

УДК 656.02

Лебедева Ольга Анатольевна,
к.т.н., доцент кафедры «Управление на автомобильном транспорте»,
ФГБОУ ВО «Ангарский государственный технический университет»,
тел.: +7(952)6326611, e-mail: kravhome@mail.ru

Ерофеев Ермак Владимирович,
студент гр. ИТБ-23-1,
ФГБОУ ВО «Ангарский государственный технический университет»,
тел.: +7(950)0535465, e-mail: ermak.080@mail.ru

ОБЗОР СОВРЕМЕННЫХ МЕТОДОВ СБОРА ДАННЫХ В ТРАНСПОРТНОМ МОДЕЛИРОВАНИИ

Lebedeva O.A., Erofeev E.V.

A REVIEW OF MODERN DATA COLLECTION METHODS IN TRANSPORT MODELING

Аннотация. В этой статье были рассмотрены существующие методики сбора данных для процедуры транспортного моделирования. Приведены источники классических и альтернативных методов сбора. Рассмотрены достоинства и недостатки вариантов сбора данных с мобильных телефонов, смарт-карт, приложений и агрегаторов приложений, GPS-навигации, Bluetooth, Wi-Fi.

Ключевые слова: транспортное моделирование, методики сбора данных, GPS-навигация, генерация транспортного спроса.

Abstract. This article reviewed existing data collection methods for the transport modeling procedure. Sources of classical and alternative collection methods are provided. The advantages and disadvantages of data collection options from mobile phones, smart cards, applications and application aggregators, GPS navigation, Bluetooth, Wi-Fi are considered.

Keywords: transport modeling, data collection methods, GPS navigation, transport demand generation.

В классической теории сбора данных в течение последних 50 лет специалисты по моделированию транспорта использовали одни и те же методы. Они применяются как к стратегическим, так и к тактическим моделям с небольшими изменениями. Это касается как классических агрегированных моделей, так и более современных моделей на основе агентов. Наиболее важные этапы сбора данных выражаются в дорогостоящем мероприятии – обследовании транспортной подвижности, традиционно включающем тщательное планирование и контроль выборки, длительные опросы и обширную обработку данных. Несмотря на изменения, позволяющие в определенной степени автоматизировать заполнение и/или поддержку (посредством телефонного опроса), на устранение ошибок требуется значительное время.

Из-за своей стоимости они обычно охватывают менее 2% населения в регионе. Качественный опрос генерирует много полезной информации о наличии транспортных средств, уровне дохода, видах деятельности и схемах поездок (поездки и туры).

Опрос является ключевым источником данных для оценки способа передвижения и других вариантов поездок, но небольшой размер выборки не раскрывает общую картину перемещений в форме матриц корреспонденций. Небольшую выборку поездок, полученную из опросов в местах проживания, необходимо дополнить с помощью опросов, собираемых на улично-дорожной сети (УДС), в общественном транспорте, станциях и остановках [1]. Высокого качества выборки вряд ли удастся достичь, поскольку данные могут быть собраны только на выбранных улицах или пунктах. Поскольку эти опросы включают меньше вопросов, они являются экономически эффективным методом сбора данных о поездках по цели, способу передвижения и времени суток. Опросы на участках УДС часто проводятся для оценки перемещения товаров. Методы «предпочтений» были разработаны в восьмидесятих годах для облегчения оценки моделей выбора, когда производился сравнительный анализ вариантов (при рассмотрении инвестиций в систему рельсового транспорта или скоростного автобуса). Они

включают опрос населения об осуществлении выбора между гипотетическими альтернативами, представленными в виде таких атрибутов, как время, деньги и частота. Опросы требуют «активного» взаимодействия с респондентами. Практики сталкиваются с растущим числом проблем с опросами о транспортной подвижности населения. Доступ к некоторым жилым зданиям становится проблематичным. Население неохотно отвечает на длинные опросы и часто пропускают или упрощают ответы. Некачественный опрос поездок стал серьезной проблемой. Пассивные методы сбора данных не страдают от этого конкретного ограничения, но они не предлагают полные поведенческие данные обследований поездок населения. С точки зрения транспортного моделирования необходимо как представление перемещений в базовом году, так и более глубокое понимание поведения при поездках, заложенное в функциональных формах и параметрах моделей выбора.

