

A Semi-Deterministic Approach for Modelling of Urban Travel Demand

Nam Huynh^{a*}
Nagesh Shukla^a
Albert Munoz^b
Vu Lam Cao^a
Pascal Perez^a

Abstract: This paper presents a methodology to construct travel related activity schedules for individuals in a synthetic population. The resulting list of activity schedules are designed as an input into a micro-simulator for urban transport dynamics analysis. The methodology involves two main steps. The first step generates a synthetic population based on census data sourced from the Australian Bureau of Statistics (ABS). The second step assigns individuals in the synthetic population activity schedules using Household Travel Survey (HTS) data related to the geographical area of interest (in this case, the Sydney Greater Metropolitan area). Each individual is assigned an ordered set of trips, travel purpose, travel mode, departure time and estimated trip time. The significance of the methodology is twofold in that it generates a synthetic population aligned with area demographics, as well as generating activity schedules that realistically represent how the population uses existing transport infrastructure. The methodology also preserves the inter-dependencies (in terms of the sequence, travel times and purpose of trips) of individual's daily trips, in contrast to many trip generators for transport micro-simulation purposes. A case study of Randwick area in southern Sydney is presented where the proposed methodology is applied. Case study data is validated against real world results and the scalability and applicability to other urban areas are discussed.

Key words: Travel diary; Synthetic population; Agent based modelling; Travel demand; Household travel survey data

I. Introduction

Of critical importance to efficient urban transport planning is an understanding of the interdependencies between populations and transport infrastructure. The daily activities of populations, where they go, using what transport infrastructure and why is a topic typically addressed through static, aggregate models to represent complex urban dynamics. These models are tasked with informing the policies that influence much of the transport

^a SMART Infrastructure Facility, University of Wollongong, NSW 2522, Australia.

^b School of Management and Marketing, University of Wollongong, NSW 2522, Australia

*Corresponding author: Nam Huynh nhuynh@uow.edu.au

<http://dx.doi.org/10.14453/isngi2013.proc.25>

infrastructure investments of an area. As such, these models need to incorporate the detailed interactions a heterogeneous population would have with existing or proposed transport infrastructure. In many cases, the models employed lack the depth to enlighten some of the hidden feedbacks transport policy may have on urban transport networks. A critical component for models of urban transport is the construction of individual-level activity schedules (travel diary) that, when aggregated, realistically represent population travel demand. Such activity schedules should comprise the sequence of trips each individual in the population makes as well as trip attributes such as travel mode, trip purpose, and departure time.

State-of-the-art models in travel demand modelling can be classified as trip based; tour based; and activity based. In trip based approach, each individual trips is modelled as independent and isolated trips *i.e.* no connections between the different trips. In the tour-based approach, trips start and end from the same location (home, work, etc) and are modelled as independent tours. As such, tour based approach lacks temporal granularity, and ignore inter-relations among independent tours. Over the past two decades, researchers have largely adopted activity based modelling to overcome these drawbacks, by deriving travel demand from the activities that individuals need or wish to perform¹⁻⁴. Activity-based approaches offer the advantage of incorporating spatial, temporal, transportation and interpersonal interdependencies (in a household) to model activity/travel behaviour. Furthermore, this approach reflects scheduling of activities in time and space and has been adopted in various operational land use and transport simulation models such as *ALBATROSS*⁴, *TRANSIMS*⁵⁻⁶, *AMOS*², *PCATS*¹.

The emerging field of activity-based models for travel demand modelling has received much attention in the literature^{4,7-10}. However, the majority of the activity based modelling methods presented are based on utility maximisation models¹¹, Markov models¹²⁻¹³ and rule based models^{1,2,4,14}. In utility maximisation models, a set of integrated tours form the basis for individual activity and travel demand modelling. For each tour, the choices of destination, mode and time of day are modelled as nested logit models with random utility maximisation objective functions¹¹. Others have employed Markov models (with activities as state) to estimate the daily, activity patterns, encompassing the interdependency of sequential activity types, timing and duration¹². On the other hand, rule based models execute the process of decision-making by using heuristic rules^{15,4}.

