Multistep Location Assignment for MATSim Demand Generation in a Regional Australian City with Integrated Local and External Trips

Sapan Tiwari[1,*], Afshin Jafari[1], Alan Both[1], and Dhirendra Singh[2]

[1]Centre for Urban Research, RMIT University, Melbourne, Australia
[2]Data61, CSIRO, Australia

[*]Corresponding author, Email address: sapan.tiwari@rmit.edu.au, Postal address: RMIT University, GPO Box 2476, Melbourne VIC 3001

Introduction

Activity-based models, such as MATsim (Horni et al., 2016), rely on synthetic populations that reflect the real-world population's demographic, spatial, and temporal characteristics. Among the key steps of population generation, location assignment plays a critical role in determining the realism of travel patterns and in generating realistic agent activity chains within activity-based transport models.

This study presents a comprehensive and multistep methodology for optimizing location assignment for the demand generation for the MATSim model of a regional city in Australia. While the previous studies have used gravity-based approaches for all activities, this study distinguishes between mandatory (work, education) and secondary (shopping, recreation) activities, acknowledging their unique behavioral patterns. By addressing both mandatory and secondary activity locations with distinct strategies, the framework enhances the spatial and behavioral realism of synthetic demand, contributing to more accurate and meaningful transport simulations.

Methodology

The methodology builds on the existing demand generation workflow by Both et al. (2021), which involves generating agents with demographic attributes, assigning them activity chains (Wang et al., 2021). The case study used is Greater Bendigo, a regional city in Australia. The activity chains used are aligned with the Victorian Integrated Survey of Travel and Activity (VISTA)[1] data, ensuring consistency with observed travel behavior. The location assignment process begins with generating home locations and allocating work and education locations as mandatory activities. Secondary activities are then inserted into the itinerary using a dynamic, gravity-based model. Figure 1 outlines the study's detailed methodology.

Mandatory activities are allocated using dedicated algorithms. Work locations are assigned using adjusted probabilities derived from journey-to-work (JTW)[2] data. This includes a combination of local (SA1)[3], global (SA3), and distance-based movement patterns. For each worker, the final assignment probability is a weighted function of these three travel patterns. This empirical weighting scheme ensures spatial diversity while maintaining alignments with the observed patterns (Agriesti et al., 2022; Casati et al., 2015). Adjustments ensure that underutilized locations are prioritized while reflecting overall movement trends.

[1]VISTA is an ongoing survey of household travel activity in Victoria, where all members of surveyed households are asked to complete a travel diary for a single specified day. `https://public.tableau.com/app/profile/vista/vizzes/#!/`

[2]https://www.abs.gov.au/articles/australias-journey-work

[3]Statistical Areas (SA) are sub-State regions classified under the Australian Statistical Geography Standard (ASGS), ranging from SA1 (smallest) to SA4 (largest).

Education location assignment balances proximity and school capacity constraints. Primary students are deterministically assigned to their nearest available school. A probabilistic model is used for secondary and tertiary students, where assignment likelihoods are based on inverse distances. An optimization problem minimizes total travel while respecting capacity limits. This approach reduces travel distances while ensuring feasible enrollment patterns. This dual strategy ensures both realism and equity in student distribution.
Secondary locations are allocated dynamically using a gravity-based model. The model calculates the probability of assigning a secondary activity to a particular SA1 based on its distance from already assigned locations (e.g., home or work) and its attraction value, determined by the land-use type (e.g., park, commercial). A distance decay parameter ensures that closer, more attractive destinations are preferred, and sequential assignment approach ensures that the final activity returns the agent near home, preserving spatial coherence in the daily itinerary.