Для выявления и учета занижения информации о поездках требуются данные об объемах перевозок и о пассажиропотоках. Они дополняются обследованиями времени в пути и предложениями, необходимыми для разработки транспортных моделей. В настоящее время иногда используется автоматическое распознавание номерных знаков, когда для микромоделей требуется только местное обследование пункта отправления и назначения транспортного средства.

Обычные методы сбора данных, указанные выше, являются финансово затратными и занимают значительную часть времени, доступного для его проведения. Данные имеют ограниченный срок, устаревая очень быстро. Риски некачественной выборки реальны и требуют тщательного контроля, обнаружения и исправления. Процесс перехода от обычного сбора данных к валидной модели не обходится без корректировок и устранения ошибок. Его можно обобщить на рис. 1.



Рис. 1. Этапы сбора данных в процедурах транспортного моделирования [2]

Целевая транспортная модель обычно представляет собой «средний день какого-то месяца», генерирующий средние потоки и выгоды. Это примечательно, поскольку изменчивость схем поездок считается источником нерегулярности времени в пути, что является серьезной проблемой в крупных городах.

Для мультимодальной модели города процесс от сбора данных до внедрения и прогнозирования занимает несколько лет. Затем модель будет использоваться в процессе планирования и выбора проектов для внедрения. Обычно рекомендуется не использовать модель, основанную на данных, которым больше пяти лет. В этом случае модель необходимо обновить за счет использования ограниченных опросов и агрегированных данных, таких как продажи билетов и объем транспортных потоков.

Сегодня существует возможность использовать новые источники данных для замены или поддержки традиционных методов сбора данных, облегчая обновление моделей и изучение новых измерений спроса на поездки [3, 4].

Эти источники данных предлагают альтернативные методы сбора, которые дают возможность получения более крупных размеров выборки и более длительного охвата времени. Некоторые из них основаны на датчиках Bluetooth и Wi-Fi, другие – на смарт-картах общественного транспорта, приложениях для смартфонов, данных мобильных сетей и телематике транспортных средств. Все это пассивные формы сбора данных, которые используют преимущества цифровых следов, генерируемых различными устройствами, приведены на рис. 2.



Рис. 2. Источники данных транспортной мобильности [2]

Но не все источники способны определить перемещения от пункта отправления до пункта назначения. Bluetooth и WiFi собирают данные в основном со смартфонов и транспортных средств (только локальные перемещения). Данные смарт-карт общественного транспорта охватывают личные перемещения от остановки/станции отправления до остановки/станции назначения. На железнодорожных перевозках производится регистрация станции отправления и назначения, но на автобусах идентифицируется только посадка, а место высадки должно оцениваться другим методом. При использовании GPS-треков навигации транспортных средств характер выборки, охватывает только перемещения транспортных средств, а не населения.

Существует два источника данных, которые предлагают перспективу предоставления перемещений «от двери до двери», поскольку они основаны на мобильных телефонах: устройствах, которые большинство населения носит с собой постоянно. Один из них повторно использует данные, генерируемые оператором мобильной сети (MNO) во время нормальной работы сети мобильной связи. Другой использует данные, генерируемые микрочипом глобальной навигационной спутниковой системы, встроенным во все смартфоны; наиболее распространенным из них является GPS. Большинство смартфонов используют A-GPS для повышения точности определения местоположения путем объединения GPS с данными сотовой связи и сигналами сети Wi-Fi.

В этом обзоре первый источник назы-

вается «данными мобильной сети» или «данными MNO», а второй – «данными смартфона».

Мобильные телефоны генерируют два основных типа цифровых следов. Один набор создается в рамках нормальной работы сети мобильной связи, и эти данные просто повторно используются для предоставления показателей мобильности и других индикаторов. Смартфоны также имеют функцию GPS, к которой обращаются несколько приложений, и они генерируют дополнительные данные.

