Improving bike travel demand generation with dynamic data: an application to the Paris Metropolis

Guoxi Feng*[1], Azise Oumar Diallo[1], Alexandre Chasse[1]

[1] Digital Science and Technology Division, IFP Energies Nouvelles, France

Keywords: bike travel demand; data fusion; OD matrix estimation; MATSim;

The mobility behaviors and transport systems have complex interactions. In particular, mobility patterns have evolved dynamically during and after the Covid-19 crisis. The rise of the home office and the rise in bike commuting in Paris may have a significant impact on travel behaviors. Moreover, people could move differently during weekdays. The existing approach [1] generates the travel demand from the survey data to model an average day. Such mobility surveys are carried out each year or even each decade. This methodology may not be able to capture short-term behavior. Modeling dynamically travel demand provides valuable insights into transportation needs, enabling effective planning, management, and resource allocation for public decision makers. Therefore, there is a great need for discussing how to improve the travel demand generation with dynamic information such as crowdsensing data, measures from counters.

In this study, we specifically focus on the bike travel demand modeling for the following reasons:

- The calibration of multi-agent simulation with desegregated data such as bike counter data is a challenging task [2]. A travel demand modeling by considering desegregated data could provide an initial scenario which is expected to lead a more consistent simulation results regarding to the observed bike flow.
- Bike considered as sustainable low-carbon transport get more and more intention recently [3].

The bike travel demand remains marginal compared to other transportation means such as car, public transport and walk. Then, the study of bike travel demand tendency can be isolated by the assumption other mode travel demand being equal. With such consideration, we aim to validate the proposed approach on the bike travel demand as a proof-of-concept. The obtained result could provide a solid base to generalize the approach to other mode transport modes later.

The bike travel demand generation that we propose is composed of two main steps. The first step is the travel demand origin-destination (OD) matrix estimation by data fusion. In this step, three types of data are used such as survey data, bike counter station data and GPS trace data. The survey data (population census, household travel survey) is used by EQASim [1] for generating the initial OD matrix of bike travel represented by M_{OD} in (1). The notation OD_{ij} represents the number of trips from the zone i to the zone j where $i, j \in \{1,2, \dots n\}$. We decompose the study zone e.g. Paris into 80 zones $(n = 80)$. Since the mobility survey data is relatively old, the generated M_{OD} may not be representative. The idea is to adjust this OD matrix with a correction matrix λ. To search this correction matrix, bike counter data and the GPS trace data are used. The GPS trace data from mobile application allows to understand how the cyclists move. We can estimate the probability P_{ijk} in (1) where we travel from a given origin i to a given destination j by going through the counter station k. We compute such probability of each pair of OD for a same station

k. We obtain a kernel matrix $M_{kernel-k}$ in (1) which describes how the cyclists could move statistically regarding to this counter station k.

$$M_{OD} = \begin{bmatrix} OD_{ij} & \cdots & OD_{in} \\ \vdots & \ddots & \cdots \\ OD_{nj} & \cdots & OD_{nn} \end{bmatrix},\ \lambda = \begin{bmatrix} \lambda_{ij} & \cdots & \lambda_{in} \\ \vdots & \ddots & \cdots \\ \lambda_{nj} & \cdots & \lambda_{nn} \end{bmatrix},\ M_{kernel-k} = \begin{bmatrix} P_{ijk} & \cdots & P_{ink} \\ \vdots & \ddots & \cdots \\ P_{njk} & \cdots & P_{nnk} \end{bmatrix} \quad (1)$$

We compute the quadric sum $I^T[M_{kernel-k} \circ \lambda \circ M_{OD}]I$ where the vector I is the column vector whose entries are 1 and the operator $\circ$ is the Hadamard product. This term provides the estimated flow on the counter station k after the adjustment of M_{OD}. This flow should be equal to the flow F_k observed by the counter k. We obtain the following flow constraint for the counter station:

$$I^T(M_{kernel-k} \circ \lambda \circ M_{OD})I = F_k \quad (2)$$

There are 90 counters inside of Paris. Then $k \in \{1,2 \ldots, 89, 90\}$ in (2). To avoid overestimating the number of the long-distance trips, the distribution of Euclidean distance between OD obtained from the survey data and GPS data should be respected. These distance constraints are considered. To search the correction matrix, we can formulate the problem as a linear optimization under constraints. To obtain correction factors λ most homogeneous possible, the criteria to minimize is the total distance between each correction parameter subject to the flow constraints in (2) and the distance constraints. The obtained correction matrix λ adjust the initial demand matrix M_{OD} by respecting the two constraints. Such estimation allows to compute the total travel trips inside of Paris by summing the adjusted number of trips. We apply the proposed method to the counter data for 09/2021. The estimated number of trips inside of Paris varies between 300,000 and 400,000 per day. These obtained results remain in the same order of magnitude as the aggregated results published by the recent EGT on the same period.

The second step is to use the estimated OD matrix to generate activity-based travel demand. We update first the weight of bicycle trips of the household travel survey according to the correction matrix λ. The updated weights enable EQASim to generate coherent number of bike trips inside of Paris. To distribute consistently these bike trips, we use the estimated OD matrix to attribute secondary locations.

The obtained activity-based bike travel demand fits correctly with the estimated OD matrix which is considered as the reference. There is a 10% error regarding between the generated bike travel demand and the estimated OD matrix. The next experiments will focus on i) improving the OD matrix estimation by taking into account additional data such as bikeshare data (station vélib), ii) extending the proposed methodology to other transport such that the entire activity-based travel demand is consistent with dynamic mobility data, iii) applying the estimated travel demand for different mobility simulation studies such as MATSim and SUMO.

REFERENCES

[1] Hörl, S., & Balac, M. (2021). Open synthetic travel demand for Paris and Île-de-France: Inputs and output data. *Data in Brief*, *39*, 107622.

[2] Feng, G., Jean, M., Chasse, A., & Hörl, S. (2021). Pre-calibration of a Discrete Choice Model and Evaluation of Cycling Mobility for Île-de-France. *Procedia Computer Science*, *184*, 172-177.

[3] Ziemke, D., Metzler, S., & Nagel, K. (2017). Modeling bicycle traffic in an agent-based transport simulation. *Procedia Computer Science*, *109*, 923-928.