УДК 519.87

В.И. Швецов

Институт системного анализа РАН

Проблемы моделирования передвижений в транспортных сетях

В работе дается краткий обзор последних направлений развития в области расчета корреспонденций. Эти направления связаны с переходом от матриц корреспонденций к рассмотрению цепочек передвижений, а также цепочек «дневной активности» (activity-based models) и применением методов Монте-Карло для моделирования корреспонденций. В работе также описана проблема, связанная с получением оценки межрайонной дальности, усредненной по способам передвижения их решения. Описан способ корректного усреднения. Далее обсуждаются различные алгоритмы поиска равновесия в транспортной сети. Принцип равновесия однозначно определяет значения потоков на дугах сети, однако оставляет большой произвол в выборе распределения корреспонденций по путям. Для устранения этой неопределенности может быть использован принцип максимизации энтропии.

Ключевые слова: транспортные потоки, межрайонные корреспонденции, модели активности, расщепление корреспонденций, равновесное распределение.

I. Введение

Основная задача комплексной математической модели транспортной сети — определение и прогноз всех параметров функционирования сети, таких как интенсивность движения на всех элементах сети, объемы перевозок в сети общественного транспорта, средние скорости движения, задержки и потери времени и т.д. Решение этой задачи основано на воспроизведении в рамках компьютерного моделирования всей структуры передвижений, осуществляемых в городе на всех видах транспорта в течение дня. Соответственно моделирование работы транспортной системы является комплексной задачей, включающей в себя большое количество связанных между собой задач.

В настоящее время «классической» считается схема моделирования транспортных потоков в сети крупного города, включающая четыре основных этапа (т.н. «4-х-стадийная схема»; см., например, [1]):

- оценка общих объемов прибытия и отправления из каждого района города (Trip generation);
- расщепление по способам передвижений, таким как пешие передвижения, передвижения с использованием общественного транспорта, передвижения на личном автомобиле и др. (Modal split);
- определение матриц корреспонденций, определяющих объем передвижений между каждой парой расчетных районов города (Trip distribution);
- распределение корреспонденций по транспортной сети, то есть определение всех путей, выбираемых участниками движения, и определение количества передвижений по каждому пути (Trip assignment).

В настоящей работе обсуждаются различные аспекты этой комплексной схемы. Основным недостатком этой схемы является то, что матрицы корреспонденций для передвижений, совершаемых с разными целями, рассчитываются независимо друг от друга. Более адекватные реальности модели корреспонденций должны рассматривать цепочки связанных друг с другом передвижений. В разделе II дается краткий обзор развития моделей корреспонденций в этом направлении. Развитие так называемых «моделей активности» (activity-based models) связано с переходом от матриц корреспонденций к рассмотрению цепочек передвижений, а также схем «дневной активности» и применением методов Монте–Карло для моделирования корреспонденций.

В разделе III описаны некоторые проблемы классической схемы, связанные с получением усредненных по видам транспорта оценок межрайонной дальности.

Основой для моделирования процесса распределения корреспонденций является принцип равновесия в транспортной сети. В разделе IV обсуждаются различные алгоритмы поиска равновесия. Принцип равновесия однозначно определяет значения потоков на дугах сети, однако оставляет большой произвол в выборе распределения корреспонденций по путям. Для устранения этой неопределенности может быть использован принцип максимизации энтропии.[32]

[32] Материал этого раздела несколько более подробно изложен в отдельной статье [2].

II. Модели прогноза передвижений в транспортной сети

Все многообразие передвижений, совершаемое в сети, может быть разбито на группы по цели передвижения. Наиболее важными и многочисленными являются

- передвижения от мест жительства к местам приложения труда и обратно (так называемые трудовые корреспонденции);
- передвижения от мест жительства к местам культурно-бытового обслуживания (магазинам и др.) и обратно;
- передвижения, совершаемые между местами приложений труда (деловые поездки);
- передвижения, совершаемые между объектами культурно-бытового обслуживания.

Для каждой группы передвижений рассчитывается своя матрица межрайонных корреспонденций. Входной информацией к модели расчета корреспонденций являются общие объемы прибытия и отправления в каждом районе ПО. Оценка объема прибытий и отправлений по разным группам связана с пространственным размещением объектов посещения и *подвижностью населения*, то есть средним количеством поездок, совершаемых с теми или иными целями. Эта оценка строится на основе имеющихся демографических и социально-экономических данных и результатов обследований. Расчет матриц корреспонденций обычно осуществляется с применением гравитационных или энтропийных моделей [3–5].

Матрицы корреспонденций расщепляются по способам передвижений в зависимости от соотношения обобщенных цен передвижений разными способами. Под различными способами передвижений понимают, например, передвижение пешком, с использованием общественного транспорта или личного автомобиля. Расщепление по способам обычно осуществляется с применением стандартных моделей дискретного выбора.

Изложенная принципиальная схема моделирования в настоящее время применяется на практике повсеместно. Основным недостатком этой схемы можно считать тот факт, что отдельные матрицы корреспонденций вычисляются и расщепляются по способам передвижений независимо друг от друга. В результате, например, может не совпадать количество людей, использующих автомобиль для поездки из дома на работу и обратно. Более того, общее количество передвижений «дом–работа» и «работа–дом» для каждой пары районов, вообще говоря, разное, поскольку жители обычно совершают попутные посещения магазинов и других мест по дороге домой (или на работу). Для корректного моделирования реальной структуры передвижений необходимо, таким образом, учитывать, что передвижения объединены в цепочки. Цепочка — это последовательность передвижений, которая начинается и заканчивается в одном месте (обычно дома), например, «дом–работа–магазин–дом». Как правило, способ передвижения не меняется для всех звеньев цепочки. Исключение могут составлять «промежуточные» цепочки, то есть цепочки, начинающиеся и заканчивающиеся в некотором промежуточном пункте. Например, в течение рабочего дня человек мог совершить деловую или бытовую поездку с возвращением на рабочее место.

