $\mathcal{R} = \sum_{i=1}^n \sum_{j=1}^m \rho_{ij}$ общее количество индивидуумов в системе, через $\nu_{ij} > 0$ — вероятность выбора индивидуумом коммуникации ρ_{ij} .

По аналогии со схемой Бернулли значение $\nu(\rho)$ определяется формулой

$$\nu(\rho) = \nu_{11}^{\rho_{11}} \cdot \nu_{12}^{\rho_{12}} \cdot \dots \cdot \nu_{nm}^{\rho_{nm}} = \prod_{i=1}^n \prod_{j=1}^m \nu_{ij}^{\rho_{ij}}.$$

Вычислим количество состояний $Q(\rho)$. Если объем корреспонденции из источника 1 в сток 1 равен ρ_{11} , то количество способов достижения этого объема равно $C_{\mathcal{R}}^{\rho_{11}}$. Далее, из оставшейся части индивидуумов количество

способов достижения объема корреспонденции ρ_{12} равно $C_{\mathcal{R}-\rho_{11}}^{\rho_{12}}$, объема корреспонденции ρ_{13} равно $C_{\mathcal{R}-\rho_{11}-\rho_{12}}^{\rho_{13}}$ и так далее. В итоге получаем следующую формулу для $Q(\rho)$:

$$\begin{aligned} Q(\rho) &= C_{\mathcal{R}}^{\rho_{11}} \cdot C_{\mathcal{R}-\rho_{11}}^{\rho_{12}} \cdot C_{\mathcal{R}-\rho_{11}-\rho_{12}}^{\rho_{13}} \cdot \dots \cdot C_{\mathcal{R}-\sum_{i=1}^{n-1} \sum_{j=1}^{m-1} \rho_{ij}}^{\rho_{nm}} = \\ &= \frac{\mathcal{R}!}{(\mathcal{R}-\rho_{11})! \rho_{11}!} \cdot \frac{(\mathcal{R}-\rho_{11})!}{(\mathcal{R}-\rho_{11}-\rho_{12})! \rho_{12}!} \times \\ &\quad \times \frac{(\mathcal{R}-\rho_{11}-\rho_{12})!}{(\mathcal{R}-\rho_{11}-\rho_{12}-\rho_{13})! \rho_{13}!} \cdot \dots \cdot \frac{\left(\mathcal{R}-\sum_{i=1}^{n-1} \sum_{j=1}^{m-1} \rho_{ij}\right)!}{\rho_{nm}!} = \frac{\mathcal{R}!}{\prod_{i=1}^n \prod_{j=1}^m \rho_{ij}!}. \end{aligned}$$

Очевидно, что результат не зависит от того, в каком порядке берутся корреспонденции ρ_{ij} для вычисления количества способов распределения индивидуумов в системе.

Подставив рассчитанные значения $\nu(\rho)$ и $Q(\rho)$ в формулу (47), получаем критерий выбора наиболее вероятного состояния системы:

$$P(\rho) = \mathcal{R}! \prod_{i=1}^n \prod_{j=1}^m \frac{\nu_{ij}^{\rho_{ij}}}{\rho_{ij}!} \rightarrow \max. \quad (48)$$

Помимо требования максимизации вероятности $P(\rho)$ на значения ρ_{ij} , как правило, накладываются дополнительные условия. Самыми естественными из них являются балансовые ограничения и условия неотрицательности. Пусть в каждой зоне-источнике $i \in S$ задан общий объем выезжающих s_i , в каждой зоне-стоке $j \in D$ — общий объем въезжающих d_j . Рассмотрим только те корреспонденции ρ_{ij} , которые удовлетворяют следующим условиям:

$$\sum_{j=1}^m \rho_{ij} = s_i, \quad \sum_{i=1}^n \rho_{ij} = d_j, \quad \rho_{ij} \geq 0, \quad i \in S, \quad j \in D. \quad (49)$$

Очевидно, для совместности системы суммарный объем выезжающих должен быть равен суммарному объему въезжающих:

$$\sum_{i=1}^n s_i = \sum_{j=1}^m d_j = \mathcal{R}. \quad (50)$$

Дополнительно к условиям баланса (49) введем ограничение на общие затраты при проезде:

$$\sum_{i=1}^n \sum_{j=1}^m c_{ij} \rho_{ij} = C, \quad (51)$$

где c_{ij} — удельные затраты на передвижения из источника i в сток j , C — полные затраты в транспортной системе.

Таким образом, проблема построения матрицы корреспонденций $\rho = (\rho_{ij}: i \in S, j \in D)$ сводится к задаче условной оптимизации (48), (49), (51).

Нет сомнений, что в заданной форме (48) функция $P(\rho)$ весьма неприятна для оптимизации. Для удобства максимизации можно воздействовать на $P(\rho)$ любым монотонным оператором, например, прологарифмировать $P(\rho)$ и вместо (48) использовать критерий

$$\ln P(\rho) = \ln \mathcal{R}! + \sum_{i=1}^n \sum_{j=1}^m (\rho_{ij} \ln \nu_{ij} - \ln \rho_{ij}!) \rightarrow \max. \quad (52)$$

Проводя параллель между физической и транспортной системами, отметим наличие большого количества неуправляемых элементов, что позволяет предположить, что значения ρ_{ij} достаточно велики. Поэтому вполне правомерно для дальнейшего преобразования критерия (52) использовать формулу Стирлинга $\ln z! = z \ln z - z$, которая справедлива при больших z . Имеем

$$\ln P(\rho) \approx \mathcal{R} \ln \mathcal{R} + \sum_{i=1}^n \sum_{j=1}^m \rho_{ij} \ln \frac{\nu_{ij}}{\rho_{ij}}.$$

При фиксированных объемах выездов s_i и въездов d_j и выполнении равенства (50) величина $\mathcal{R} \ln \mathcal{R}$ постоянна и может быть исключена из критерия.

