



Е.Р. Майоров, О.Н. Сапрыкин

АНАЛИЗ ДАННЫХ UBER MOVEMENT НА ВОЗМОЖНОСТЬ ИСПОЛЬЗОВАНИЯ В ИМИТАЦИОННОМ МОДЕЛИРОВАНИИ ТРАНСПОРТНЫХ ПОТОКОВ

(Самарский университет)

Жизнь в современном мегаполисе практически невозможно представить без личного автомобиля. Однако их массовое использование превратило этот вид транспорта из самого удобного и быстрого, в самый медленный и затратный по времени. Одной из основных причин этого является несоответствие городской транспортной инфраструктуры транспортному спросу. В настоящее время многие города пытаются исправить это и занимаются реконструкцией улично-дорожной сети в целях: увеличения пропускной способности, качества дорог, безопасности и др. [1] Однако, часто возникает вопрос оценки и прогнозирования таких изменений (например, количество полос или светофорный цикл). Современным и практичным способом прогнозирования и поиска сбалансированного решения как на конкретном участке дороги, так и на городской сети в целом, является моделирование проектов [2]. Однако в процессе создания имитационной модели часто возникает проблема соответствия этой модели реальной мобильности населения города. Многие современные исследователи занимаются анализом данных и определением наиболее подходящих данных для оценки соответствия модельных данных реальным.

Один из основных критериев при поиске данных является их открытость. Многие государства заботятся о защите прав их граждан на неприкосновенность частной жизни и поэтому ограничивают распространение персональных данных. Однако некоторые компании, которые специализируются на сборе данных и Big Data анализе, анонимизируют персональные данные, что позволяет их использовать в исследованиях.

Одной из таких компаний, предоставляющих открытые данные о мобильности населения в городах является Uber со своим сервисом Movement. Компания Uber уже много лет является лидером среди приложений такси и за это время собрала большое количество статистической информации. В сервисе Uber Movement находится агрегированная информация о среднем времени затрачиваемого на передвижения из одного транспортного района в другой (Рисунок 1).

Эти данные можно экспортировать и использовать для анализа собственными средствами и методами. Формат данных представлен на рисунке 2: sourceid – номер транспортного района отправления, dstid – номер транспортного района назначения, mean_travel_time – среднее время, необходимое для передвижения из района отправления в район назначения.

Для проведение сравнительного анализа и анализа соответствия этих данных реальным, был разработан экспериментальный метод, основанный на



проверке имитационном моделировании транспортных потоков. Метод состоит из следующих основных шагов:

1. Импорт и предварительная обработка данных Uber Movement.
2. Преобразование временной матрицы в матрицу корреспонденций.
3. Создание имитационной модели по матрице корреспонденций.
4. Сравнение интенсивности транспортных потоков в модели с данными детекторов.

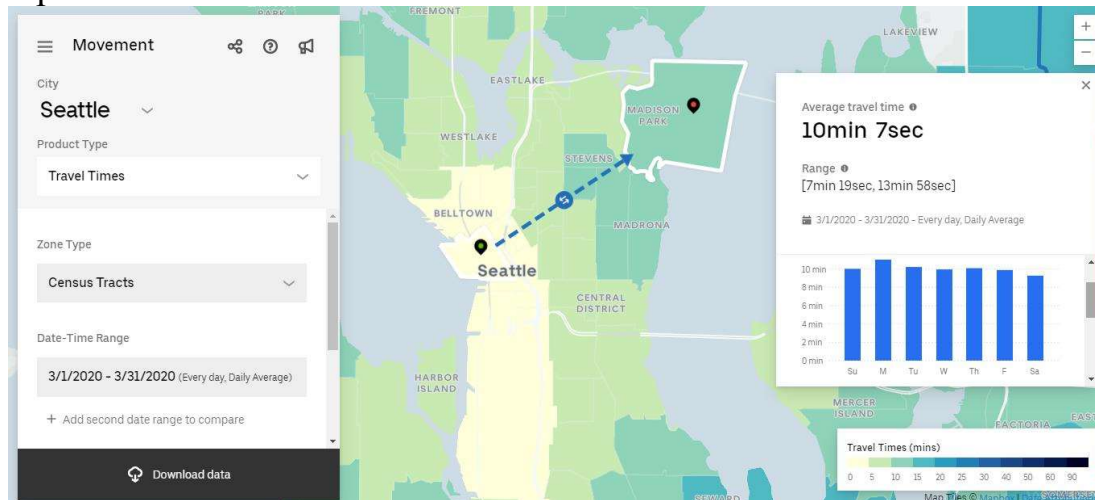


Рисунок 1 – Сервис Uber Movement

	A	B	C	D	E
1	sourceid	dstid	hod	mean_travel_time	standard_deviation_travel_time
2	658	101	19	1088.21	397.11
3	241	83	18	1552.0	367.42
4	702	227	8	2620.78	652.56
5	610	581	19	453.22	79.15
6	240	93	18	2086.2	658.23
7	267	319	10	1064.8	277.81
8	641	271	19	2150.0	493.76
9	66	109	20	708.88	179.24
10	6	193	11	1588.25	709.3
11	17	83	11	896.5	183.64
12	285	139	10	828.89	126.75
13	136	536	10	1130.44	142.9
14	166	236	10	802.7	213.03
15	134	556	10	1589.38	116.89
16	152	376	10	1510.2	451.87

Рисунок 2 – Цифровой формат данных Uber Movement

Преобразование временной матрицы Uber Movement в матрицу корреспонденций выполнялось по формуле:

$$T_{ij} = \frac{L_{ij} * p}{t_{ij}} \quad (1)$$

- T_{ij} – объем перемещений между районами;
- L_{ij} – манхэттенское расстояние между районами;
- p – суммарная плотность населения в районе отправления и районе назначения;



- t_{ij} – среднее время, затрачиваемое между районами.