По сложившейся терминологии модели, в которых используются отдельные матрицы корреспонденций, называются «trip-based», а модели, в которых явно рассматриваются цепочки передвижений, называются «tour-based». Среди известных практически реализованных «tour-based» моделей можно упомянуть модель Стокгольма [6] или Нью–Гемпшира [7].

Объединение отдельных передвижений в цепочке позволяет учесть влияние этих передвижений друг на друга. Например, если посещение магазина совершается из дома в рамках простейшей цепочки «дом–магазин–дом», то наиболее вероятными будут цели посещения неподалеку от дома. Если же это посещение совершается в составе цепочки «дом–работа–магазин–дом», то велика вероятность целей, находящихся на пути с работы домой, даже если они расположены далеко от дома. При таком подходе, однако, тот факт, что житель совершил попутное посещение магазина на пути с работы домой, не будет влиять на вероятность совершения дополнительной поездки в магазин после возвращения домой. Логическим шагом является рассмотрение полных «схем дневной активности» («day-pattern»), представляющих собой совокупность всех цепочек передвижений, совершаемых в течение дня. При таком подходе в качестве элементов, формирующих структуру передвижений, выступают схемы, например: 1) «дом–работа–дом»; 2) «дом–работа–магазин–дом»; 3) «дом–работа–дом; дом–магазин–дом». Каждая схема имеет свою вероятность реализации.

Модели такого типа носят общее название *моделей активности* (activity-based). Под «активностью» понимается тот или иной тип деятельности, требующий передвижения. Например, работа или посещение магазина. Одной из первых моделей активности, получившей широкое практическое применение, была модель SAMS (Sequenced Activity Mobility Simulator) [8]. Дальнейшее развитие моделей активности, происходившее в течение последнего десятилетия, было связано с попытками учета все большего числа тонких аспектов процесса формирования передвижений. Среди последних практически реализованных моделей упомянем модель Сан-Франциско [9], Нью-Йорка и ряда других городов [10], Далласа и Сиднея [11].

Математическая формулировка и программная реализация моделей активности и классической 4-стадийной модели существенно отличаются. В классической модели элементами структуры являются отдельные передвижения, которые составляют матрицы передвижений (корреспонденций). В моделях активности в качестве элементов выступают целые последовательности передвижений. Для одновременного вычисления объемов совершаемых передвижений потребовались бы многоиндексные массивы, требующие чрезмерной компьютерной памяти. Проблема еще более усугубляется при учете распределения передвижений по времени суток. В этом случае у каждого передвижения в составе каждой цепочки появляется дополнительная характеристика — время начала передвижения.

При практической реализации моделей активности вместо расчета многоиндексных массивов производится микромоделирование отдельных передвижений методом Монте-Карло. Для каждого отдельного жителя производится случайный выбор характеристик передвижений — размещения целей, времени совершения, способа передвижения. При классификации цепочек передвижений различают основную и побочные цели. Каждая цепочка передвижений фактически представляет собой передвижение к основной цели с попутным посещением побочных целей. Для упрощения процедуры выбора применяют ее *декомпозицию.* Обычно применяется декомпозиция до 3 уровней [12]:

1. Выбор схемы дневной активности.
2. Моделирование цепочек: определение основной цели, основного способа и времени начала основного передвижения для каждой цепочки.
3. Моделирование звеньев: выбор расположения промежуточной цели, выбор способа и времени начала передвижения для всех звеньев.

Одной из важных особенностей моделей активности является возможность учета согласованного поведения членов семьи [13]. Например, в семьях, имеющих автомобиль, возможны передвижения типа «подвоза» членов семьи до некоторых промежуточных пунктов. Для жителей пригородов такие поездки являются типичными. С другой стороны, в семьях с единственным автомобилем факт совершения поездки на автомобиле одним из членов семьи, естественно, влияет на вероятность совершения такой же поездки другим членом.

Классификация схем активности и определение наиболее распространенных схем требует широких социологических обследований. Основной методикой таких обследований должно быть ведение «дневников передвижений». В существующих моделях количество учитываемых схем дневной активности колеблется от нескольких десятков до нескольких сотен.

Несмотря на значительные достигнутые успехи в развитии моделей активности, их практическое применение связано с трудностями:

- Для создания моделей требуется проведение масштабных и трудоемких социологических обследований.
- Реализация моделей стоит дорого. Создание известных моделей для крупных городов за рубежом занимало в типичном случае 2–3 года и стоило 600–800 тыс. долл.
- Очень большие времена счета (сравнительно с классическими моделями) — от 0.5 до 2-х суток на современном ПК.

III. Оценка межрайонных дальностей и расщепление по видам транспорта

Ввиду того, что продвинутые модели передвижений являются в настоящее время затратными и сложными в эксплуатации, сохраняет актуальность задача разработки упрощенных моделей, позволяющих адекватно воспроизводить структуру передвижений. Такую возможность предоставляет классическая 4-стадийная схема. Здесь мы обсудим некоторые проблемы, возникающие при реализации этой схемы, а также возможные способы их преодоления.

Обычно при построении комплексной модели транспортной системы города сначала вычисляются суммарные матрицы корреспонденций, а затем они расщепляются по способам передвижений. Ключевую роль в решении обеих задач играют межрайонные *дальности*, то есть показатели, характеризующие удаленность районов друг от друга с точки зрения скорости и удобства передвижения в данной транспортной системе. Обычно в качестве меры дальности используется обобщенная цена передвижения между районами. Обобщенная цена — это взвешенный критерий, включающий в себя разные факторы, такие как время передвижения, комфорт, оплата и другие. В дальнейшем для краткости будем называть обобщенную цену передвижения просто ценой.