В результате проведенных преобразований наиболее вероятное состояние транспортной системы будет соответствовать такой матрице корреспонденций ρ , элементы которой удовлетворяют условиям (49), (51) и критерию

$$\sum_{i=1}^n \sum_{j=1}^m \rho_{ij} \ln \frac{\nu_{ij}}{\rho_{ij}} \rightarrow \max. \quad (53)$$

При построении энтропийной модели (49), (51), (53) предполагалось, что известна априорная информация о предпочтении индивидуумом одной коммуникации другой. Если же любое состояние система принимает с равной вероятностью, то есть для любых пар (i, j) значение ν_{ij} постоянно и определяется как $\nu_{ij} = \frac{1}{mn}$, то вместо критерия (53) рассматривают

$$\sum_{i=1}^n \sum_{j=1}^m \rho_{ij} \ln \frac{1}{\rho_{ij}} \rightarrow \max. \quad (54)$$

Допустимая область, задаваемая условиями (49), (51), образует полиэдральное множество. Целевая функция критерия (53) на допустимой области является строго вогнутой. В самом деле, матрица Гессе для (53)

имеет вид диагональной матрицы размерности $mn \times mn$ с элементами на главной диагонали $\left\{-\frac{1}{x_{ij}}\right\}$. Такая матрица отрицательно определена для любых m, n и $x_{ij} \geq 0$. Таким образом, задача (49), (51), (53) относится к классу задач выпуклой гладкой оптимизации. Строгая вогнутость целевой функции гарантирует единственность ее решения. Несмотря на свои хорошие свойства, для реальных транспортных сетей задача (49), (51), (53) имеет большую размерность, что в свою очередь серьезно усложняет применение на практике стандартных для этого класса задач численных методов. Так, например, для расчета трудовых миграций в УДС Владивостока [18, 19] территория города была поделена на зоны 800×800 метров. В результате получилась сетка 22×29 квадратов, каждый из которых одновременно являлся зоной-источником и зоной-стоком, при этом размерность задачи (49), (51), (53) составила 407 044 переменных, 1277 ограничений равенств (49), (51).

Для решения задачи (49), (53) разработана простая итерационная схема [27, 30]: начиная с матрицы $\rho^0 = (\rho_{ij}^0 = v_{ij} : i \in S, j \in D)$ на каждой итерации метода попеременно достигается выполнение балансовых ограничений для выездов и въездов:

$$\rho_{ij}^k = \rho_{ij}^k s_i \left[\sum_{j \in D} \rho_{ij}^k \right]^{-1}, \quad \rho_{ij}^{k+1} = \rho_{ij}^k d_j \left[\sum_{i \in S} \rho_{ij}^k \right]^{-1}. \quad (55)$$

В работе [4] доказана сходимость процесса (55) к оптимальному решению задачи (49), (53). Существуют и другие подходы к решению энтропийных моделей (см., например, [10, 40]).

Подробнее генезис и феноменология энтропийных моделей для поиска равновесного состояния макросистем, в том числе транспортных, рассмотрены в приложении Е. В. Гасниковой. Особый интерес тут представляет связь энтропийного критерия с динамикой достижения равновесного состояния.

1.5.3. Связь между гравитационной и энтропийной моделями

Количество переменных в задаче (49), (51), (53), как правило, во много раз превышает число ограничений. Традиционно в такой ситуации вместо исходной решается двойственная задача, которая в данном случае заключается в максимизации функции Лагранжа:

$$L(\rho, \lambda, \mu, \gamma) = \sum_{i=1}^n \sum_{j=1}^m \left[\rho_{ij} \ln \frac{v_{ij}}{\rho_{ij}} + \lambda_i (s_i - \rho_{ij}) + \mu_j (d_j - \rho_{ij}) + \gamma (C - c_{ij} \rho_{ij}) \right],$$

где $\lambda = (\lambda_i : i \in S)$ — вектор двойственных переменных, соответствующих балансовым ограничениям (49) для источников, $\mu = (\mu_j : j \in D)$ — вектор

двойственных переменных, соответствующих балансовым ограничениям (49) для стоков, γ — двойственная переменная, соответствующая ограничению по затратам (51).

Точка максимума для $L(\rho, \lambda, \mu, \gamma)$ должна удовлетворять условиям (49), (51) и системе уравнений

$$\ln \frac{v_{ij}}{\rho_{ij}} - 1 - \lambda_i - \mu_j - \gamma c_{ij} = 0, \quad i \in S, \quad j \in D. \quad (56)$$

Из системы (56) можно выразить корреспонденции

$$\rho_{ij} = v_{ij} \exp(-1 - \lambda_i - \mu_j - \gamma c_{ij}). \quad (57)$$

Видим, что для $v_{ij} \geq 0$ условие неотрицательности корреспонденций ρ_{ij} выполнено автоматически, поэтому может не учитываться при построении двойственной задачи и применении к ней численных методов. Однако заметим, что случай, когда $v_{ij} = 0$, означает отсутствие корреспонденции между парой (i, j) , следовательно, $\rho_{ij} = 0$ и максимизация функции Лагранжа должна рассматриваться в пространстве меньшей размерности.

Введем обозначения

$$\alpha_i = \frac{\exp(-1 - \lambda_i)}{s_i}, \quad \beta_j = \frac{\exp(-\mu_j)}{d_j}.$$

Тогда выражение (57) переписывается в виде

$$\rho_{ij} = \alpha_i \beta_j s_i d_j v_{ij} \exp(-\gamma c_{ij}). \quad (58)$$

При подстановке (58) в балансовые ограничения (49) определяются параметры α_i и β_j :

$$\alpha_i = \left[\sum_{j \in D} \beta_j d_j v_{ij} \exp(-\gamma c_{ij}) \right]^{-1}, \quad \beta_j = \left[\sum_{i \in S} \alpha_i s_i v_{ij} \exp(-\gamma c_{ij}) \right]^{-1}.$$

Отметим, что величина C на практике, как правило, неизвестна, поэтому лагранжевый множитель γ нельзя определить из решения уравнения (51). Значение γ определяется обычными методами калибровки.

Сравнивая выражение (58) с гравитационной моделью (44), видим, что отличие между ними состоит только в аналитическом задании функции тяготения $f(c_{ij})$. При $f(c_{ij}) = v_{ij} \exp(-\gamma c_{ij})$ гравитационная (44) и энтропийная (49), (51), (53) модели эквивалентны. Таким образом, при однородной цели поездок, при заданных объемах выездов s_i , въездов d_j , затратах на передвижение c_{ij} , при фиксированных полных затратах C существует наиболее вероятное распределение поездок между зонами (i, j) и это распределение совпадает с тем, которое задается гравитационной моделью с экспоненциальной функцией притяжения.

1.6. Парадоксы транспортного равновесия

В данном разделе рассматривается ряд антиинтуитивных примеров транспортных ситуаций, в которых применение принципа равновесия приводит к неожиданным решениям.

