

ОСНОВНЫЕ ЭТАПЫ ПРОЦЕССА ПРОГНОЗИРОВАНИЯ И ПЛАНИРОВАНИЯ ТРАНСПОРТНЫХ ПОТОКОВ

Г. А. Омарова

Институт вычислительной математики и математической геофизики СО РАН,
630090, Новосибирск, Россия

УДК 519.179.2+512.23

Рассмотрены основные этапы процесса прогнозирования и планирования транспортных потоков: моделирование поездки, выбор места назначения, выбор способа передвижения и выбор маршрута. Приведен пример построения логит-модели основных этапов моделирования.

Ключевые слова: модель, вероятность, условная вероятность, логит-модель, принцип равновесия, сохранение потока, случайная величина.

Article is devoted to detailed consideration of the main stages of process of forecasting and planning of transport streams: trip modeling, destination choice, choice of a way and route choice. It is given an example constructions logit models in two cases.

Key words: Model, probability, conditional probability, logit model, equilibrium assignment, conservation of flow, random variable.

Введение. В большинстве случаев процесс принятия решения индивидуальным путешественником связан с различными вероятностями, которые играют центральную роль в прогнозировании и планировании транспортировки.

Цель данной работы — прогнозирование количества транспортных потоков в заданной области. Область делится на зоны в зависимости от плотности транспортных потоков. Внутри зоны могут быть выделены участки с минимальным количеством дорог. Для решения задачи проводится ее разбиение на подзадачи: моделирование поездки, выбор места назначения, выбор способа и выбор маршрута. Данное разбиение связано с вероятностным выбором путешественника перед началом поездки.

В связи с этим проведем следующую структуризацию выбора на каждом шаге [1]:

1. Даны географическое положение i в данный момент времени (час, день и т. д.), вид деятельности (покупки, работа, отдых и т. д.), первое индивидуальное решение: путешествовать или нет. Безусловная вероятность, что путешественник совершает путешествие в данный период времени, равна P_i , не совершает — $(1 - P_i)$.

2. Данный выбор осуществляется на первом уровне принятия решения и зависит от местонахождения путешественника i в момент выбора места назначения j для проведения данной деятельности. Условная вероятность этого решения равна $P_{i/j}$.

3. Дан результат первых двух решений, когда путешественник выбирает, какие виды транспорта m целесообразно использовать среди различных доступных альтернативных способов перемещения между пунктами i и j . Условная вероятность этого решения равна $P_{m/ij}$.

4. Дан результат всех предыдущих решений. Путешественник принимает решение о выборе пути r среди всех возможных для путешествия. Условная вероятность этого решения равна $P_{r/ijm}$.

Согласно данной структуре моделирование поездки — первый из четырех этапов процесса прогнозирования транспортировки [2]. Это позволяет предсказывать число поездок, происходящих в пределах данной зоны или проходящих через специфическую зону анализа трафика. Каждая поездка имеет начальный и конечный пункты, которые должны быть известны. На первом этапе моделирования определяется, сколько поездок происходит внутри зоны (начало и конец поездки находятся в одной зоне), второй этап моделирования — то, сколько поездок предполагалось между зонами (начало и конец поездки в разных зонах), так как участки могут быть разделены на две категории: населенный участок и участок, не связанный с постоянным проживанием людей. Следующие основные этапы — выбор места назначения, выбор способа передвижения и выбор маршрута, широко используемые для оптимизации путешествия.

1. Определение моделей. При построении модели возникают различные вопросы, например: как предсказать количество поездок; какое количество поездок, начинающихся в зоне или планируемых в зону, описано тарифами поездки или уравнениями? Сначала необходимо идентифицировать значимые переменные [1, 2].

1.1. Виды деятельности, связанные с поездкой. Люди участвуют в различных действиях, которые являются “целью” поездки или обслуживанием пассажиров. Главные виды деятельности — возвращение домой, поездка на работу, в магазин, на учебу, питание вне дома, отдых. Существуют также другие виды деятельности, совершаемые людьми ежедневно или еженедельно, такие как визит к доктору, в банк и т. д. Менее частые категории включены в категорию “Другое”. Каждая поездка имеет две крайние точки: начало и завершение. Поездки разделены на следующие категории: цель, деятельность, место назначения.

Люди участвуют в деятельности последовательно, совершая при этом цепочку поездок.