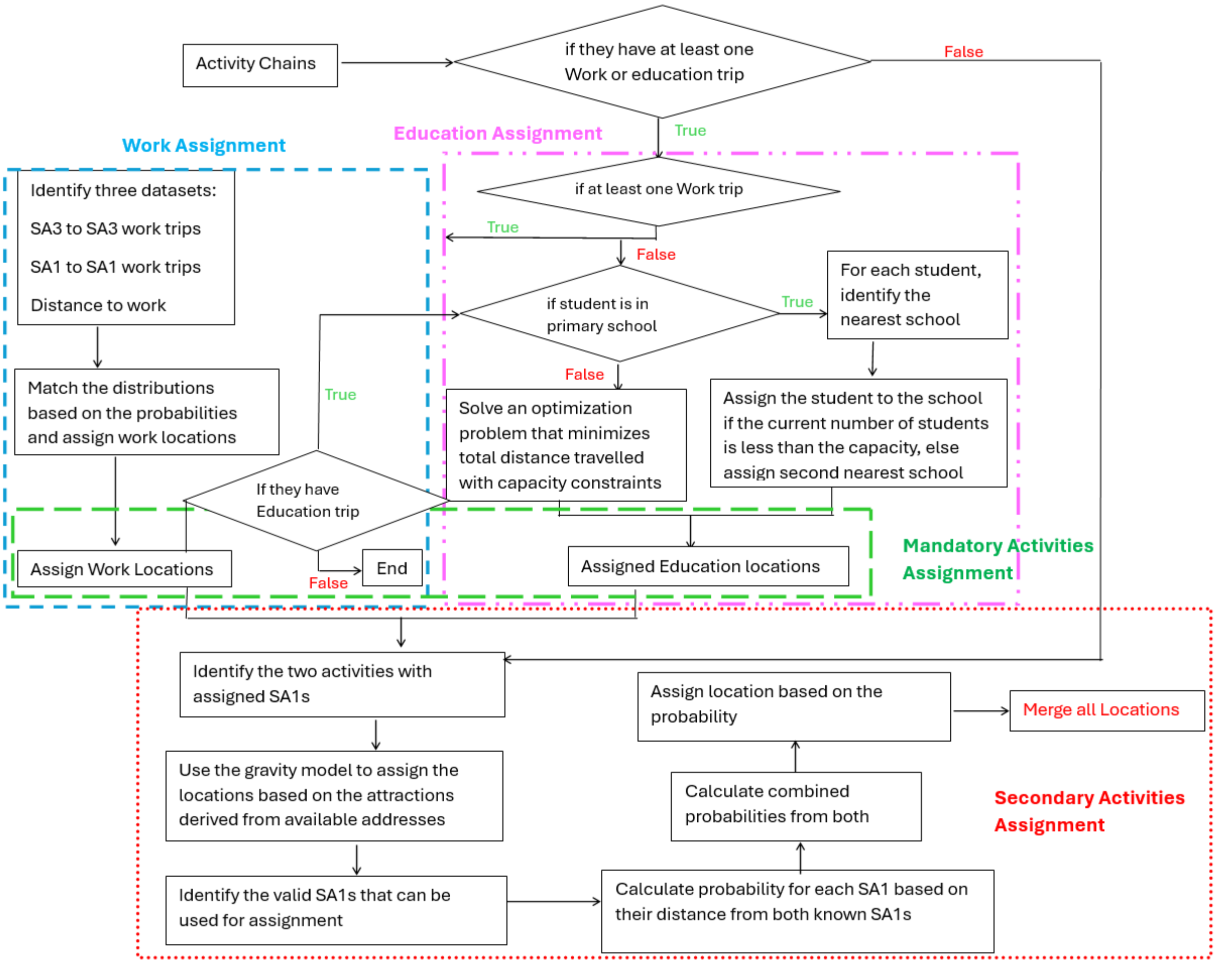

Figure 1: Schematic diagram of the detailed methodology

Results

The methodology produces locations for activities that align closely with observed spatial behavior in regional contexts. Secondary activities are allocated logically along an agent's path, with earlier activities leaning toward work locations and later getting assigned toward home. For example, if an agent has an activity between home and work, it is typically placed near the work location, especially when more activities remain in the chain. Conversely, when only one or two stops remain, other activities are increasingly assigned closer to home. These results confirm that the gravity-based approach, combined with sequential allocation, captures spatial decision-making behavior across diverse activity types. Figure 2a shows that work assignments show a high preference with most destinations located 2.5-3.5 km from home, reflecting realistic regional travel distances. Figure 2b shows that the model strongly aligns with real-world travel distances.