1.6.1. Парадокс Браесса

Пример Пигу (см. раздел 1.3) заставляет усомниться в эффективности «невидимой руки рынка» Адама Смита, которая, направляя эгоистичные действия пользователей сети, позволяет достичь общественного блага. Последующий пример Браесса показывает, что конкурентное бескоалиционное равновесие может не только отклоняться от системного оптимума, но и ухудшать ситуацию для всех участников движения.

Рассмотрим появление парадокса Браесса в результате последовательных весьма вероятных трансформаций транспортной сети окрестностей г. Владивостока, которые представлены в серии рис. 3.

Будем рассматривать ситуацию с точки зрения перевозок Аэропорт — Владивосток, с общей потребностью в перевозках 6 условных единиц. На рис. 3 изображены воображаемые этапы изменения участка транспортной сети в окрестности Владивостока, которые могут быть связаны с созданием игровой зоны на мысе Черепahi.

а) Начальное состояние. Аэропорт и Владивосток соединены двумя дорогами, одна из которых проходит через г. Артем, а другая через мыс Черепahi — место, где будет построена игровая зона «Геймланд». В начальный момент обе дороги невысокого качества и, как показано на рис. 3 а), время проезда по ним сильно зависит от нагрузки y .

Очевидно, что в силу симметрии равновесные потоки распределяются поровну между двумя маршрутами Аэропорт — мыс Черепahi — Владивосток и Аэропорт — Артем — Владивосток с соответствующими потоками $y = 3$. Пользовательские затраты на проезд — 90, системные — 540.

б) Построено шоссе Аэропорт — «Геймланд». Снизилась зависимость времени проезда из Аэропорта в «Геймланд», в затратах на проезд появилась постоянная составляющая, которая может представлять собой время проезда по пустой дороге. Часть равновесного трафика переместилась на направление Аэропорт — мыс Черепahi — Владивосток (3,17), соответственно поток по другому маршруту Аэропорт — Артем — Владивосток упал до 2,83. Пользовательские затраты на проезд составили 84,88, системные — $84,88 \cdot 6 = 509,28$.

в) Построено шоссе Артем — Владивосток. Полученные доходы от игорного бизнеса позволили модернизировать часть одного из маршрутов Владивосток — Артем — Аэропорт, в результате чего равновесные потоки

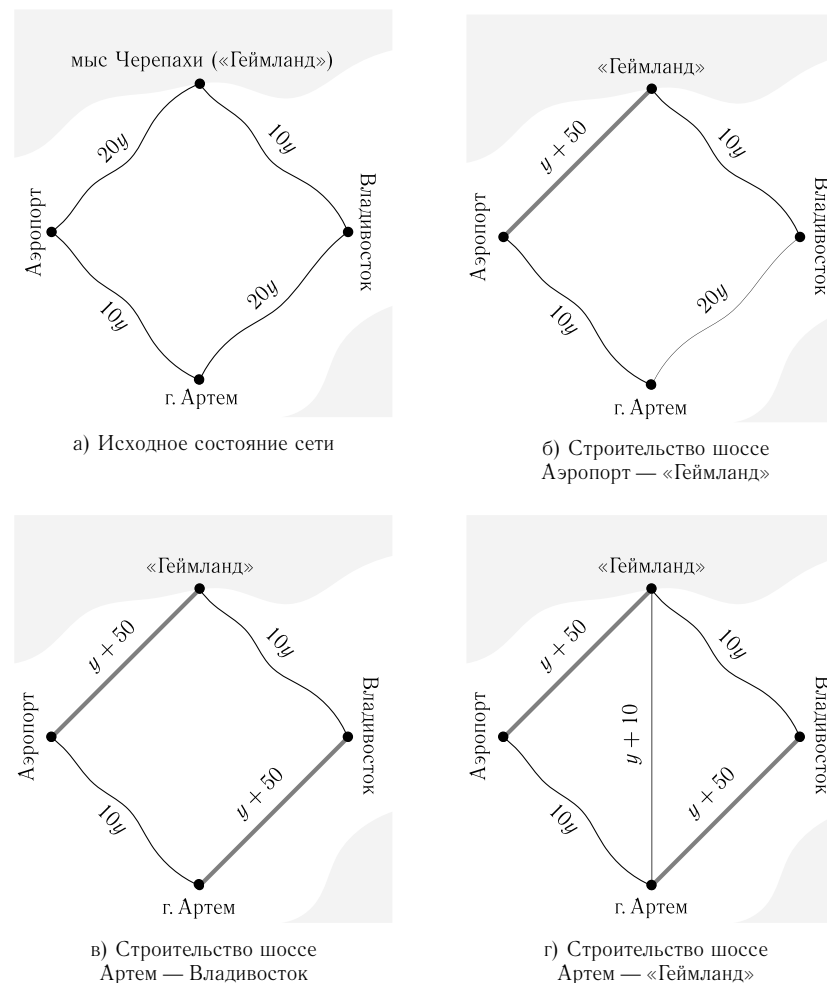


Рис. 3. Парадокс Браесса

снова стали симметричными и затраты пользователей составили 83, а системные — 498, дальнейшее снижение.

г) Построено шоссе Артем — «Геймланд». Поскольку игорная зона обслуживается в основном жителями Артема, был поставлен и положительно решен вопрос о строительстве дороги Артем — «Геймланд». Затраты на проезд соответствовали классу уже построенных дорог, а постоянное слагаемое уменьшилось в силу территориальной близости.

Равновесные потоки теперь распределятся по трем маршрутам: Владивосток — «Геймланд» — Артем — Аэропорт, Владивосток — «Геймланд» — Аэропорт, Владивосток — Артем — Аэропорт, причем нагрузка на каждый из них будет составлять 2 единицы трафика, пользовательские затраты на проезд неожиданно возросли до 92, а системные — до 552!

В результате как системные, так и личные затраты превзошли даже первоначальный уровень а) несмотря на то, что на каждом предыдущем этапе мы улучшали транспортную ситуацию как с пользовательской, так и с системной точки зрения.

Причиной этого эффекта является то, что постройка на этапе г) шоссе Артем — «Геймланд» создала оппортунистическую возможность проехать по маршруту Аэропорт — Артем — «Геймланд» — Владивосток. При нулевом потоке на маршруте Артем — «Геймланд» временные затраты составляют 70 и провоцируют водителей на выбор именно этого маршрута. Однако когда эта идея овладеет массами, то поток по сегменту Артем — «Геймланд» будет уже ненулевой, что увеличит соответствующие общие затраты. Равновесная ситуация установится при одинаковых затратах (временах) по всем маршрутам, что и приводит к этому парадоксальному результату.