Имитационное моделирование проводилось в программном продукте SUMO [3], который обладает средствами моделирования синтетического транспортного спроса по матрице корреспонденций.

Проверка корректности имитационной модели осуществляется при помощи GEN-анализа:

$$GEN = \sqrt{\frac{2(M-C)^2}{M+C}} \quad (2)$$

где M – данные об интенсивности Uber Movement, C – данные об интенсивности с государственного сайта статистики. Считается, что модель адекватно описывает реальные транспортные потоки, если 85% и более замеров имеют значение $GEN < 5$ [4].

Метод был опробован на открытых данных города Сиэтл (США). С официального сайта Сиэтла были экспортированы статистические данные в формате среднего значения интенсивности на участках улично-дорожной сети за 24 часа в рабочие дни (с понедельника по пятницу) – AWDТ (Average Weekday Day Traffic). Эти данные использовались для валидации модели, построенной по данным Uber Movement.

Анализ результатов имитационного моделирования показал, что 52% данных взятых с Uber Movement входят в адекватное значение $GEN < 5$, 76% - входят в диапазон $GEN < 10$.

Помимо GEN-анализа так же был проведен регрессионный анализ, сравнивающий моделируемый и реальный транспортные потоки (рисунок 3) [5].

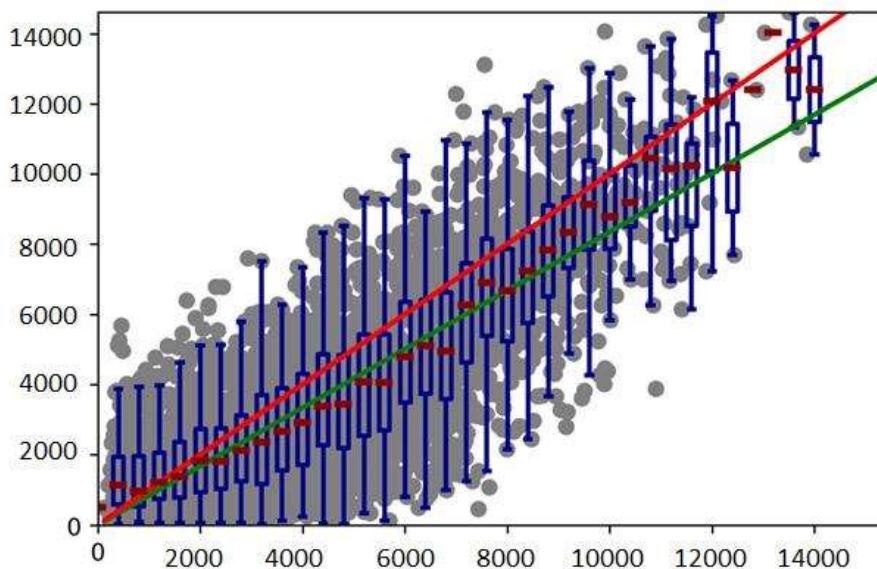


Рисунок 3 – График линейной регрессии

Так же были вычислены значение коэффициента $b=0.8389$, значение среднеквадратичной ошибки = 19292674.63 и R^2 (коэффициент детерминации) = 0.60.



Полученные результаты говорят о том, что модель построенная на основе данных Uber Movement не отвечает требованиям адекватности по GEN метрике, однако регрессионный анализ показывает причину таких значений. Из графика видно, что моделируемый поток в среднем на 17% ниже реальных данных. Возможными путями исправления модели являются:

- введение дополнительных коэффициентов в формулу (1)
- калибровка имитационной модели на данных детекторов.

Литература

1. Михайлов А.Ю., Головных И.М., Современные тенденции проектирования и реконструкции улично-дорожных сетей городов// – Новосибирск: Наука, – 2004. – с.267
2. Швецов В. И., Математическое моделирование транспортных потоков// Автомат. и телемех., – 2003. – №11. – с.3–46.
3. Lopez P.A. et al., Microscopic Traffic Simulation using SUMO, 2018 21st International Conference on Intelligent Transportation Systems (ITSC), Maui, HI, USA, 2018, pp. 2575-2582, doi: 10.1109/ITSC.2018.8569938
4. Feldman O., The GEN Measure And Quality Of The Highway Assignment Models, European Transport Conference At: Glasgow, October 2012
5. Шашков, В.Б. Прикладной регрессионный анализ. Многофакторная регрессия// Учебное пособие. – Оренбург: ГОУ ВПО ОГУ, – 2003. – с.363.

А.А. Мещеряков

АЛГОРИТМ ПОСТРОЕНИЯ ОПТИМАЛЬНОГО МАРШРУТА ДЛЯ МОРСКИХ СУДОВ

(Самарский университет)

Одной из важных проблем судоходства является задача построения оптимального маршрута с учетом метеорологических условий и обеспечения безопасности прохождения пути. Построение оптимального по времени маршрута с учетом данных прогноза погоды особенно важно с точки зрения сокращения экономических затрат. Поэтому проведение исследований по оптимизации маршрута по указанным критериям крайне актуально в навигации.

Несмотря на быстрое развитие информационных технологий в последние годы, до сих пор не существует универсального общепринятого метода построения оптимальных маршрутов, гарантирующего вычислительную простоту и точность прогноза.

Для решения задачи построения оптимального маршрута сначала необходимо описать математическую модель. Скорость судна при плавании в данной модели описывается следующим соотношением:

$$V = V_0 + V_{wt}\alpha + V_{wd}\beta,$$

Где V_0 - собственная скорость судна;