Очевидно, передвижения, совершаемые между фиксированной парой районов разными способами, имеют разную цену. Именно на основе сопоставления этих цен определяются коэффициенты расщепления корреспонденции по способам. Однако для расчета суммарных (не расщепленных) корреспонденций требуется в качестве меры дальности использовать цену передвижения, каким-то образом усредненную по разным способам. Далее будет изложена корректная методика такого усреднения (впервые предложенная в [14]).

Корректность методики состоит в том, что должно выполняться естественное свойство *монотонности*: мера дальности должна быть монотонно возрастающей (по крайней мере — неубывающей) функцией каждой из цен, вычисленных для отдельных способов передвижения. Другими словами, если цена передвижения одним из способов увеличивается, а цены для всех других способов не меняются, то межрайонная дальность не может уменьшиться. Она должна либо увеличиться, либо не измениться.

Далее везде, где рассматриваются цены передвижений или коэффициенты расщепления корреспонденций, подразумеваются передвижения между некоторой (произвольной) фиксированной парой районов. Индексы районов в формулах для краткости опускаются.

III.1. Условие монотонности оценок межрайонной дальности

Будем рассматривать n разных способов передвижения. Обозначим через t_1, ..., t_n цены передвижения разными способами, s_1, ..., s_n — коэффициенты расщепления: s_i — доля корреспонденции, использующая способ i. Коэффициенты расщепления предполагаются функциями цен передвижений: $s_i = s_i(t_1, ..., t_n)$. Усредненную по способам передвижения цену обозначим через t^s. Стандартная формула усреднения с учетом коэффициентов расщепления:

$$t^s = \sum_i t_i s_i(t_1, ..., t_n). \qquad (1)$$

Тогда требуемое условие монотонности имеет вид

$$\frac{\partial t^s}{\partial t_j} = s_j + \sum_i \frac{\partial s_i}{\partial t_j} t_j \geqslant 0, \quad \forall\, j. \qquad (2)$$

В корректных моделях условия (2) должны проверяться при использовании эмпирических оценок коэффициентов расщепления или специально накладываться при выводе тех или иных теоретических формул расщепления. Рассмотрим наиболее часто используемую полиномиальную (multinomial logit) модель расщепления. Модель применяется для широкого класса задач, в которых присутствует выбор из дискретного набора возможностей, для которых определена некоторая оценка «полезности» (или, наоборот, «затрат»), связанных с выбором этих возможностей. Согласно этой модели вероятность выбора одной из возможностей убывает пропорционально экспоненте от роста затрат:

$$s_i = \frac{e^{-U_i}}{\sum_j e^{-U_j}}, \quad \forall\, i, \qquad (3)$$

где U_i — функция затрат при выборе i-й возможности. В задаче расщепления передвижений по способам естественно считать затраты связанными с ценой передвижения:

$$U_1 = \alpha_i t_i + \beta_i. \qquad (4)$$

Здесь коэффициенты α_i,β_i — это калибровочные параметры. Параметр α_i характеризует чувствительность населения к фактору цены при передвижениях i-м способом. Параметры β_i характеризуют асимметрию между способами передвижений. Коэффициенты β_i определены с точностью до прибавления общей константы, поэтому всегда можно положить самый маленький коэффициент равным 0.

Условие монотонности (2) с учетом (3)–(4) принимает вид

$$\sum_i (1 + \alpha_i(t_i - t_j))e^{-U_i} \geqslant 0, \quad \forall\, j. \qquad (5)$$

Ясно, что это условие выполняется не при всех значениях цен передвижений. В частности, условие нарушается, когда одна из цен существенно превосходит другие. В общем случае можно констатировать, что монотонность всегда будет нарушаться при больших t_i, если коэффициент расщепления s_i слишком быстро убывает с ростом t_i. В этом случае соответствующее слагаемое в сумме (1) убывает, когда t_i растет.

III.2. Метод вывода корректных формул оценки дальности

Предложенная в [14] методика оценки межрайонной дальности основана на следующих простых соображениях. Рассмотрим задачу определения цены передвижения одним из способов, например, на легковом автомобиле. Теоретически из одного района в другой можно проехать большим числом путей за разное время. При этом в качестве меры дальности естественно выбирается время (обобщенная цена) самого короткого пути, никакого усреднения по всем альтернативным путям не производится. Подобный подход можно было бы применить и при сравнении разных способов передвижения, то есть не усреднять цены для разных способов, а просто выбрать минимальную из них. Заметим, что такая оценка естественно удовлетворяет свойству монотонности, поскольку функция минимума из нескольких переменных всегда монотонна по каждому аргументу.

Почему же возникает необходимость в усреднении? Как представляется, причина состоит в том, что сопоставление разных способов передвижений не сводится к простому сравнению времен. Здесь учитываются разные дополнительные факторы — комфорт, денежные затраты, просто наличие или отсутствие автомобиля. Количественное влияние этих факторов различается для разных участников движения, поэтому возникает необходимость в усреднении разных индивидуальных оценок.