1.6.2. Транспортно-экологические парадоксы

Существует ряд парадоксов [55], связанных с транспортными ситуациями, в которых помимо времени проезда учитываются и дополнительные критерии. Одним из таких критериев, важных в настоящее время, является загрязнение окружающей среды (ЗОС). ЗОС является сложным многокомпонентным понятием, включающим различные виды ущерба для окружающей среды: газовое и тепловое загрязнение, разрушение сложившихся природных ландшафтов, мест обитания редких животных и пр. В данном случае будем все же считать, что ЗОС измеряется некоторым универсальным показателем, связанным с данным участком дороги и зависящим, вообще говоря, от потока транспорта по этой дороге. ЗОС от различных участков дороги суммируются, образуя итоговый ЗОС либо от маршрутов, либо от всей транспортной сети в целом.

1.6.2.1. Экологический парадокс Браесса. В своей схематической форме сеть, реализующая парадокс Браесса, представлена на рис. 4, где на каждом ребре показаны как временные затраты $\tau(y)$, так и экологический ущерб $e(y)$ как функции потоков по этим ребрам y . Взяв для предполагаемого потока те же данные, что в предыдущем примере, оценим ЗОС до и после строительства новой дороги. Для исходного состояния сети ЗОС оценивается как

$$E = 2 \cdot 3 \cdot (0,2 + 0,1) = 1,8.$$

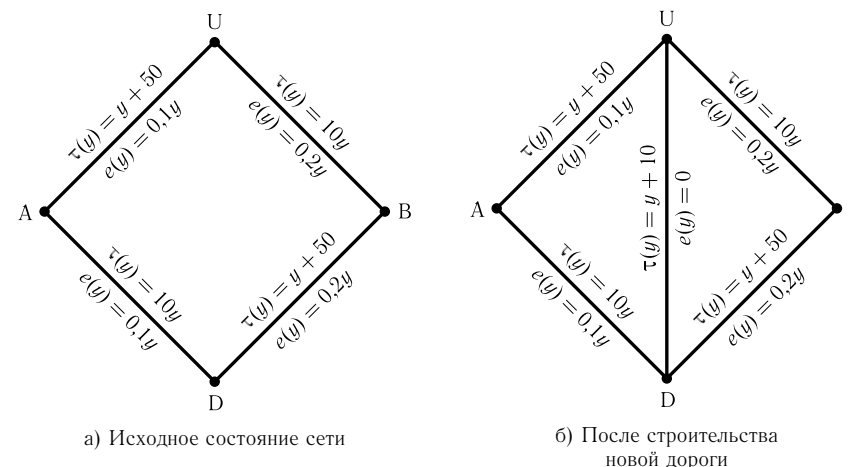


Рис. 4. Парадокс Браесса

После строительства новой дороги с нулевым экологическим ущербом ЗОС становится равным

$$E = 4 \cdot 0,2 + 2 \cdot 0,1 + 2 \cdot 0 + 4 \cdot 0,2 + 2 \cdot 0,1 = 2!$$

Как нетрудно понять, в данном случае парадокс вызван тем, что в результате перераспределения потоков увеличились потоки именно по тем дугам, которые имеют максимальные удельные приращения ЗОС.

1.6.2.2. Экологический треугольник. Рассмотрим теперь еще более простую транспортную сеть, представленную на рис. 5. Так же как и ранее,

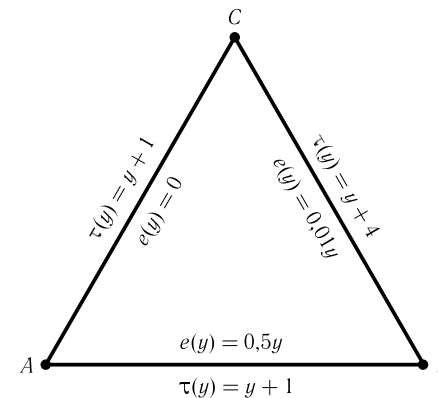


Рис. 5. Уменьшение перевозок вызывает увеличение ЗОС

на дугах этой сети приведены формулы, описывающие временные затраты $\tau(y)$ и экологический ущерб $e(y)$ как функции потока y по этой дуге. Пусть требуется перевезти 2 единицы груза из C в B и одну единицу из C в A . Очевидно, что достаточно рассмотреть 3 маршрута: $p_1 = C \rightarrow A$, $p_2 = C \rightarrow A \rightarrow B$ и $p_3 = C \rightarrow B$. Условия равновесия совместно с условиями удовлетворения спроса на перевозки дают систему уравнений

$$\begin{aligned}(x_1 + x_2) + 1 + x_2 + 1 &= x_3 + 4, \\ x_2 + x_3 &= 2, \quad x_1 = 1,\end{aligned}$$

где x_i , $i = 1, 2, 3$, — это потоки по маршрутам p_i , $i = 1, 2, 3$. Решение этой системы дает равновесные потоки $x_1^\dagger = x_2^\dagger = x_3^\dagger = 1$ с общим экологическим ущербом $E = 0,51$.

Теперь предположим, что спрос на перевозки по маршруту $C \rightarrow A$ упал до $1/2$. Тогда решение аналогичной системы дает $x_1^\dagger = 1/2$, $x_2^\dagger = 7/6 > 1$, $x_3^\dagger = 5/6 < 1$ с общим экологическим ущербом $E = 0,5 \cdot 7/6 + 0,01 \cdot 5/6 \approx 0,591 > 0,51$.

Заметим, что увеличение экологического ущерба в этом случае вызвано уменьшением нагрузки на экологически чистую дугу сети. С одной стороны, это вызвало переход части трафика с маршрута p_3 , не проходящего через дугу $C \rightarrow A$, на маршрут p_2 , проходящий через эту дугу, однако, с другой стороны, это вызвало увеличение трафика по дуге $A \rightarrow B$ с высоким экологическим ущербом, и суммарный эффект оказался негативным.

1.6.2.3. Рокадная экология. Последующий пример показывает, как такая популярная мера, как строительство рокадной дороги улучшенного качества, может на самом деле ухудшить экологическую ситуацию.

