

**Европейская экономическая комиссия****Комитет по внутреннему транспорту****Рабочая группа по статистике транспорта****Семьдесят вторая сессия**

Женева, 9–11 июня 2021 года

Пункт 2 предварительной повестки дня

Большие данные для статистики транспорта**Использование плавающих данных об автомобилях
для подготовки транспортной статистики****Записка секретариата***Резюме*

В настоящем документе описывается опыт Германии в части использования плавающих данных об автомобилях — потенциальном новом источнике данных, который может быть использован для получения статистики о поездках пассажиров. Результаты показывают, что данные о пунктах отправления и назначения, полученные на основе плавающих данных об автомобилях, сопоставимы с данными, собранными с помощью традиционного обследования мобильности. В приложении приводится более подробная техническая информация о методологии исследования.

I. Справочная информация

1. Как отмечалось в ходе недавних вебинаров, совместно организованных секретариатом и Международным транспортным форумом (МТФ) (ECE/TRANS/WP.6/2021/4), эпидемия COVID-19 ускорила переход от подготовки транспортной статистики на основе данных обследований к использованию новых источников данных, включая источники больших данных. Одним из таких источников являются плавающие данные об автомобилях (ПДА), которые представляют собой данные, собираемые непосредственно самим транспортным средством во время его движения, и обычно включают данные о местоположении и скорости ТС. В настоящем документе излагается опыт Германии в этой области.

**II. Краткая информация об опыте Германии в области
работы с плавающими данными об автомобилях**

2. Использование ПДА, которые генерируются в больших количествах и практически непрерывно во время логистических и навигационных процессов, ставит



новые задачи в контексте проведения исследований, касающихся транспортных перевозок. Наличие обширного массива данных создает ожидание, что эти данные позволят в будущем предоставлять необходимую информацию для планирования условий движения быстрее и в более широком объеме и что это может снизить затраты на проведение традиционных, эмпирических обследований. В этой связи Федеральное министерство транспорта и цифровой инфраструктуры (ФМТИ) поручило Горному университету Вупперталя провести исследование на тему «Матрицы “пункт отправления — пункт назначения” для автомобильного и грузового транспорта в общенациональном масштабе за 2015 год: анализ на основе спутниковых данных» (FE 97.372/2016).

3. Использование ПДА, которые генерируются с помощью навигационных приборов и основаны на системе спутникового позиционирования (обеспечивающей высочайшую пространственную точность), позволяет получить информацию о пунктах отправления и назначения, а также о скорости движения и маршрутах транспортных средств. Эта информация может быть проанализирована в различных временных и пространственных контекстах. Целью данного исследовательского проекта было выяснить, можно ли получить действительные оценочные данные о матрицах «пункт отправления — пункт назначения» в общенациональном масштабе за полный год, и если да, то каким образом.

4. Исследовательский проект был реализован на основе набора ПДА за 2015 год, который состоял из данных, полученных от нескольких поставщиков навигационных услуг, как по пассажирским, так и по грузовым транспортным средствам. Пространственные и темпоральные кластеры этого набора данных были изучены на различных уровнях. Сведения о принадлежности транспортных средств к соответствующим типам были проверены и скорректированы, после чего на их основе полученные данные о пространственном местоположении (цепочка сигналов) были разбиты на поездки. Затем определялась цель поездок на основе проводившихся ранее эмпирических обследований с помощью анализа общности, но только для данных по легковым автомобилям; для грузовых перевозок представлялось невозможным аналогичным образом выявить цели поездок.

5. Поскольку поездки, данные о которых были получены в формате ПДА, составляли лишь небольшую выборку из всех поездок по стране в течение полного года, а также поскольку этот набор данных характеризовался лишь выборочным использованием навигационных приборов, то матрицы «пункт отправления — пункт назначения» для автомобильного транспорта были экстраполированы с использованием параметров, полученных в ходе обследования «Мобильность в Германии» (Mobilität in Deutschland (MiD)). Для этой цели использовались только маргинальные суммы для матрицы «пункт отправления — пункт назначения», так что распределение общего трафика по индивидуальным соотношениям картировалось исключительно на основе набора данных ПДА. Ввиду отсутствия подходящих методов оценки маргинальных сумм для данных о грузах, матрица по грузовому транспорту была экстраполирована с помощью данных об объемах движения, собираемых на постоянных счетных станциях федеральной сети автомобильных дорог.

6. Последующий анализ экстраполированных матриц опирался на агрегированные параметры (пробег, количество поездок, пройденное расстояние) и показал в целом удовлетворительные результаты для матрицы по автомобильному транспорту, оцененной с помощью ПДА. Следует отметить, что все использованные сравнительные значения также были сгенерированы на основе случайных выборок и модельных расчетов. В обеих матрицах трансграничные и транзитные поездки не удалось репрезентативно отобразить ввиду сильных пространственных кластеров в наборе данных ПДА. Что касается грузовых перевозок, то результаты экстраполяции имеют недостаточную валидацию маргинальными суммами. Короткие поездки без проезда через постоянную счетную станцию затрудняли оценку действительных коэффициентов экстраполяции для соответствующих соотношений.