Предлагаемая методика основана на достаточно прямолинейной реализации изложенных выше соображений. Предположим, что каждый участник движения при сравнении способов передвижения использует индивидуальную оценку дальности передвижения каждым возможным способом:

$$\tilde{t}_i = t_i + a_i, \tag{6}$$

где t_i — это «объективная» оценка цены передвижения, общая для всех участников, а a_i — это субъективное «штрафное» время, которое добавляется к цене передвижения i-м способом при сравнении разных способов. Коэффициенты a_i являются, по существу, поведенческими характеристиками участников движения, причем эти характеристики можно считать случайными величинами. Индивидуальная оценка межрайонной дальности каждым участником движения дается формулой

$$t^s(a_1, ..., a_n) = \min(t_1 + a_1, ..., t_n + a_n). \tag{7}$$

Оценку (7) можно назвать *микроскопической* оценкой межрайонной дальности, поскольку она описывает поведение отдельного элемента системы (участника движения). Соответственно *макроскопическая* оценка может быть получена путем усреднения микроскопической оценки по всем участникам движения. Для усреднения необходимы дополнительные предположения о виде распределения значений поведенческих параметров a_i. Пусть $\rho(a_1, ..., a_n)$ — известная плотность этого распределения. Тогда макроскопическая оценка дальности дается интегралом

$$t^s = \int \rho(a_1, ..., a_n) t^s(a_1, ..., a_n) da_1 ... da_n. \tag{8}$$

Предложенный подход может быть развит и конкретизирован в двух основных направлениях. Во-первых, вместо (6) можно применить более сложную методику вычисления индивидуальной оценки дальности: $\tilde{t}_i = \tilde{t}_i(t_i, a_i)$, где a_i не обязательно является аддитивной добавкой к цене, но играет роль поведенческого параметра, так или иначе влияющего на субъективную оценку передвижений i-м способом. Во-вторых, можно сделать разные предположения о характере распределения значений поведенческих параметров среди населения.

III.3. Формулы оценки дальности с учетом двух видов транспорта

В качестве примера выведем явные формулы для оценки межрайонных дальностей с учетом двух возможных способов передвижения: поездок на легковом автомобиле и на общественном транспорте. Соответственно используется единственный поведенческий параметр a, влияющий на сравнение цен поездок на этих видах транспорта. Будем обозначать через t_{car} цену движения на легковом автомобиле, t_{pub} — цену движения на общественном транспорте. В соответствии с (7) индивидуальная оценка межрайонной дальности имеет вид

$$t^s(a) = \min(t_{car} + a, t_{pub}), \tag{9}$$

где параметр a имеет смысл добавочного штрафного времени, которое добавляется к цене движения на автомобиле для сопоставления с ценой движения в общественном транспорте. Отметим, что параметр a может принимать как положительные, так и отрицательные значения.

Обозначим плотность распределения параметра a через $\rho(a)$, соответствующую функцию распределения — через $F(a) = \int\limits_{-\infty}^{a} \rho(a')da'$. Также будем обозначать разность цен на двух видах транспорта через $\Delta t = t_{pub} - t_{car}$. Тогда формула (8) для межрайонной дальности будет иметь вид

$$t^s = t_{pub} - F(\Delta t)\Delta t + \int\limits_{-\infty}^{\Delta t} \rho(a) a da. \tag{10}$$

Для дальнейших вычислений необходимы предположения о виде распределения значений параметра a. Рассмотрим здесь простой случай, когда параметр имеет нормальное (Гауссово) распределение:

$$\rho(a) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(a-\mu)^2}{2\sigma^2}}, \tag{11}$$

где μ — математическое ожидание, а σ — дисперсия случайной величины a. С учетом (11) выражение для межрайонной дальности (10) окончательно примет вид

$$t^s = t_{pub} - F(\Delta t)(\Delta t - \mu) + \frac{\sigma}{\sqrt{2\pi}} e^{-\frac{(\Delta t-\mu)^2}{2\sigma^2}}, \tag{12}$$

где $F(\Delta t)$ — функция нормального распределения:

$$F(\Delta t) = \frac{1}{\sqrt{2\pi}\sigma} \int\limits_{-\infty}^{\Delta t} e^{-\frac{(a-\mu)^2}{2\sigma^2}} da. \tag{13}$$

В предыдущих формулах предполагалось, что параметр a может иметь отрицательные значения. Отрицательное значение a соответствует ситуации, когда участник движения оценивает поездку на легковом транспорте как предпочтительную, даже если t_{car} равно или немного превосходит t_{pub}. При этом индивидуальная оценка дальности может для некоторых пар районов принимать отрицательные значения (а именно, возможен случай $t_{car} + a < 0 < t_{pub}$). Этот эффект можно признать нежелательным с интуитивной точки зрения, поскольку межрайонная дальность обычно мыслится как существенно положительная величина (впрочем, такой подход имеет право на существование, если воспринимать дальность как абстрактный критерий, который, в частности, может включать в себя отрицательные «штрафные» слагаемые).

Можно обеспечить положительность оценки дальности, если внести небольшую модификацию в формулу для индивидуальной оценки (9). А именно, в случае отрицательного значения a можно не уменьшать цену поездки на легковом транспорте, а вместо этого увеличить на ту же величину цену поездки на общественном транспорте, т. е. вместо (9) используется формула

$$t^s(a) = \begin{cases} \min(t_{car} + a, t_{pub}), & a \geqslant 0, \\ \min(t_{car}, t_{pub} - a), & a < 0. \end{cases} \tag{14}$$

Можно показать, что формула средней дальности (10) при этом заменяется на следующую:

$$t^s = t_{pub} - F(\Delta t)\Delta t + \int\limits_0^{\Delta t} \rho(a) a da, \tag{15}$$

причем в случае $\Delta t < 0$ следует изменить знак перед интегралом и поменять местами пределы интегрирования в правой части (15). Поскольку модифицированная индивидуальная оценка (14) удовлетворяет требованию монотонности, это свойство сохраняется и для средней дальности (15) (это можно также проверить прямым дифференцированием).

В предположении, что параметр a имеет нормальное распределение, интегрирование в (15) дает следующее выражение для средней дальности:

$$t^s = t_{pub} - F(\Delta t)(\Delta t - \mu) - F(0)\mu +$$
$$+ \frac{\sigma}{\sqrt{2\pi}} \left(e^{-\frac{-\mu^2}{2\sigma^2}} - e^{-\frac{(\Delta t - \mu)^2}{2\sigma^2}} \right).$$

IV. Принцип равновесия

В основе алгоритмов распределения транспортных потоков по путям в сети лежит предположение о том, что каждый участник движения стремится минимизировать обобщенную цену своего пути. Цена пути представляет собой сумму цен движения по дугам, из которых состоит путь, а также цен переходов с дуги на дугу (маневров на пересечениях). В дальнейшем для сокращения формул будем опускать цены маневров и говорить только о ценах дуг, которые будем обозначать через c_a, где a — номер дуги.