Начало и конец поездки. Общее количество поездок домой может быть представлено в виде функции $T_h = f_h(x_1, x_2, x_3, x_4, x_5, x_6, x_7)$, где x_1 — единица измерения; x_2 — домашний адрес; x_3 — возраст; x_4 — доход; x_5 — доступность; x_6 — владелец транспортного средства; x_7 — другие элементы.

В конце рабочего дня число поездок с работы — это число сгенерированных поездок, которые можно описать функцией $T_w = f_w(x_1, x_2, x_3, x_4)$, где x_1 — вид работы; x_2 — область рабочего пространства; x_3 — род занятий; x_4 — другие элементы, связанные с работой.

Поход по магазинам. Аналогично поход по магазинам зависит от различных факторов, поэтому его можно представить в виде функции $T_s = f_s(x_1, x_2, x_3, x_4, x_5, x_6)$, где x_1 — число мест продажи вне магазина; x_2 — тип доступной розничной продажи; x_3 — область доступной розничной продажи; x_4 — местоположение; x_5 — конкуренция; x_6 — другие элементы, связанные с розничной продажей.

1.2. Входные данные и оценка модели. Прогнозирование деятельности, выполняемое планировщиками или экономистами на основе экономического анализа, позволяет получить данные о составляющих развития деятельности населения. Прогноз использования участков влияет на изменение прогнозирования трафика внутри зоны.

При оценке модели возникает проблема выбора средних чисел, которая состоит в индивидуальности каждого человека, совершающего поездку.

Начало и конец поездки. Чтобы оценить качество моделирования поездки домой, можно использовать модель поперечной классификации, в основе которой лежит построение таблицы, строки и столбцы которой содержат различные характеристики, а каждая ячейка показывает прогнозируемое число поездок, полученное непосредственно из эмпирических данных. Например, в модели поперечной классификации зависимая переменная — количе-

ство поездок одного человека, а независимые переменные — место регистрации, фактическое место проживания, возраст человека.

Поездка, не являющаяся возвращением домой. Оценки моделирования поездки на “работу” и поездки “с другими целями”, имеющие отношение к поездке, роду занятий и другим характеристикам путешествия, могут быть получены с помощью метода наименьших квадратов регрессионного анализа [3, 4].

При оценке норм поездки в конце рабочего дня используются следующие переменные: занятость в офисах E_{off} , розничная продажа E_{ret} , прочее E_{oth} .

Типичное уравнение моделирования поездки на “работу” и поездки “с другими целями” можно записать в виде

$$T_i = a_1 E_{off,i} + a_2 E_{oth,i} + a_3 E_{ret,i},$$

где T_i — количество поездок человека, работающего в зоне i ; $E_{off,i}$ — занятость в офисах в зоне i ; $E_{oth,i}$ — другая занятость в зоне i ; $E_{ret,i}$ — продажа в розницу и занятость в зоне i ; a_1, a_2, a_3 — коэффициенты.

1.3. *Нормализация.* Независимо от целей поездки (например, домой, на работу, на учебу) число поездок из дома должно быть равно числу возвращений домой. Две различные модели могут дать разные результаты. Можно либо предположить, что одна из этих моделей правильная, и корректировать другую, либо использовать сочетание моделей. Необходимо гарантировать, что общее количество начальных поездок равно общему количеству поездок к месту назначения, так как по определению каждая поездка должна иметь две конечные точки. Предполагается, что разработанные оценки для поездки домой являются точными [3, 4]. Основное уравнение нормализации имеет вид

$$T'_j = T_j \frac{\sum_{i=1}^I T_i}{\sum_{j=1}^J T_j}.$$

Введем следующие переменные: T_i — поездки, происходящие в зоне i ; T_j — поездки, направленные в зону j ; T'_i — нормализованные поездки, происходящие в зоне i ; T'_j — нормализованные поездки, направленные в зону j ; T_h — поездки домой (как правило, утром — из дома, вечером — домой); T_w — поездки, смоделированные в конце работы (как правило, вечером — домой, утром — на работу); T_s — поездки, смоделированные в конце шопинга; H_i — число домашних хозяйств в зоне i ; $E_{off,i}$ — область деятельности организаций, офисы которых расположены в зоне i ; $E_{ret,i}$ — продажа в розницу и занятость в зоне i ; $E_{oth,i}$ — другая занятость в зоне i .

Результаты, полученные на первом этапе прогнозирования транспортировки, позволяют построить имитационную модель поездки с учетом всех входных данных для дальнейших разработок систем управления автотранспортными потоками.