7. Несмотря на выявленные проблемы, ПДА считаются многообещающим источником данных с точки зрения определения звеньев «место отправления — место назначения» в крупном масштабе — данных, которые было бы трудно получить с

помощью проведения обследования. В будущем следует ожидать дальнейшего расширения использования навигационных устройств, особенно применительно к более коротким и регулярным поездкам. Кроме того, гармонизация европейского рынка мобильной связи позволит устранить барьеры, которые на момент проведения исследования по-прежнему препятствовали использованию навигационных услуг с помощью смартфонов и, таким образом, стали причиной того, что трансграничные поездки были картированы лишь выборочно. Существует также значительная потребность в дальнейших методологических исследованиях, касающихся использования ПДА.

III. Выводы

8. Результаты опыта Германии в области работы с ПДА показывают, что этот источник данных может в будущем либо дополнить, либо даже заменить данные, получаемые с помощью традиционных обследований мобильности. Кроме того, в будущем следует ожидать совершенствования процедур для отслеживания трансграничных поездок с помощью ПДА в связи с изменениями на европейском рынке мобильной связи. Требуется дальнейшая работа для обеспечения того, чтобы ПДА могли точным образом охватить все население.

Приложение

Техническая информация о методологии исследования

Цель проекта заключалась в том, чтобы использовать набор данных ПДА за 2017 год для проверки и валидации метода (а именно, матричной оценки взаимозависимости), разработанного в рамках предыдущего исследовательского проекта FE 97.372/2016 «Общегосударственная транспортная взаимозависимость в области автомобильного индивидуального транспорта в 2015 году: анализ на основе спутниковых данных»¹. Поэтому в ходе исследования был поставлен вопрос о том, в какой степени распределение транспортных потоков остается стабильным с течением времени и могут ли указанные данные использоваться для оценки общенациональной транспортной взаимозависимости, т. е. моделей «пункт отправления — пункт назначения». В рамках данного исследовательского проекта разработанные методы были применены к набору плавающих данных об автомобилях Немецкого автомобильного клуба (Allgemeiner Deutscher Automobil-Club e.V. (ADAC)) за 2017 год. В ходе реализации проекта были усовершенствованы существующие алгоритмы сегментации поездок и экстраполяции данных.

По сравнению с 2015 годом объем базовых данных увеличился примерно на 30 процентов. Благодаря гибридной процедуре экстраполяции, которая основана на показателях, зависящих от типа региона, класса дистанции и назначения поездок и полученных в ходе исследования «Мобильность в Германии» (Mobilität in Deutschland (MiD)) 2017 года, сгенерированные показатели выглядят убедительно. Оценка показала, что при этом ни пространственное, ни темпоральное распределение данных не претерпело фундаментальных изменений. Дополнительные поездки, охваченные увеличенным объемом данных, распределяются аналогичным образом. Данные о выборе места назначения, а также симметрия данных о поездках, генерируемых на основе ПДА, демонстрируют весьма сильную корреляцию между отдельными годами как для автомобильного индивидуального транспорта, так и для большегрузного транспорта. Исследование также показало, что, несмотря на отмену общеевропейских тарифов на мобильный роуминг, уровень картирования трансграничных поездок пока еще недостаточен для получения полностью убедительных результатов. В дополнение к делению ячеек на уровнях НТЕС-3² и МАЕ-1³, которое использовалось для генерации матриц, была разработана динамическая модель ячейки сетки, совместимая с INSPiRE⁴, на основе сетевых ячеек 100 x 100 м, применявшихся в ходе обследования 2011 года. Исходя из заранее заданных параметров для движения по пунктам отправления и назначения, данный метод обеспечивает — на основе данных ПДА — динамическую генерацию ячеек сетки, кратных ячейкам, которые использовались в ходе обследования, что позволяет добиться гораздо более точного разделения на ячейки в плане соответствия данным.

¹ Доступен на немецком языке по адресу https://www.bmvi.de/SharedDocs/DE/Anlage/G/verkehr-in-zahlen-praesentation-3.pdf?__blob=publicationFile.

² Номенклатура территориальных единиц для статистики Европейской комиссии. Уровень НТЕС-3 примерно совпадает с муниципальным уровнем; в Германии насчитывается 429 регионов уровня НТЕС-3.

³ Местные административные единицы (также определены Европейской комиссией). МАЕ-1 находится ниже уровнем, чем НТЕС-3.

⁴ INSPiRE — инфраструктура пространственных данных Европейского союза.