Рассмотрим дорожную сеть, изображенную на рис. 6, часть а). Предположим, что эта сеть предназначена для перемещения автомобиль-

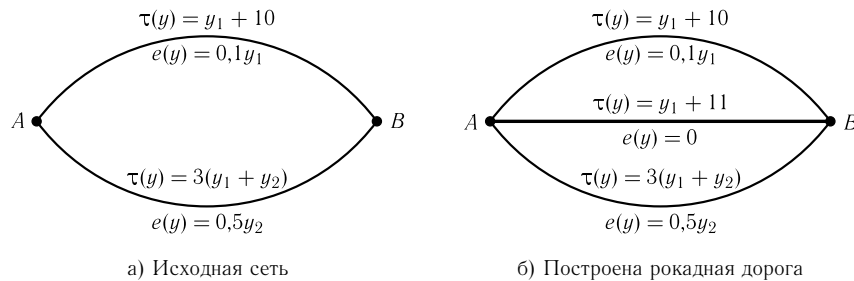


Рис. 6. Новая рокадная дорога вызывает увеличение ЗОС

ного трафика в объеме 5 условных единиц из точки A в точку B по двум дублирующим дорогам различного качества, временные затраты по

которым задаются соотношениями на соответствующих дугах. В этих соотношениях y_1 означает поток по верхней дуге, y_2 — по нижней. Зависимость времени проезда по нижней дуге от потока по верхней может быть вызвана указанием приоритета на соответствующем перекрестке.

Определяющая система уравнений для равновесных потоков имеет вид

$$y_1 + y_2 = 5, \quad y_1 + 10 = 3(y_1 + y_2),$$

откуда $y_1^\dagger = 5$, $y_2^\dagger = 0$. Затраты пользователей на проезд составляют при этом $\tau^\dagger = 15$, а экологический ущерб составляет $0,5$.

Если, как показано на правой части рис. 6, построена новая дорога с нулевым ущербом для окружающей среды и временными характеристиками $\tau(y) = y + 11$, то новая определяющая система будет иметь вид

$$y_1 + y_2 + y = 5, \quad y_1 + 10 = 3(y_1 + y_2) = y + 11,$$

и ее решение: $y_1^\dagger = y_2^\dagger = 2$, $y^\dagger = 1$. Затраты пользователей на проезд при этом уменьшились до $\tau^\dagger = 12$, а экологический ущерб возрос до $1,2!$

1.6.2.4. Перераспределение спроса. Еще один пример показывает, что перенос части пассажиров с транспортного средства экологически более вредного (скажем, автобус) на менее вредный (например, трамвай) может в действительности увеличить ЗОС.

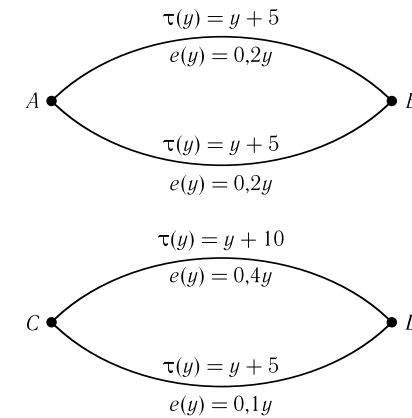


Рис. 7. Сети автомобильных дорог и трамвайных линий

На рис. 7 изображена транспортная сеть, состоящая из двух изолированных подсетей: автомобильной AB и трамвайной CD . В пунктах A и C происходит разделение потока пассажиров: 10 единиц потока выбирают автобус, 5 единиц выбирают трамвай. Это распределение объемов перевозки считается заданным. Далее пассажиры, пользуясь свободой выбора,

определяют, каким из двух автобусных и двух трамвайных маршрутов воспользоваться, и делают это основываясь на теории равновесия. В данном случае равновесие будет заключаться в равенстве временных затрат по каждой паре маршрутов и соответствующие потоки будут равны: в автомобильной подсети в силу симметрии общий объем перевозок поделится поровну: $y_{\text{авт}}^1 = y_{\text{авт}}^2 = 5$ с временными затратами для каждого пассажира в 10 временных единиц. Суммарный ЗОС в автомобильной подсети будет равен 2. В трамвайной подсети возникнет ситуация, описанная фактически в примере Пигу (см. раздел 1.3), и поток $y_{\text{трам}}^1$ по верхней дуге будет равен нулю, а по нижней дуге $y_{\text{трам}}^2 = 5$ также с одинаковыми временными затратами на пассажира, равными 10. Суммарный ЗОС в трамвайной сети будет равен 0,5 и существенно ниже автомобильного. Ориентируясь на эти объемы ЗОС, может возникнуть идея перенести часть потока из автомобильной сети в трамвайную и уменьшить тем самым суммарный ЗОС, который равен первоначально 2,5. Новое равновесное решение будет иметь вид:

Автомобильная сеть. В силу симметрии потоки равны и составляют $y_{\text{авт}}^1 = y_{\text{авт}}^2 = 3,75$. Экологический ущерб $E = 1,5$.

Трамвайная сеть. $y_{\text{трам}}^1 = 1,25$, $y_{\text{трам}}^2 = 6,25$ с одинаковыми временными затратами 11,25 и суммарным экологическим ущербом $0,4 \cdot y_{\text{трам}}^1 + 0,1 \cdot y_{\text{трам}}^2 = 1,125$.

Суммарный экологический ущерб составляет 2,625, что превосходит (!) ущерб в предыдущем варианте распределения нагрузки для двух видов трафика.

То, что перенос трафика осуществляется в размере 2,5 единиц, на самом деле несущественно, любое уменьшение объема перевозок по автомобильной сети вызывает появление ненулевого объема перевозок по экологически затратной дуге 3 в трамвайной сети, что не компенсируется уменьшением объемов перевозок по дугам 1, 2. Увеличение объемов перевозок по автомобильной сети также не приводит к уменьшению ЗОС, так как вызывает в два раза больший экологический ущерб, чем уменьшение ЗОС от трамвайного трафика, который будет продолжать концентрироваться на дуге 4.

1.7. Практическая работа

Упражнение 1. На входе в Великий Федеральный Университет есть две двери, через которые входят и выходят студенты. При подходе к ним перед каждым студентом возникает проблема, какой дверью воспользоваться, если он заинтересован в скорейшем входе или выходе. Найти равновесное распределение потоков в двух случаях:

1) задержка в дверях одинакова для входящих и выходящих и пропорциональна произведению интенсивностей входящего и выходящего потоков;

2) задержка в дверях пропорциональна интенсивности потока студентов,двигающихся в том же направлении, плюс задержка, пропорциональная произведению интенсивностей входящего и выходящего потоков.

Общее количество входящих и выходящих за единицу времени студентов считать одинаковым.

Упражнение 2. Рассматривается транспортная сеть $\Gamma = (V, E)$ (пример сети взят из работы [36]), состоящая из 25 вершин ($|V| = 25$) и 40 ориентированных дуг ($|E| = 40$). Топология сети с направлением дуг представлена на рис. 8.