Равновесным называется распределение потоков, при котором ни один из участников движения не может изменить свой путь так, чтобы уменьшить цену пути. Согласно современным представлениям распределение потоков в транспортной сети в установившейся ситуации является равновесным в указанном смысле. Сформулированное условие известно как условие Вардрупа (или условия Вардрупа, поскольку оно часто формулируется в виде двух предложений [14]).

Ключевым предположением о цене является следующее: цена движения по дуге является неубывающей функцией суммарного потока по этой дуге: $c_a = c_a(u_a)$. Другими словами, чем больше автомобилей движется по дуге, тем более медленным и дискомфортным становится движение и соответственно возрастает цена пути. При данном предположении можно показать, что в любой транспортной системе существует равновесное состояние. Это состояние является единственным в том смысле, что условие равновесия однозначно определяет потоки на всех дугах сети. Другими словами, существует единственная загрузка элементов сети, удовлетворяющая условию равновесия. Однако равновесное распределение не является единственным в терминах распределения корреспонденций по путям в сети. Т. е. разные распределения корреспонденций по путям могут порождать одинаковую загрузку всех элементов сети (на самом деле, для данной загрузки существует бесконечное количество порождающих ее распределений).

Равновесное распределение дается решением следующей задачи оптимизации с ограничениями [15]:

$$\sum_a \int\limits_0^{u_a} c_a(v) dv \to \min. \tag{16}$$

Потоки на дугах u_a, входящие в это выражение, не являются независимыми переменными задачи. Они представляют собой суммы потоков u_{kpq} по всем путям k, использующим данную дугу a. Ограничения в задаче состоят в том, что потоки по путям u_{kpq} должны в сумме давать корреспонденции F_{pq} между всеми парами районов p и q:

$$\begin{aligned} u_a &= \sum_{p,q} \sum_{k \in K_{pq}, a \in k} u_{kpq}, \\ F_{pq} &= \sum_{k \in K_{pq}} u_{kpq}, \\ u_{kpq} &\geqslant 0. \end{aligned} \tag{17}$$

IV.1. Алгоритмы вычисления равновесного распределения

Хотя принцип равновесия в транспортной сети был впервые сформулирован еще в 50-е гг., практическое применение этого принципа началось значительно позже, в середине 70-х гг. [16]. Это связано с тем, что численная реализация модели требует значительного объема памяти компьютера. Для хранения всей информации о распределении корреспонденций по путям в крупном городе требуется значительный объем (до нескольких гигабайт) памяти. Компьютеры с такими возможностями стали широко распространены только в последнее время. Однако возможность моделировать равновесное распределение появилась значительно ранее, благодаря разработке алгоритма Франке–Вульфа [17]. Особенностью этого алгоритма является то, что он дает возможность вычислить итоговую равновесную загрузку транспортной сети, не сохраняя в памяти компьютера само распределение корреспонденций по путям.

Схема работы алгоритма Франке–Вульфа состоит в следующем. На начальном шаге все корреспонденции распределяются на оптимальные (по цене) пути, рассчитанные по свободной сети. На каждом последующем шаге алгоритма $n+1$ каждая корреспонденция уже распределена на некоторое количество путей (не более n), вычисленных на предыдущих n шагах. В результате возникает загрузка дуг сети u_a^n. С учетом этой загрузки рассчитываются цены дуг c_a^n и строится новая система оптимальных путей. Данная система путей является «опорной» для перераспределения корреспонденций. При этом для некоторых корреспонденций рассчитанные пути могут оказаться действительно новыми, в то время как для других корреспонденций рассчитанный оптимальный путь уже содержится среди путей, полученных на предыдущих шагах. Далее вычисляется доля корреспонденций λ, которая будет перераспределена со всех прежних путей на новые оптимальные пути (способ вычисления λ для последующего анализа несущественен).

Особенностью алгоритма является то, что со всех прежних путей перераспределяется одна и та же доля корреспонденций λ. Это означает, что загрузка всех дуг уменьшится ровно в $(1-\lambda)$ раз. Отсюда следует, что для вычисления новой загрузки всех дуг нет необходимости запоминать распределение по путям, а достаточно сохранять в памяти компьютера только загрузку всех дуг от предыдущего шага. С учетом того, что цены дуг зависят только от загрузки дуг $c_a = c_a(u_a)$, для вычисления самой доли λ также не требуется знать распределение по путям.

Описанное достоинство алгоритма Франке–Вульфа объясняет его «долгожительство». Этот алгоритм до сих пор является наиболее распространенным в практике моделирования загрузки автомобильных сетей. Однако алгоритм обладает существенным недостатком. Хотя теоретически алгоритм всегда сходится, на практике скорость сходимости существенно ухудшается в ходе итераций. В течение первых итераций наблюдается достаточно быстрое приближение к равновесию, однако после нескольких десятков итераций приближение может практически прекратиться. Особенно сильно этот эффект проявляется на больших сетях. Ухудшение сходимости тесно связано с так называемым эффектом «застревающих потоков» (residual flows). Эффект проявляется в сильной неравномерности сходимости потоков к равновесным значениям на отдельных дугах, т. е. при общей хорошей сходимости к равновесию в сети может сохраняться небольшое количество дуг, на которых поток сильно отличается от равновесного, причем эти «выбросы» не устраняются последующими итерациями.