Approaches to date have used sets of decision making algorithms, such as determining activity patterns, travel time of day, activity durations, travel mode choice, etc. for assigning travel details to individuals in a synthetic population. Each decision making step is modelled based on statistical models (e.g. nested logit, Markov) or decision trees (e.g. classification and regression trees), which has an associated error term. As a result, the overall error rate of an activity list assignment is compounded as the output of a particular decision making step is the input to the next. Furthermore, these approaches consider individual level travel details assignment, ignoring the interdependencies that exist among individuals in a household. With a view to address these limitations, a single-step approach using household level semi-deterministic search method is proposed in this paper. The purpose is to assign travel diary to each individual in a synthetic population.

II. Synthetic population (SP) construction

There are two major approaches to generating synthetic populations, synthetic reconstruction and the combinatorial optimisation (CO). For in-depth reviews of each approach, interested readers are referred to the work of Huang and Williamson¹⁶, Ryan et al.¹⁷, Muller and Axhausen¹⁸, and Kurban et al.¹⁹. One issue that remains unaddressed in population synthesis is the incorporation of household resident relationships. Such a synthetic population would have to simultaneously synthesise the correlations between individuals and households against the real population in order to facilitate the collective decision making critical to agent based models. The value in these models lies in their ability to capture realistic behaviours of individuals in their interactions with infrastructure systems, and the subsequent value to urban policy design. For example, a household with a single parent with two children under 15 years old would have a considerably different transport need and behaviours than a married couple household with no kids.

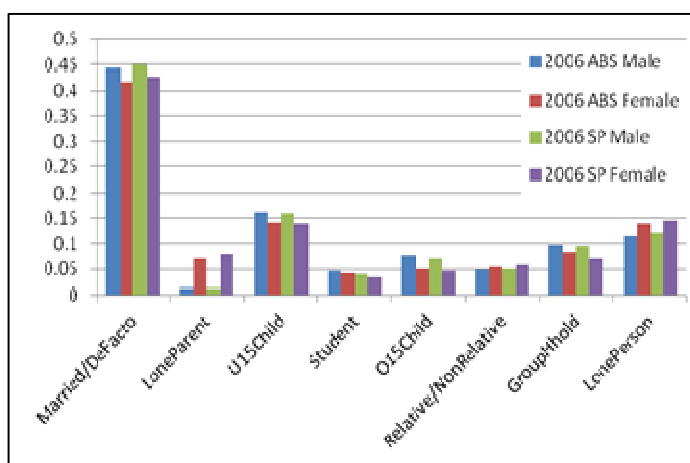


Figure 1. Percentage of males and females by household relationship of Randwick area in 2006.

In this study, a synthetic population is generated for agent based modelling purposes using a variation of the CO approach²⁰. Individuals are selected from an individual pool and allocated into households in a household pool to satisfy the distribution of household compositions in the study area. Each record in the individual pool represents an individual of the synthetic population and has four attributes, age, gender, household relationship and income. In contrast to the CO approach, the pool of individuals is instantiated from an aggregate data set representing the demographics distribution of the study area rather than extracted from an existent disaggregate survey data. The pool of households is instantiated from a different aggregate data set. Each record in this pool represents a household and has three attributes, number of males and number of females of the residents, and household type.

Using this algorithm, a synthetic population was constructed for Randwick area in Sydney using the 2006 ABS census data. In 2006, the area had approximately 106000 individuals living in around 47000 households. Visitors were not included in the synthetic population as they were not permanent residents of the study area. Figure 1 compares the proportion of household relationships of male and female individuals in the synthetic population of Randwick area against the original ABS data. Figure 2 compares the 17 household types in the synthetic population by number of households and number of residents against the original ABS data. Household types HF1 to HF16 are family households, distinguished by

the number of parents (i.e. ‘Married’/‘DeFacto’ individuals) and the number of children types (i.e. ‘U15Child’, ‘Student’, ‘O15Child’). Group household members and lone persons live only in households of type NF (non-family households). The correlations between these values validate that the methodology as one that can construct a realistic synthetic population for agent based modelling purposes that matches well with key statistics of the real population in the study area.

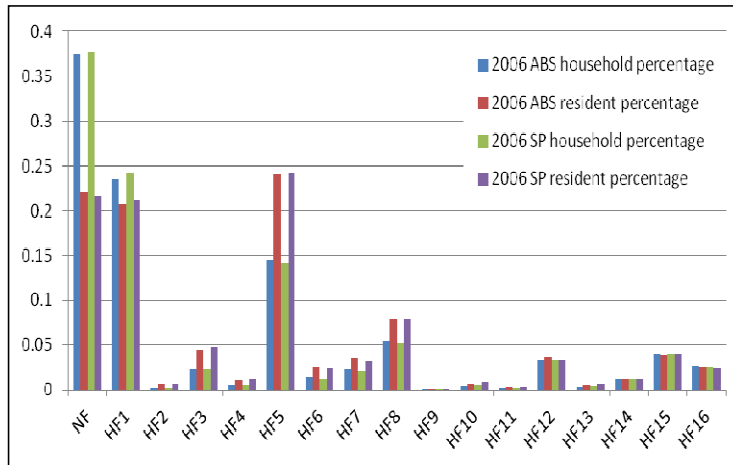


Figure 2. Percentage of 2006 household types in Randwick area.