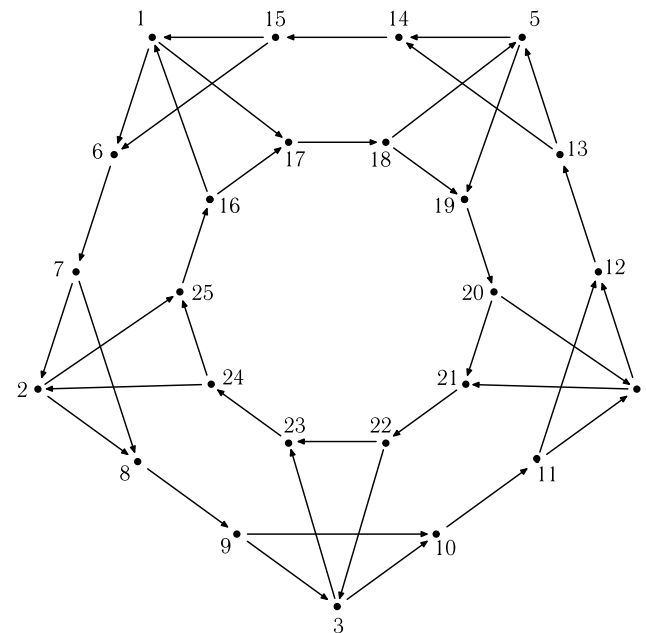


Рис. 8. Транспортная сеть $\Gamma = (V, E)$

Дороги (дуги) транспортной сети Γ поделены на четыре категории:

- 1) магистрали $E_h = \{(6 \rightarrow 7), (8 \rightarrow 9), (10 \rightarrow 11), (12 \rightarrow 13), (14 \rightarrow 15), (17 \rightarrow 18), (19 \rightarrow 20), (21 \rightarrow 22), (23 \rightarrow 24), (25 \rightarrow 16)\}$;
- 2) выезды $E_{ex} = \{(16 \rightarrow 1), (15 \rightarrow 1), (24 \rightarrow 2), (7 \rightarrow 2), (22 \rightarrow 3), (9 \rightarrow 3), (20 \rightarrow 4), (11 \rightarrow 4), (18 \rightarrow 5), (13 \rightarrow 5)\}$;

3) въезды $E_{en} = \{(1 \rightarrow 6), (1 \rightarrow 17), (2 \rightarrow 25), (2 \rightarrow 8), (3 \rightarrow 23), (3 \rightarrow 10), (4 \rightarrow 21), (4 \rightarrow 12), (5 \rightarrow 19), (5 \rightarrow 14)\}$;

4) второстепенные дороги $E_s = \{(15 \rightarrow 6), (7 \rightarrow 8), (9 \rightarrow 10), (11 \rightarrow 12), (13 \rightarrow 14), (16 \rightarrow 17), (18 \rightarrow 19), (20 \rightarrow 21), (22 \rightarrow 23), (24 \rightarrow 25)\}$.

Пропускная способность магистралей равна 140 единицам потока, остальных дуг — 70 ед.

Категория дороги влияет на затраты при передвижении. Минимальные транспортные затраты τ_e^0 по каждому из участков $e \in E$ определяются по формуле $\tau_e^0 = \kappa_e l_e$, где l_e — длина дуги e , данные приведены в таблице 1, $\kappa_e > 0$ — коэффициент, зависящий от категории, которой принадлежит дуга e :

$$\kappa_e = \begin{cases} 0,011, & e \in E_h, \\ 0,025, & e \in E_{ex} \cup E_{en}, \\ 0,033, & e \in E_s. \end{cases}$$

Таблица 1. Длины дуг сети $\Gamma = (V, E)$

Дуга	Длина	Дуга	Длина	Дуга	Длина
6 → 7	4	16 → 1	6	1 → 6	3
8 → 9	10	15 → 1	9	1 → 17	7
10 → 11	3	24 → 2	3	2 → 25	6
12 → 13	3	7 → 2	8	2 → 8	2
14 → 15	5	22 → 3	5	3 → 23	6
17 → 18	1	9 → 3	1	3 → 10	5
19 → 20	2	20 → 4	10	4 → 21	6
21 → 22	6	11 → 4	8	4 → 12	8
23 → 24	9	18 → 5	5	5 → 19	7
25 → 16	2	13 → 5	3	5 → 14	4
15 → 6	1	7 → 8	6	9 → 10	4
11 → 12	3	13 → 14	9	16 → 17	10
18 → 19	4	20 → 21	6	22 → 23	10
24 → 25	1				

В сети Γ выделено пять вершин-источников $S = \{1, 2, 3, 4, 5\}$ с заданными объемами выходящего трафика $s = (69, 90, 10, 100, 53)$ и пять вершин-стоков $D = \{17, 19, 21, 23, 25\}$ с заданными объемами входящего трафика $d = (128, 59, 34, 61, 40)$.

Используя аппарат математического моделирования и численные методы, рассчитать объемы корреспонденций ρ_w для всех пар источник-сток $w \in W = S \times D$.

При моделировании необходимо учесть, что предпочтения при выборе пары w определяются функцией $f(c_w) = \exp(-0,065c_w)$, где c_w — минимальные транспортные затраты на передвижение для пары w . Предполагается, что величина c_w характеризует длину кратчайшего пути для

каждой пары w и определяется из соотношений: $c_w = 0,05$, если источник и сток совпадают, в противном случае

$$c_w = \min_{p \in P_w} \sum_{e \in E} \theta_{ep} \tau_e^0, \quad \theta_{ep} = \begin{cases} 1, & \text{если путь } p \text{ проходит через дугу } e; \\ 0 & \text{в противном случае,} \end{cases}$$

где через P_w обозначено множество всех допустимых маршрутов передвижения для пары w .

Для рассчитанных корреспонденций сравнить загрузку транспортной сети, индивидуальные и общие транспортные затраты при нормативном и дескриптивном распределении потоков. Под нормативным распределением понимается централизованное управление движением, имеющее своей целью минимизацию совокупных транспортных затрат (*второй поведенческий принцип Вардрона*). Под дескриптивным — отсутствие централизованного управления, каждый пользователь выбирает маршрут следования исходя из минимизации собственных транспортных затрат (*первый поведенческий принцип Вардрона*).