Причину эффекта «застревающих потоков» можно понять исходя из принципа работы алгоритма, показанного на рис. 1. В итоге работы алгоритма каждая корреспонденция должна быть распределена по набору альтернативных путей в таких пропорциях, чтобы цены всех путей в наборе в точности уравнялись. Если минимальная и максимальная цены путей в наборе мало отличаются, то для приближения к равновесию значение λ нужно выбирать малым. Большее значение λ только отдалит распределение данной корреспонденции от равновесия. Предположим, что на ранних стадиях алгоритма в силу случайных причин небольшое количество корреспонденций «отстало» от основной массы корреспонденций по сходимости к равновесию. Для этих корреспонденций требуется перераспределить существенно бо̀льшую долю λ. Поскольку значение λ выбирается как среднее (в некотором смысле) для всех корреспонденций, оно будет выбрано малым. В результате основная часть корреспонденций еще более приблизится к равновесию, в то время как «отставание» небольшой их части сохранится. При последующих итерациях значение λ, диктуемое основной частью корреспонденций, будет становиться все меньше и меньше, так что отдельные «застрявшие» корреспонденции так и останутся неперераспределенными.

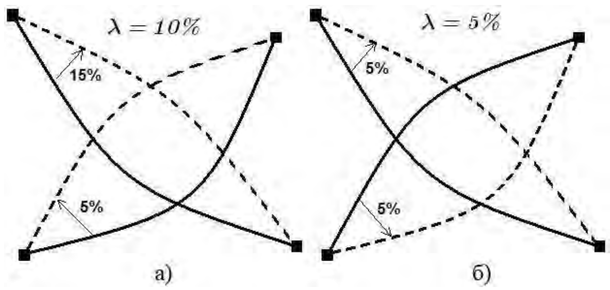

Рис. 1. Принцип работы алгоритма Франке–Вульфа

В последние годы были разработаны новые алгоритмы поиска равновесного распределения, свободные от указанных выше недостатков алгоритма Франке–Вульфа [18, 19]. Основная идея состоит в индивидуальной работе с корреспонденциями, что позволяет своевременно перераспределять отдельные «застрявшие потоки». Разумеется, перераспределение одной корреспонденции среди альтернативных путей изменяет цены дуг и косвенно влияет на распределение других корреспонденций, использующих те же дуги. Отсюда возникает необходимость в многократном проходе всего массива корреспонденций для приведения системы в равновесное состояние.

Опишем простой алгоритм такого типа, использованный для моделирования Московской агломерации. Алгоритм может быть назван «балансировкой по путям». Основная идея алгоритма состоит в следующем. В равновесном состоянии каждая корреспонденция распределена среди некоторого количества альтернативных путей, причем это количество может быть разным для разных корреспонденций и заранее не известно. Если фиксировать некоторые наборы путей между всеми парами районов, то можно сформулировать задачу поиска равновесия «в узком смысле»: найти равновесие в системе при дополнительном условии, что для передвижения используются только пути из фиксированных наборов, и никакие другие. Очевидно, что если в эти наборы включены все пути, используемые при настоящем равновесном распределении, то равновесие «в узком смысле» совпадает с равновесием в обычном смысле.

Будем называть фиксированный набор путей между парой районов «пучком» путей. На первом шаге алгоритма можно сформировать пучки, состоящие всего из одного пути — оптимального пути между соответствующей парой районов, рассчитанного по пустой сети. На каждом шаге алгоритма пучки пополняются новыми путями следующим образом. Сначала решается задача равновесия «в узком смысле» для имеющихся пучков (на первом шаге эта задача тривиальна). После этого каждый пучок пополняется оптимальным путем, рассчитанным при сложившейся загрузке дуг сети. Разумеется, в качестве «нового» может выступать один из ранее полученных путей. Фактически алгоритм закончит свою работу, когда ни один по-настоящему новый путь не будет найден. Отметим, что описанный алгоритм находит равновесное состояние за конечное число шагов, хотя оно, разумеется, не будет точным, поскольку балансировка пучков на каждом шаге ведется до достижения некоторой конечной точности.

Согласно установившейся в последние годы терминологии алгоритмы поиска равновесия в транспортной сети можно разделить на «дуговые» (link-based), то есть основанные на работе с потоками по дугам, и «маршрутные» (route-based), работающие с потоками по всем путям (или маршрутам) в сети. Типичным примером дугового алгоритма является алгоритм Франке–Вульфа, маршрутного типа — изложенный выше алгоритм балансировки по путям. Основным преимуществом маршрутных алгоритмов является существенное повышение равномерности сходимости потоков на дугах, в частности, решение проблемы «застревающих потоков». Однако это достигается за счет кардинального увеличения необходимого объема компьютерной памяти.

В этих условиях заслуживает внимания промежуточный подход, предложенный Bar–Gera [19]. Этот подход представляет собой нечто среднее между дуговым и маршрутным подходами. Основной идеей является различение на каждой дуге представителей корреспонденций, двигающихся из общего района отправления, без различения районов прибытия. Таким образом, вместо потоков на дугах u_a или путях u_{kpq} в качестве независимых переменных используются потоки из «общего источника» (origin-based) u_{ap}, где a — номер дуги, p — номер района отправления. Очевидно, этот массив существенно (в количество раз порядка числа районов) меньше массива потоков по путям и может помещаться в оперативную память. Вместе с тем он позволяет достаточно индивидуально работать с корреспонденциями от разных районов. Как показали численные эксперименты, алгоритм обеспечивает требуемую равномерность сходимости и решает проблему «застрявших потоков».

IV.2. Неединственность равновесного распределения и энтропия

Как уже было сказано выше, условие Вардрупа однозначно определяет значения суммарных потоков на всех дугах сети u_a. Это следует из вида целевой функции (16). Действительно, если цены дуг $c_a(u_a)$ являются возрастающими функциями потоков, то интегралы в (16) — выпуклые функции от u_a, поэтому равновесное распределение существует и единственно как точка минимума выпуклой функции при линейных ограничениях. Однако равновесное распределение не является единственным в терминах распределения корреспонденций по путям в сети. Причина этой неединственности наглядно показана на простом примере на рис. 2.