2. Выбор места назначения. Распределение поездки или выбор места назначения является вторым этапом процесса прогнозирования. “Таблица поездок” представляет собой матрицу, отображающую число поездок от каждой начальной зоны до других зон. Распределение поездки является наименее разработанным компонентом транспортировки в планируемой модели. Рассмотрим модель Fratar и модель Gravity [5].

Наиболее простая модель распределения поездки — модель Fratar (модель роста) — экстраполирует таблицу поездки базисного года к будущему с учетом коэффициента, характеризующего рост и вычисляемого по формуле $T_{ijy+1} = gT_{ijy}$, где T_{ijy} — количество поездок из i в j в году y ; g — фактор роста.

Модель Fratar не учитывает изменение пространственной доступности вследствие увеличения поставок или изменений в подмодели “Путешествие”, что приводит к некорректности некоторых данных.

Модель Gravity описывает макроскопические отношения между местами (дом и рабочие места). Недостатком модели Gravity является ее составной характер. Несмотря на то что идеология модели остается эффективной на агрегированном уровне, более точные исследования позволяют сохранить уровень детализации информации максимально долго. Модель Gravity достаточно точно объясняет выбор большого количества людей, однако выбор конкретного человека существенно отличается от предсказанного значения.

Модель Gravity имеет вид [5]

$$T_{ij} = K_i K_j T_i T_j f(C_{ij}), \quad \sum_j T_{ij} = T_i, \quad \sum_i T_{ij} = T_j, \quad K_i = \frac{1}{\sum_j K_j T_j f(C_{ij})}, \quad K_j = \frac{1}{\sum_i K_i T_i f(C_{ij})},$$

где T_{ij} — поездки между начальным i и конечным j пунктами назначения; T_i — поездки, начинающиеся в зоне i ; T_j — поездки, заканчивающиеся в зоне j ; C_{ij} — стоимость путешествия между зонами i и j ; K_i, K_j — балансирующие факторы, определяемые итерационно; f — отклонение от первоначальной стоимости.

3. Выбор способа передвижения. Зональный анализ маршрутов поездок позволяет построить ряд таблиц данных о конечных пунктах поездок, в то время как анализ выбора способа позволяет разработчику модели определять, какой режим транспортировки будет использоваться. Например, модель, разработанная Chicago area transportation study (CATS), основана на транзите, для чего необходимо знать длительность путешествия транзитом [2]. Данная модель разделяет поездки транзитом на два класса: поездки, предварительно оплаченные наличными (главным образом, поездки в метро, экспресс-автобусами и пригородными поездами), и другие (в основном по местной системе магистралей). Подобные методики использовались во многих исследованиях.

Для решения некоторых задач модель CATS использует известные методы отклонения кривой. Авторы CATS исследовали отклонение трафика от второстепенных улиц к выбранным участкам скоростных автомагистралей. Кроме того, модель CATS использует кривые отклонения, чтобы установить, какой процент трафика будет использовать обход, построенный вокруг городов. Анализ выбора кривой отклонения выполняется по формуле $C_{transit}/C_{auto} = R$, где C_m — время путешествия способом m (m — *transit* или *auto* (*transit* — число пассажиров, использующих фиксированные маршруты; *auto* — число пассажиров, определяющих маршрут переезда самостоятельно)); R — полученные эмпирические данные.

Кривые строятся на основе эмпирических наблюдений, их можно изменить, используя улучшенные данные. Нетрудно получить результаты обработки массива данных, так как кривые доступны для исследователей. В некотором смысле анализ кривой аналогичен анализу экспертной системы. Планировщики могли оценивать количество пассажиров, использующих фиксированные маршруты в зависимости от маршрута и времени дня. Данные получены опытным путем, поэтому возможно построение соответствующих диаграмм.

3.1. Эконометрическая формулировка. Эконометрическая формулировка выбора способа передвижения сводится к нахождению функции максимальной полезности, в данном кон-

тексте имеется в виду полное удовлетворение от сделанного выбора или использованного сервиса.