III. Travel Diary Assignment to Synthetic Population

The household travel survey (HTS) data is the largest and most comprehensive source of information on individual travel patterns for the Sydney Greater Metropolitan Area (GMA). The data is collected through face to face interviews with approximately 3000-3500 households each year (out of 5000 households in the Sydney GMA randomly invited to participate in the survey). Details recorded include (but are not limited to) departure time, travel time, travel mode, purpose, origin and destination, of each of the trips that each person in a household makes over 24 hours on a representative day of the year. Socio-demographic attributes of households and individuals are also collected. The total number of trips included in the HTS data used in this paper is approximately 161000.

HTS data was used in this study to assign travel diary to individuals in each household in the synthetic population constructed. The method proposed for activity schedule assignment comprises two steps. The first deterministically searches in HTS data for households that best match the household type, the number of children under 15 years old, and the number of adults of a synthetic population household. This stage is described in steps 2, 7, 9, 10, and 11 in Figure 3. The deterministic search carried out in those steps gradually relaxes the constraints on exact matching of the number of children younger than 15 years old so that the search always returns at least one HTS household. The second step randomly selects a HTS household from the list of households identified in stage 1 and assigns travel diary to individuals in the HTS household to those in the synthetic household. The random selection follows a uniform distribution, see steps 3, 4, 5, 6, 8, and 12 in Figure 3.

At the conclusion of this process, each individual in the synthetic population has an activity schedule with a sequence of trips for a typical week day, as well as purpose, mode, departure time, and estimated trip time of each trip, totalling 509000 trips made in the synthetic population of Randwick area. Because the sampling