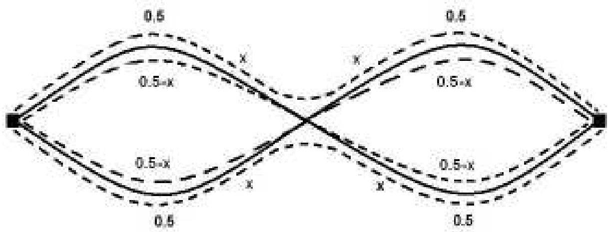

Рис. 2. Неединственность равновесного распределения

Пусть два района A и B соединены двумя парами идентичных путей с точкой пересечения посередине, как показано на рис. 2. Распределяется единственная корреспонденция из района A в район B. Ясно, что в условиях равновесия загрузки всех показанных дорог должны быть равны, то есть в обеих точках развилки корреспонденция должна делиться пополам. Существует ровно 4 разных пути из A в B. При этом существует бесконечное число способов распределения корреспонденции по этим путям, при которых корреспонденция на обеих развилках делится пополам. А именно, пусть x — это доля водителей, избравших на обеих развилках левый путь. Как показано на рисунке, при любом значении $0 \leqslant x \leqslant 0{,}5$ мы получаем распределение по четырем путям, порождающее равновесную загрузку. Различие этих распределений состоит только в том, сколько водителей из числа выбравших левый путь на первой развилке примет такое же решение на второй развилке.

В реальной сети с большим количеством районов ситуация более запутанная, однако природа неоднозначности распределения по путям остается такой же, как в этом примере. В сети присутствует большое количество подобных развилок, при этом с точки зрения принципа равновесия важны только общие объемы потоков по дугам. При соблюдении нужных общих объемов совершенно не важно, из каких именно водителей составляются эти потоки.

Таким образом, приходим к заключению, что условие Вардрупа не определяет однозначно распределение корреспонденций по путям. На самом деле, каждому возможному набору значений потоков по дугам u_a соответствует многомерное линейное пространство распределений по путям u_{kpq}. Следовательно, для корректной постановки задачи поиска равновесного распределения требуется сформулировать какие-то предположения, дополнительные к условию Вардрупа.

Представляется естественным вероятностный подход к решению вопроса. Рассмотрим какую-либо пару районов p,q, соединенных пучком путей K_{pq}. Можно предположить, что водители (представители корреспонденции f_{pq}) выбирают один из путей $k \in K_{pq}$ случайно и независимо друг от друга. Если задать значения вероятности выбора того или иного пути, то несложно вычислить вероятность реализации того или иного распределения корреспонденции f_{pq} по путям. В частности, в условиях равновесия цены всех используемых путей равны, и поэтому можно принять, что каждый водитель выбирает любой из используемых путей с одинаковой вероятностью.

В терминах макросистемного подхода [20] эту ситуацию можно описать так. Водители являются элементами макросистемы, при этом выбор того или иного пути определяет состояние элемента. Микросостояние системы определяется совокупностью состояний элементов, то есть указанием пути для каждого водителя индивидуально. Макросостояние определяется «числами заполнения», то есть общим количеством элементов в том или ином состоянии. В нашем случае числа заполнения — это потоки на путях u_{kpq}. Таким образом, разные распределения корреспонденций по путям соответствуют разным макросостояниям системы. Каждое макросостояние может быть реализовано различным количеством микросостояний, что определяет его «статистический вес». В нашем случае, когда число элементов и число состояний конечны, статистические веса макросостояний просто пропорциональны вероятностям их реализации.

Если сеть находится в равновесии и известен набор равновесных значений потоков на дугах u_a^*, то это накладывает линейные ограничения на возможные значения u_{kpq} (см. ограничения (17)). Предполагая, что в реальности должно осуществляться состояние с наибольшим статистическим весом, приходим к модели максимизации энтропии (логарифма статистического веса) системы. Энтропийная модель равновесного распределения потоков может быть сформулирована в следующем виде [21]:

$$\sum_{p,q} \sum_{k \in K_{pq}} \left(u_{kpq} \ln \left(\frac{u_{kpq}}{f_{pq}} \right) - u_{kpq} \right) \to \max \quad (18)$$

при ограничениях

$$\begin{aligned} u_a^* &= \sum_{p,q} \sum_{k \in K_{pq}, a \in k} u_{kpq}, \\ f_{pq} &= \sum_{k \in K_{pq}} u_{kpq}, \\ u_{kpq} &\geqslant 0. \end{aligned} \quad (19)$$

В связи с изложенным представляется интересным сравнить распределения по путям, возникающие при применении алгоритма Франке–Вульфа и маршрутных алгоритмов. Работа этих алгоритмов схематично показана на рис. 3. Пространство возможных наборов значений потоков на дугах u_a представлено в виде одномерной линии. В этом пространстве имеется единственная точка u_a^*, соответствующая равновеснам потокам. При этом пространство возможных распределений по путям u_{kpq} оказывается расслоенным на подпространства, каждое из которых соответствует некоторому фиксированному набору u_a. В каждом из этих подпространств имеется наиболее вероятное (максимизирующее энтропию) распределение u_{kpq}^*. Алгоритмы поиска равновесия обычно начинают работу с распределения корреспонденций по оптимальным путям, вычисленным по свободной сети. Далее осуществляется итеративное перераспределение корреспонденций между путями. Логика этого перераспределения основана на стремлении привести точку u_a к точке равновесия u_a^*. При этом совершенно не учитывается степень близости u_{kpq} к u_{kpq}^*. Вследствие этого отклонение итогового распределения по путям от наиболее вероятного для всех этих алгоритмов выглядит произвольным, и сделать теоретически обоснованные выводы трудно.