Удельные транспортные затраты для каждой пары источник-сток w складываются из затрат по дугам τ_e , входящим в маршрут следования из источника в сток. В свою очередь на значение τ_e влияет величина потока, проходящего по дуге $e \in E$, такая зависимость описывается функцией $\tau_e(y_e) = \tau_e^0(1 + (y_e/c_e)^4)$, где c_e — пропускная способность дуги e .

Литература

1. Антипин А. С. Равновесное программирование: проксимальные методы // Автоматика и телемеханика. 1997. № 8. С. 125–137.
2. Ашманов С. А. Математические модели и методы в экономике. М.: Изд-во МГУ, 1980.
3. Байокки К., Капело А. Вариационные и квазивариационные неравенства. Приложения к задачам со свободной границей. М.: Наука, 1984.
4. Брэгман Л. Д. Доказательство сходимости метода Шелейховского для задачи с транспортными ограничениями // ЖВМ и МФ. 1967. Т. 7, № 1. С. 147–156.
5. Бенсусан А., Лионс Ж.-Л. Импульсное управление и квазивариационные неравенства. М.: Наука, 1987.
6. Васильева Е. М., Левит Б. Ю., Лившиц В. Н. Нелинейные транспортные задачи на сетях. М.: Финансы и статистика, 1981.
7. Васильева Е. М., Игудин Р. В., Лившиц В. Н. Оптимизация планирования и управления транспортными системами. М.: Транспорт, 1987.
8. Васин В. В., Еремин И. И. Операторы и итерационные процессы фейеревского типа. Теория и приложения. Екатеринбург: УрО РАН, 2005.
9. Вильсон А. Дж. Энтропийные методы моделирования сложных систем. М.: Наука, 1978.

10. Гасникова Е.В. Двойственные мультипликативные алгоритмы для задачи энтропийно-линейного программирования // ЖВМ и МФ. 2009. Т.49, №3. С.453–464.
11. Данскин Дж. Теория максимина и ее приложение к задачам распределения вооружений. М.: Сов. радио, 1970.
12. Еремина И.И., Мазурова В.Д. Нестационарные процессы математического программирования. М.: Наука, 1979.
13. Корпелевич Г.М. Экстрарадиентный метод для отыскания седловых точек и других задач // Экономика и матем. методы. 1976. Т.1, №4. С.747–756.
14. Никайдо Х. Выпуклые структуры и математическая экономика. М.: Мир, 1972.
15. Нурминский Е.А. Использование дополнительных малых воздействий в фейеровских моделях итеративных алгоритмов // ЖВМ и МФ. 2008. Т.48, №12. С.2121–2128.
16. Нурминский Е.А. Фейеровские процессы с малыми возмущениями // Доклады РАН. 2008. Т.422, №5. С.601–605.
17. Нурминский Е.А. Фейеровские алгоритмы с адаптивным шагом // ЖВМ и МФ. 2011. Т.51, вып. 5. С.1–11.
18. Нурминский Е.А., Шамрай Н.Б. Моделирование транспортных потоков г. Владивостока на основе теории равновесия // Systeme de transport si logistica: Materialele Conf. Int., Chisinau, 22–23 octombrie 2009; red. Resp. Dumitru Solomon; Acad. de Transporturi, Informatica si Comunicatii. Ch.: Evrica, 2009. P.334–348.
19. Нурминский Е.А., Шамрай Н.Б. Прогнозное моделирование автомобильного трафика Владивостока // Труды МФТИ. 2010. Т.2, №4(8). С.119–129.
20. Попков Ю.С., Посохин М.В., Гутнов А.Э., Шмультян Б.Л. Системный анализ и проблемы развития городов. М.: Наука, 1983.
21. Попков Ю.С. Макросистемные модели пространственной экономики. М.: КомКнига, 2008.
22. Стенбринк П.А. Оптимизация транспортных сетей. М.: Транспорт, 1961.
23. Попов Л.Д. Введение в теорию, методы и экономические приложения задач о дополнителности. Екатеринбург: Изд-во Урал. ун-та, 2001.
24. Тодд М.Дж. Вычисление неподвижных точек и приложения к экономике. М.: Наука, 1983.
25. Шамрай Н.Б. Решение задач транспортного равновесия с декомпозицией по ограничениям // Труды Всероссийской конференции «Равновесные модели в экономике и энергетике». Иркутск: Изд-во ИСЭМ СО РАН. 2008. С.618–624.
26. Шамрай Н.Б. Поиск потокового равновесия проективными методами с использованием декомпозиции и генерации маршрутов // Автоматика и телемеханика. 2012. №3.
27. Шацкий Ю.А. Расчет схемы расселения и трудовых корреспонденций при разработке генерального плана города // Развитие системы городского транспорта. Киев, 1971. №4. С.3–14.
28. Швецов В.И. Математическое моделирование транспортных потоков // Автоматика и телемеханика. 2003. №11. С.3–46.

29. Швецов В.И. Алгоритмы распределения транспортных потоков // Автоматика и телемеханика. 2009. №10. С.148–157.
30. Шелейховский Г.В. Транспортные основания композиции городского плана. Л., 1936.
31. Хоботов Е.Н. О модификации экстрарадиентного метода для решения вариационных неравенств и некоторых задач оптимизации // ЖВМ и МФ. 1987. Т.27, №10. С.1462–1473.
32. Agdeppa R.P., Yamashita N., Fukushima M. The traffic equilibrium problem with nonadditive costs and its monotone mixed complementarity problem formulation // Transportation Research Part B. 2007. №41. P.862–874.
33. Arrowsmith G.A. A behavioural approach to obtaining a doubly constrained trip distribution model // Operational Research Quarterly. 1973. V.24, №1. P.101–111.
34. Bar-Gera H. Origin-based algorithm for the traffic assignment problem // Transportation Science. 2002. V.36, №4. P.398–417.
35. Beckmann M., McGuire C.B., Winsten C.B. Studies in the economics of transportation. RM-1488. Santa Monica: RAND Corporation, 1955.
36. Bertsekas D., Gafni E. Projection methods for variational inequalities with application to the traffic assignment problem // Mathematical Programming Study. 1982. №17. P.139–159.
37. Boyce D., Ralevic-Dekic B., Bar-Gera H. Convergence of traffic assignments: how much is enough? // Journal Transport Engineer. 2004. V.130, №1. P.49–55.
38. Chen M., Bernstein D.H., Chien S.I.J., Mouskos K. Simplified formulation of toll design problem // Transportation Research Record. 1999. №1667. P.88–95.
39. Dafermos S. Traffic equilibrium and variational inequalities // Transportation Science. 1980. V.14, №1. P.42–54.
40. Fang S.-C., Rajasekera J.R., Tsao H.-S.J. Entropy optimization and mathematical programming. Kluwer Academic Publisher, 1997.
41. Facchinei F., Pang J.-S. Finite-Dimensional Variational Inequalities and Complementarity Problems (V.I, II). Springer, 2003.
42. Frank M., Wolfe P. An algorithm for quadratic programming // Naval Research Logistics Quarterly. 1956. V.3. P.95–110.
43. Gabriel S.A., Bernstein D. The traffic equilibrium problem with nonadditive path costs // Transportation Science. 1997. V.31, №4. P.337–348.
44. Iusem A.N. An iterative algorithm for the variational inequality problem // Comput. and Appl. Mathematics. 1994. V.13, №2. P.103–114.
45. Janson B., Zozaya-Gorostiza C. The problem of cyclic flows in traffic assignment // Transportation Research Part B. 1987. V.21, №4. P.299–310.
46. Leventhal T., Nemhauser G.L., Trotter L.Jr. A column generation algorithm for optimal traffic assignment // Transportation Science. 1973. №7. P.168–176.
47. Michelot C. A finite algorithm for finding the projection of a point onto the canonical simplex of \mathbf{R}^n // J. of Optimization Theory and Appl. 1986. V.50, №1. P.195–200.
48. Knight F.H. Some fallacies in the interpretation of social cost // The Quarterly Journal of Economics. 1924. V.38, №4. P.582–606.