process

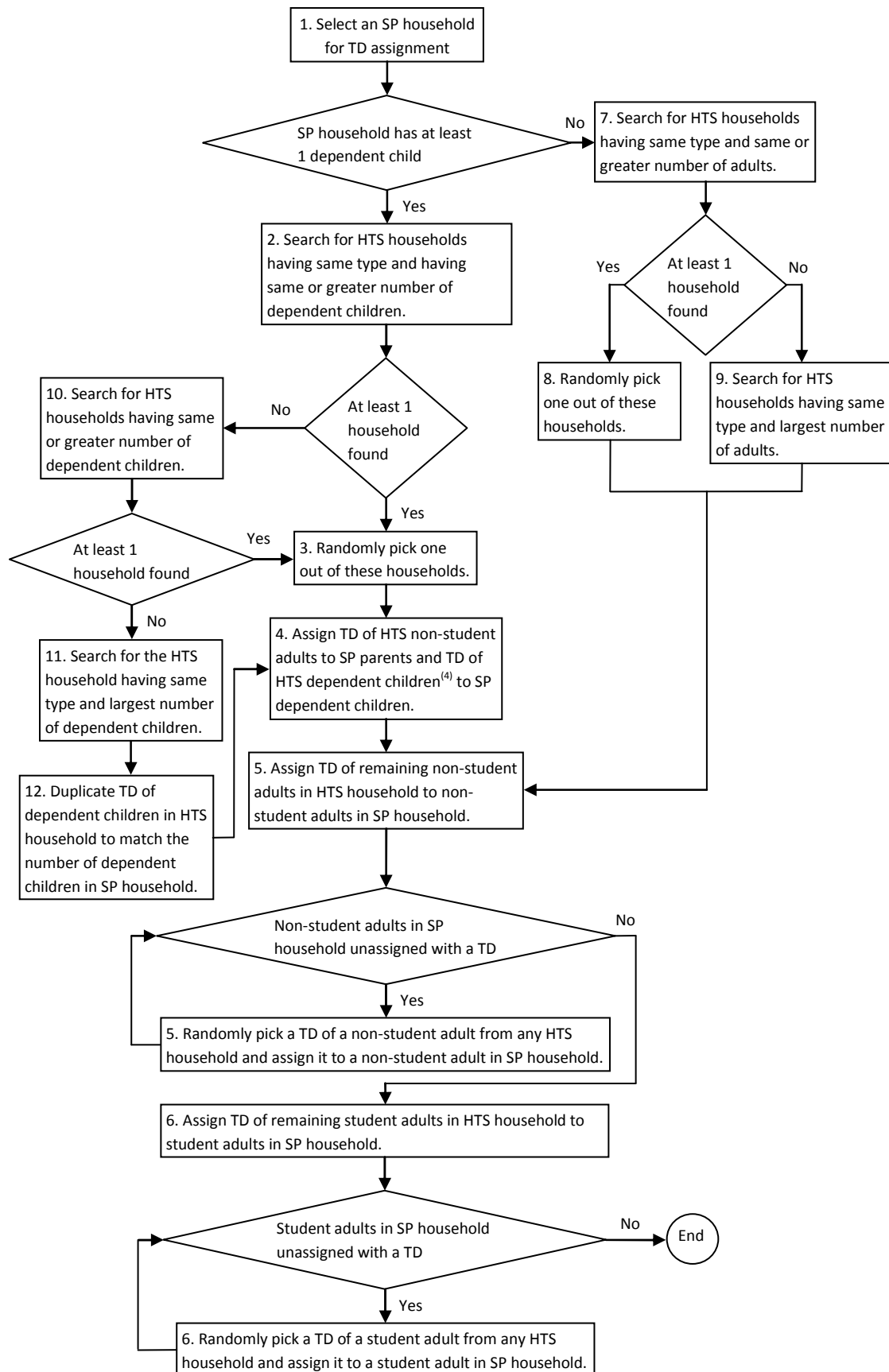


Figure 3. Operational design to assign travel diary from HTS data to synthetic population

from HTS data was carried out at household level, the inter-dependencies among individuals in a synthetic population household are preserved.

Activity schedules needs to realistically represent the patterns of travel demand of that area in order to be deemed suitable for input into traffic models. While no survey data is available to specifically detail the travel demand of Randwick area, Randwick synthetic population activity schedules were validated against HTS data of the whole Sydney GMA. Figure 4 compares the trip count proportions by trip purpose in synthetic population travel diary with HTS data. Figure 5 compares the proportion of trips counts by trip modes. Figure 6 compares the percentage of individuals in the synthetic population against that in HTS data by the number of trips made daily.