Существует несколько типов данных MNO, собираемых для разных целей и содержащих разные виды информации. Они генерируются при взаимодействии мобильного телефона с антенной или базовой приемопередающей станцией (BTS). MNO собирают записи о вызовах, генерируемые во время вызовов и сообщений, и записи о протоколах Интернета, генерируемые при передаче данных на смартфоны и со смартфонов для взимания платы с пользователей. MNO также генерируют данные для мониторинга услуг и обеспечения быстрого поиска телефона, когда требуется подключение. Одни из таких данных генерируются, когда телефон перемещается из одной группы ячеек (локальной области) в другую. Другой тип данных получается в результате зондирования, используемого MNO для дальнейшего улучшения качества обслуживания. В этом случае каждое устройство зондируется через регулярные интервалы времени, чтобы обеспечить быстрый ответ и контролировать другие аспекты качества сети. Эти два типа данных часто объединяются для предостав-

ления более частых точек, определенных как данные мобильной сети. Каждое обновление локальной области и зонда генерирует временную метку и идентификатор, с которой был установлен контакт; следовательно, пространственная точность этой точки данных является точностью, участвующей в контакте. Многие приложения для смартфонов используют данные о местоположении для улучшения предоставляемых ими услуг. Данные о местоположении, полученные таким образом, затем собираются и объединяются компаниями, поддерживающими рекламу идентификатора устройства (ID). Операционные системы мобильных телефонов предоставляют маркетинговый идентификатор устройства для поддержки рекламы. Они используются для связывания меток местоположения и времени (и других пользовательских данных) из разных приложений на одном телефоне. Маркетинговые идентификаторы используются для таргетинга рекламы на определенное устройство в зависимости от местоположения и профиля пользователя и могут быть повторно использованы для доставки данных о перемещении. Пользователь может ограничить и сбросить эти идентификаторы, но мало кто это делает. Пространственная точность этих данных соответствует точности службы определения местоположения на телефоне, обычно комбинации GPS и других сигналов.

Источники данных обрабатываются для предоставления коммерческих продуктов мобильности различными предприятиями. Существует по крайней мере два типа поставщиков матриц поездок из данных мо-

бильных сетей. Один из них – сам оператор мобильной связи, стремящийся напрямую монетизировать имеющиеся у него данные. Иногда консультанты берут эти матрицы поездок MNO и адаптируют их для более точного соответствия потребностям конечных пользователей. Есть также компании по анализу данных; они обрабатывают данные от MNO, объединяют их с другими источниками, добавляют контекст и предоставляют матрицы поездок и показатели мобильности, напрямую адаптированные к требованиям конечного пользователя. Данные приложений для смартфонов обычно агрегируются из ряда приложений. Матрицы поездок как из данных MNO, так и из данных приложений были доступны и используются более пяти лет, но их принятие не было всеобщим. Существует множество пользователей, разочарованных качеством полученных данных. Следует задаться вопросом, насколько близко представление мобильности, получаемое в следствии оценки этих матриц, к реальности, что иллюстрирует неотъемлемые ограничения метода. Кроме этого, практически невозможно различать вид транспорта, используемый в густонаселенных городских районах. Но есть категория пользователей, которые используют эти данные. Важно изучить причины этой неоднозначной оценки, попытаться понять сильные и слабые стороны этих источников данных и выработать рекомендации по их спецификации и использованию. Для этого необходимо более подробно исследовать природу и характеристики этих источников данных.

СПИСОК ЛИТЕРАТУРЫ

1. **Полтавская, Ю.О.** Методы сбора данных о продолжительности движения на маршруте и требования к объему выборки / Ю.О. Полтавская // Вестник Ангарского государственного технического университета. 2018. № 12. С. 192-195.
2. **Kusumastuti, D.** Mixed-Use Urban Planning and Development / D. Kusumastuti, A.J. Nicholson, // BRANZ Study Report 481. BRANZ Ltd. Judgeford, New Zealand, 2017.
3. **Лебедева, О.А.** Алгоритм восстановления маршрута с помощью геоинформационных данных / О.А. Лебедева, З.Ф. Та // Современные технологии и научно-технический прогресс. 2021. № 8. С. 179-180.
4. **Лебедева, О.А.** Оценка времени в пути с использованием бортового оборудования автобусов / О.А. Лебедева // Современные технологии и научно-технический прогресс. 2020. Т. 1. № 7. С. 163-164.
5. **Полтавская, Ю.О.** Применение геоинформационных систем для обеспечения устойчивого развития транспортной системы города / Ю.О. Полтавская // В сборнике: Информационные технологии в науке, управлении, социальной сфере и медицине. Сборник научных трудов VI Международной научной конференции. Под редакцией О.Г. Берестневой, В.В. Спицына, А.И. Труфанов, Т.А. Гладковой. 2019. С. 164-167.