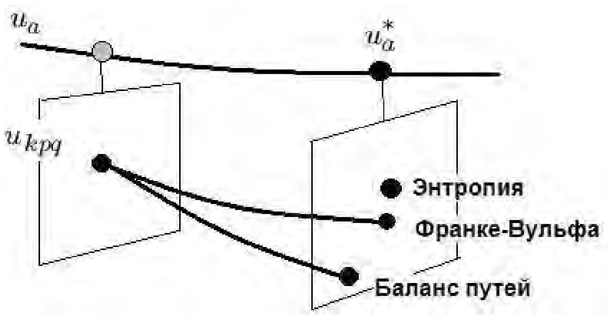

Рис. 3. Сопоставление разных алгоритмов поиска равновесия

Однако численные эксперименты свидетельствуют о том, что алгоритм Франке–Вульфа часто дает распределение по путям, более близкое к u_{kpq}^*, чем маршрутный алгоритм. В частности, в маршрутном алгоритме резко снижается расщепление корреспонденций по разным путям, то есть алгоритм стремится минимизировать число разных путей. Такие же результаты получены в [22] при сравнении алгоритма Франке–Вульфа с алгоритмом, использующим потоки из общего источника. Возникает вывод, звучащий, на первый взгляд, несколько парадоксально: маршрутные алгоритмы по сравнению с дуговыми обеспечивают лучшую сходимость на дугах, но имеют тенденцию приводить к «худшему» распределению по маршрутам (путям).

V. Заключение

В работе дан краткий обзор моделей корреспонденций, основанных на понятии «активности». Модели этого типа могут обеспечить более реалистичное воспроизведение структуры передвижений в городах, чем классическая 4-стадийная схема. Однако применение моделей такого типа ограничивается сложностью реализации, дороговизной и слишком большими временами счета. Поэтому сохраняет актуальность задача построения упрощенных моделей передвижений, в особенности на ранних стадиях проектирования.

В работе рассмотрена также задача корректного формирования оценки межрайонной дальности на основе усреднения цен передвижений, совершаемых разными способами (разными видами транспорта или пешком). Изложен метод вывода формул для такого усреднения, который обеспечивает автоматическое выполнение естественного свойства монотонности оценки как функции цен передвижений разными способами. В рамках данного подхода могут быть получены формулы для вычисления межрайонной дальности существенно разного вида. Для практического применения результатов требуется дальнейшее изучение свойств получаемых оценок и сопоставление теоретических функций с имеющимися статистическими данными.

Также рассмотрено применение принципа равновесия для моделирования распределения корреспонденций по транспортной сети. В задаче моделирования транспортных потоков можно выделить два уровня: расчет потоков на дугах сети, то есть расчет загрузки сети, а также расчет распределения всех корреспонденций по путям в сети, которое порождает эту загрузку. Принцип равновесия однозначно определяет потоки на дугах сети, но оставляет большой произвол в выборе распределения по путям. Для устранения этой неопределенности принцип равновесия можно дополнить принципом максимизации энтропии, позволяющим рассчитать наиболее вероятное распределение по путям при известных потоках на дугах, хотя такой подход сопряжен со значительными вычислительными трудностями в связи с большой размерностью задач.

Литература

1. *Швецов В.И.* Математическое моделирование транспортных потоков // Автоматика и Телемеханика. — 2003. — № 11. — С. 3–46.

2. *Швецов В.И.* Алгоритмы распределения транспортных потоков // Автоматика и Телемеханика. — 2009. — № 10. — С. 148–157.

3. *Carrothers G.A.P.* An historical review of the gravity and potential concepts of human interaction // Journal of the American Institute of Planners. — 1956. — V. 22. — P. 94–102.

4. *Wilson A.G.* A statistical theory of spatial distribution models // Transpn. Res. — 1967. — V. 1. — P. 253–270.

5. *Wilson A.G.* A family of spatial interaction models and associated developments // Envir. & Plan. A. — 1971. — V. 3. — P. 255–282.

6. *Algers S., Daly A., Kjellman P., Widlert S.* Stockholm model system (sims): Application // 7th World Conference of Transportation Research. — Sydney, Australia, 1995.

7. *Rossi T., Shiftan Y.* Tour based travel demand modelling in the u.s. // Proceeding of the 8th IFAC /IFIP/ IFORS Symposium on Transportation Systems. — Chania, Greece, 1997.

8. *Kitamura R., Pas E.I., Lula C.V., Lawton T.K., Benson P.E.* The sequenced activity mobility simulator (sams): An integrated approach to modelling transportation, land use and air quality // Transportation. — 1996. — V. 23(3). — P. 267–291.

9. *Jonnalagadda N., Freedman J., Davidson W.A., Hunt J.D.* Development of a microsimulation activity-based model for san francisco – destination and mode choice models // 81st Annual Meeting of the Transportation Research Board. — Washington, 2002.

10. *M. Bradley, Vovsha P.* A model for joint choice of daily activity pattern types of household members // Transportation. — 2005. — V. 32. — P. 545–571.

11. *Arentze T., Timmermans H.J.P.* A co-evaluation approach to extracting and predicting linked sets of complex decisions rules from activity diary data // 80Th Annual Meeting of the Transportation Research Board. — Washington, D.C., 2001.

12. *Shiftan Y., Ben–Akiva M.* A practical policy sensitive activity-based model // 11th International Conference on Travel Behaviour Research. — Kyoto, 2006.

13. *Zhang J., Timmermans H.J.P., Borgers A.* A model of household task allocation and time use // Transpn. Res. B. — 2005. — V. 39. — P. 91–95.

14. *Котельникова А.Г., Швецов В.И.* Оценка межрайонных дальностей в транспортных моделях // Сб. Трудов ИСА РАН «Прикладные про-