Характеристика полезности u объекта x определяется следующим образом: $u(x) = v(x) + e(x)$ [2], где $v(x)$ — абсолютная полезность; $e(x)$ — терм ошибки. Термы ошибки независимы и одинаково распределены распределением Гумбеля или двойным экспоненциальным распределением. Это приводит к многочлену логит-модели, имеющему следующий вид: $\log(P_i/(1 - P_i)) = v(x_i)$ [3, 4]. В этом случае формула полезности имеет вид

$$\log\left(\frac{P_A}{1 - P_A}\right) = \beta_0 + \beta_1(C_A - C_T) + \beta_2(t_A - t_T) + \beta_3I + \beta_4N = v_A,$$

где P_A — вероятность выбора способа передвижения на личном автомобиле; C_A, C_T — стоимость передвижения; t_A, t_T — время прохождения маршрута; I — уровень дохода путешественника; N — рассматриваемое число путешественников.

Наиболее широко используемый вид модели

$$\frac{P_A}{1 - P_A} = e^{v_A}$$

или

$$P_A = \frac{e^{v_A}}{1 + e^{v_A}}.$$

В случае учета выбора места назначения j функция полезности имеет более сложный вид

$$\tilde{U}_{ij} = b_i - c_{ij} + \sum_k a_j^k X_j^k + \sum_l b_i^l Y_i^l + \sum_m d_{ij}^m Z_{ij}^m \quad \forall i, j, \quad (1)$$

где b_i — уровень дохода пассажира; c_{ij} — стоимость проезда из пункта i в пункт j ; X_j — оценка прочих атрибутов поездки с точки зрения пассажира; Y_i — мера измерения атрибутов поездки путешественником; Z_{ij} — матрица комбинаций альтернатив, включающая индивидуальные расхождения между альтернативами, выраженные в цене или иных характеристиках переезда из пункта i в пункт j ; знак “ $\sim$ ” обозначает условную полезность, т. е. полезность, соответствующую сделанному выбору.

Таким образом, полезность зависит не только от атрибутов альтернатив, но и от характеристик путешественника, так как два путешественника могут по-разному оценивать множество атрибутов. Например, для первого путешественника может быть более важна стоимость путешествия, а для второго — время.

Следует отметить, что спецификация полезности (1) становится общей для всех начал маршрутов i в зоне. Это означает, что все индивидуальные путешественники в данных требованиях представлены типичными (средними) путешественниками. Эти допущения удобны и применяются в случае большого количества требований и путешественников.

Параметры a_j^k , b_i^l , d_{ij}^m определяются на основе наблюдений реальных выборов путешественников.

Будем считать, что путешественник всегда выбирает альтернативное путешествие, обладающее максимальной полезностью. Применяя этот принцип, можно вычислить среднее альтернатив путешественника, которое будет выбрано, если условные полезности всех различных альтернатив будут оценены или упорядочены с помощью чисел. В этом случае,

когда все полезности имеют различные числовые значения, будут выбраны альтернативы с наибольшей полезностью.

3.2. *Случайная полезность или полином логит-модели.* Спецификация полученных в формуле (1) условных полезностей может быть улучшена и обобщена путем введения случайных термов между спецификацией полезности и средним путешественником. Полезность, получаемую средним путешественником, можно определить, выбрав альтернативу j из i :

$$\tilde{V}_{ij} = \tilde{U}_{ij} + \varepsilon_{ij} \quad \forall i, j. \quad (2)$$

Здесь $\tilde{U}_{ij}$ — средняя полезность, полученная из формулы (1); ε_{ij} — колебание или значение функции полезности для индивидуальных путешественников (этот терм ненаблюдаемый и неизмеримый в факторах, включенных в рассмотрение). Для удобства вычислений принимаем $M(\varepsilon_{ij}) = 0$; $\tilde{U}_{ij}$ — систематическая или фиксированная полезность (“видимая” часть полезности): $M(V_{ij}) = \tilde{U}_{ij}$; $\tilde{V}_{ij}$ — случайная величина, равная сумме величины $\tilde{U}_{ij}$ и случайной величины ε_{ij} , т. е. информация о полезностях путешественников (вероятностная).

В дальнейшем вероятности будут зависеть от $\tilde{U}_{ij}$ и ε_{ij} , где $\varepsilon_{i1}, \dots, \varepsilon_{ij}$ — совместно распределенные функции. Эти вероятности должны быть совместимы также с законами максимальной полезности.

Будем рассматривать случайные величины ε_j как нормально распределенные величины.