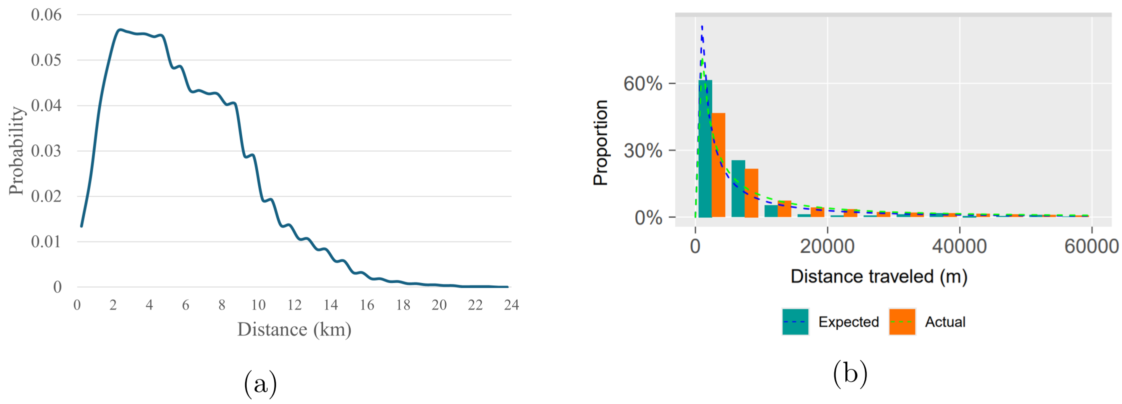

Figure 2: (a) Distance-based probability for choosing a work destination; (b) Expected and actual distance histograms for trips after activity location assignments.

Conclusions

This study introduces a multilevel methodology for optimizing location assignment in MATSim synthetic demand generation. The framework captures complex spatial behaviour by applying different mandatory and secondary activities strategies. Work and education locations are allocated using data-driven methods that reflect observed commuting patterns and capacity constraints. The gravity-based assignment of secondary activities incorporates attraction, distance, and the sequence of activity chains to ensure spatial realism. Together, these approaches result in realistic daily itineraries in a regional context. The proposed methodology is transferable to other regional contexts, especially those with sparse data and distinct spatial dynamics.

References

Agriesti, S., Roncoli, C., & Nahmias-Biran, B.-h. (2022). Assignment of a synthetic population for activity-based modeling employing publicly available data. *ISPRS International Journal of Geo-Information*, *11*(2), 148.

Both, A., Singh, D., Jafari, A., Giles-Corti, B., & Gunn, L. (2021). An activity-based model of transport demand for greater melbourne. *arXiv preprint arXiv:2111.10061*.

Casati, D., Müller, K., Fourie, P. J., Erath, A., & Axhausen, K. W. (2015). Synthetic population generation by combining a hierarchical, simulation-based approach with reweighting by generalized raking. *Transportation Research Record*, *2493*(1), 107–116.

Horni, A., Nagel, K., & Axhausen, K. W. (Eds.). (2016). *The multi-agent transport simulation MATSim*. Ubiquity, London. doi: 10.5334/baw

Wang, K., Zhang, W., Mortveit, H., & Swarup, S. (2021). Improved travel demand modeling with synthetic populations. In S. Swarup & B. T. R. Savarimuthu (Eds.), *Multi-agent-based simulation xxi* (pp. 94–105). Cham: Springer International Publishing.