49. *Konnov I. V.* Combined relaxation methods for variational inequalities. Berlin: Springer, 2001.
50. *Konnov I. V.* Equilibrium Models and Variational Inequalities. Elsevier Science, 2007.
51. *Kravchuk A. S., Neittaanmaki P. J.* Variational and Quasi-Variational Inequalities in Mechanics. Springer, 2007.
52. *Lo H. K., Chen A.* Traffic equilibrium problem with rout-specific costs: formulation and algorithms // Transportation Research Part B. 2000. V. 34, № 6. P. 493–513.
53. *Marcotte P.* Application of Khobotov's algorithm to variational inequalities and network equilibrium problems // INFOR. 1992. V. 29, № 4. P. 258–270.
54. *Nagurney A.* Network Economics: A Variational Inequality Approach. Dordrecht: Kluwer Academic Publishers, 1999.
55. *Nagurney A., Dong J.* Paradoxes in networks with zero emission links: implications for telecommunications versus transportation // Transportation Research Part D. 2001. V. 6, № 4. P. 283–296.
56. *Nemirovsky A., Yudin D.* Informational complexity and efficient methods for solution of convex extremal problems. N.Y.: Wiley, 1983.
57. *Nesterov Yu., de Palma A.* Stationary dynamic solutions in congested transportation networks: summary and perspectives // Networks and Spatial Economics. 2003. № 3. P. 371–395.
58. *Nurminski E. A.* Envelope stepsize control for iterative algorithms based on Fejer processes with attractants // Optimization Methods and Software. 2010. V. 25, № 1. P. 97–108.
59. *Patriksson M.* The traffic assignment problem — models and methods. Utrecht, Netherlands: VSP, 1994.
60. *Piugou A. C.* The economics of welfare. London: MacMillan, 1932. 4-th edition. (Русский перевод: Пигу А. С. Экономическая теория благосостояния. Т. 1–2. Сер. Экономическая мысль Запада. М.: Прогресс, 1985).
61. *Roughgarden T., Tardos E.* How bad is selfish routing? // Journal of the ACM. 2002. V. 49, № 2. P. 236–259.
62. *Sun D.* A projection and contraction method for the nonlinear complementarity problem and its extensions // Mathematica Numerica Sinica. 1994. V. 16. P. 183–194.
63. *Wilson A. G.* A statistical theory of spatial distribution models // Transportation Research. 1967. V. 1. P. 253–270.
64. *Wardrop J.* Some Theoretical Aspects of Road Traffic Research // Proceedings of the Institute of Civil Engineers. 1952.
65. *Wang Y. J., Xiu N. H., Wang C. Y.* Unified framework of extragradient-type methods for pseudomonotone variational inequalities // J. of Optimization Theory and Appl. 2001. V. 111, № 3. P. 641–656.
66. *Xiu N., Zhang J.* Some recent advances in projection-type methods for variational inequalities // J. of Comput. and Appl. Mathematics. 2003. V. 152. P. 559–585.

Глава 2

Математические модели транспортных потоков

2.1. Макроскопические модели

В разделе 2.1 приводятся основные (с исторической точки зрения, и с точки зрения возможных приложений) *макроскопические модели* транспортных потоков. Много внимания уделяется гидродинамическим аналогиям. *Транспортный поток* уподобляется *сжимаемой жидкости с мотивацией*, которая присутствует, например, в *уравнении состояния* транспортного потока (зависимости *скорости потока* от *плотности*). Ключевым понятием этого раздела является *обобщенное решение* начальной задачи Коши для закона сохранения, описывающего транспортный поток. Так, например, разрывы обобщенного решения интерпретируются как границы заторов (переход от свободного движения к заторному).

2.1.1. Модель Лайтхилла—Уизема—Ричардса (LWR)

Во второй половине 40-х годов и в 50-е годы XX века в СССР и США интенсивно занимались исследованием процессов, возникающих при взрыве бомбы (см., например, монографии [1,2]). В частности, большое внимание было уделено изучению начально-краевых задач для уравнения типа закона сохранения и систем таких уравнений. В это же время наблюдался и рост приложений, в которых встречаются схожие уравнения [3,4]. Так в 1955 г. независимо в работах [5,6] (см. также [7]) была предложена, по-видимому, первая макроскопическая (гидродинамическая) модель однополосного¹⁾ транспортного потока, названная впоследствии *моделью Лайтхилла—Уизема (Уитема)—Ричардса (LWR)*, в которой поток АТС (автотранспортных средств) рассматривается как поток одномерной сжимаемой жидкости. Часто эту модель называют *моделью Лайтхилла—Уизема*. Отметим, что вместо термина «автомобиль» и тем более «машина» в транспортной литературе принято использовать термин АТС.

В модели LWR предполагается, что

- а) существует взаимно однозначная зависимость между скоростью $v(t, x)$ и погонной плотностью $\rho(t, x)$ потока — *уравнение состояния*;
- б) выполняется *закон сохранения массы* количества АТС.

¹⁾Полоса бесконечная в обе стороны, движение происходит слева направо (для определенности), нет источников и стоков автотранспортных средств.