В общем случае множественных альтернатив можно получить приближенную оценку вероятности выбора с использованием последовательности процедур, которая, однако, быстро становится громоздкой. Для численной оценки вероятности выбора можно также использовать метод Монте — Карло, но это очень затруднительно и, как правило, не применяется на практике. Поэтому в качестве замены нормального распределения будет использоваться распределение “Гумбел” или “двойное экспоненциальное”. В этом случае вероятность равна

$$P_j = \frac{e^{\beta \tilde{U}_j}}{\sum_k e^{\beta \tilde{U}_k}} \quad \forall j. \quad (3)$$

Общее выражение для вероятности P_j с учетом того, что индивидуальный путешественник выбран случайно и предпочел альтернативу j , соответствует множественной логит-модели. В общем случае $\tilde{U}_j$ — многозначная функция от нескольких переменных. Например, в терминах общей спецификации полезности $\tilde{U}_j$, определенной по формуле (1), вероятности равны

$$P_j = \frac{e^{\beta \tilde{U}_j}}{\sum_l e^{\beta \tilde{U}_l}} = \frac{\exp \left[\beta \left(b - c_j + \sum_k a_j^k X_j^k \right) \right]}{\sum_l \exp \left[\beta \left(b - c_l + \sum_k a_l^k X_l^k \right) \right]} = \frac{e^{\beta(-c_j + A_j)}}{\sum_l e^{\beta(-c_l + A_l)}} \quad \forall j, \quad (4)$$

где $A_j = \sum_k a_j^k X_j^k$.

4. Выбор маршрута. Выбором маршрута, или назначением трафика, называется выбор маршрутов между начальной и конечной вершинами в сетях транспортировки.

Предположим, путешественнику необходимо выполнить трафик назначения (поездку) в пределах существующей сети магистралей и систем транзита с учетом множества изменений в сети. Для начала путешествия нужно знать характеристики поездки (время прохождения

транспортных потоков, обязательность посещения дополнительного пункта и др.) и рассматривать варианты изменения выбора с учетом множества изменений в сети.

Определим функцию эффективности звеньев сети. Стоимость, которую водитель устанавливает для пассажиров, называется предельной стоимостью. Однако может возникнуть ситуация, когда при принятии решения водителю необходимо принять во внимание собственную (среднюю) стоимость и игнорировать затраты, что вынудит его совершить поездку в убыток, т. е. по стоимости ниже предельной.

При выборе маршрута необходимо учитывать несколько правил, например принцип равновесия, принцип сохранения потока и др. [6].

Принцип равновесия: каждый пользователь действует так, чтобы минимизировать стоимость, подчиненную любому другому пользователю, который в свою очередь действует так же. Время прохождения принимается одинаковым на всех используемых маршрутах и меньше, чем на любом неиспользованном маршруте.

Оптимальная система: каждый пользователь действует так, чтобы минимизировать полное время прохождения по системе.

Сохранение потока: число транспортных средств, входящих в вершину, равно числу транспортных средств, выходящих из нее в течение данного промежутка времени (за исключением источников и стоков). Аналогично число транспортных средств, входящих в сток, равно числу таких средств, выходящих из источника (за длительный период времени).

Заключение. Разбиение задачи на основные части позволяет получить результаты на каждом этапе моделирования. Каждый этап можно рассматривать как отдельную задачу, их объединение дает возможность представить картину в целом. В работе рассматривается логит-модель для первых двух этапов процесса прогнозирования и построения функции полезности. При выборе маршрута учитываются принципы равновесия, выбора оптимальной системы и сохранения потока.

Список литературы

1. OPPENHEIM N. Urban travel demand modeling from individual choices to general equilibrium. N. Y.: Wiley-Intersci. Publ., 1995.
2. [http://en.wikibooks.org/wiki/Fundamentals of Transportation/Trip Generation](http://en.wikibooks.org/wiki/Fundamentals_of_Transportation/Trip_Generation).
3. Гумбель Э. Статистика экстремальных значений / Пер. с англ. М.: Мир, 1965.
4. Носко В. П. Эконометрика для начинающих (Дополнительные главы). М.: ИЭПП, 2005.
5. ГАСНИКОВ А. В. Введение в математическое моделирование транспортных потоков. М.: МФТИ, 2010.
6. КРАВЧЕНКО П. С., ОМАРОВА Г. А. Моделирование транспортного потока как задача принятия решения // Тр. Всерос. науч. конф. "Моделирование систем информатики", Новосибирск, 8–11 нояб. 2011. С. 1–6.

*Омарова Гульзира Алимовна — канд. физ.-мат. наук, науч. сотр. Института вычислительной математики и математической геофизики СО РАН;
e-mail: gulzira@rav.sccc.ru*

Дата поступления — 6.06.12