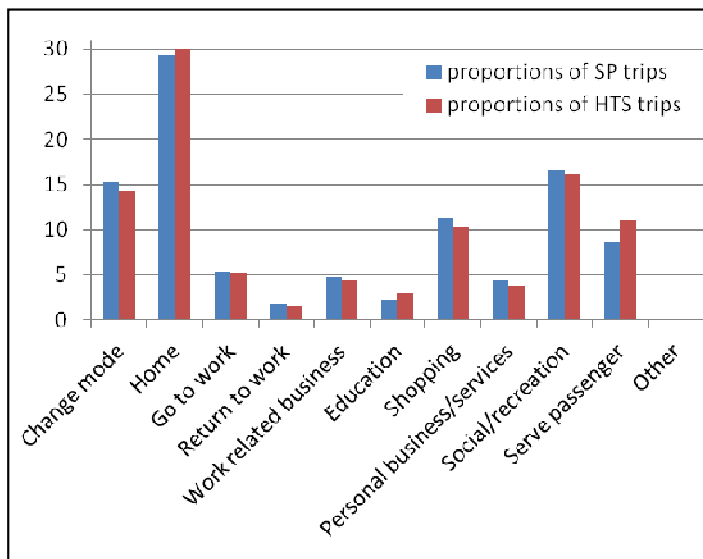


Figure 4. Proportion of trip counts by purposes

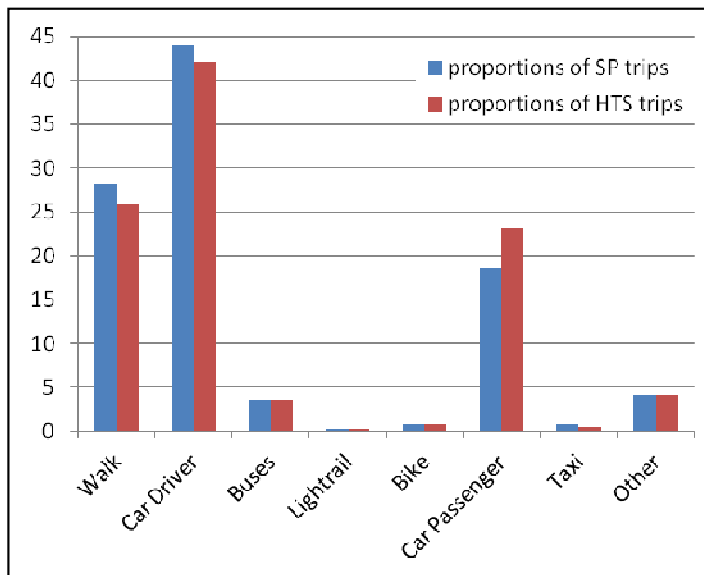


Figure 5. Proportion of trip counts by modes

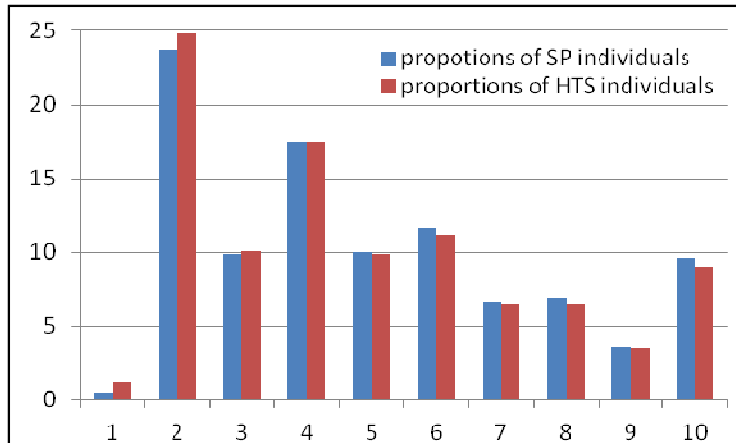


Figure 6. Percentage of individuals by number of trips made

These validations affirmed that the activity schedules generated using the semi-deterministic sampling method presented accurately mimic travel demand, satisfactorily reproduces the distribution of trip counts by purpose as well as the distribution of individuals by the number of trips made a day. The methodology also indicates that driving, walk and car passenger are the three dominant travel modes, consistent with HTS data. There are some deviations, which are attributed to various factors, such as the mismatches of distribution of household types and/or household compositions (i.e. type of individuals living in a household) in the synthetic population and in HTS data. For example, a lower proportion of children under 15 years of age exist in the synthetic population compared to that in the HTS data. This discrepancy would result in a lower proportion of car passengers.

IV. Conclusions

This paper proposed a generic methodology to generate a realistic activity schedules for a synthetic population. The methodology was applied to construct a synthetic population for Randwick urban area in Sydney and assign activity schedules to each individual in the population. Comparisons of key statistics of the synthetic population against 2006 ABS census data validate the suitability of the algorithm presented in constructing a realistic synthetic population for simulation modelling purposes. Validation against the HTS data for the whole Sydney GMA confirmed that the constructed activity schedules successfully reproduce the travel demand patterns in the Randwick urban area. More importantly, it preserves the inter-dependencies (in terms of the sequence, travel times and purpose) of daily trips of individuals in a synthetic population household, which many trip generators for transport micro-simulation purposes ignore.

The activity schedule of an individual as constructed by the methodology details the sequence of trips that individual makes in a day, as well as trip attributes such as travel mode, trip purpose, departure time and estimated trip time. Given specific locations for each of the trips, they provide complete inputs into a transport micro-simulator (e.g. Transims). The execution of such a micro-simulator provides a bird's eye view of traffic dynamics on a road network as well as actual travel times. This information is not only essential to assisting urban planning but a valuable input into an agent based model to simulate mode choice of the population.

Such a simulation model facilitates a more realistic prediction of future travel demands in the study area.

References

- ¹Kitamura, R., and Fujii, S., "Two Computational Process Models of Activity-Travel Choice", *Theoretical Foundations of Travel Choice Modelling*, edited by T. Gärling, T. Laitila and K. Westin, Elsevier Science Ltd, 1998, pp. 251-279.
- ²Pendyala, R.M., Kitamura, R., and Reddy, D.V.G.P., "Application of An Activity-Based Travel Demand Model Incorporating A Rule-Based Algorithm", *Environment and Planning B*, Vol. 25, 1998, pp. 753-772. <http://dx.doi.org/10.1068/b250753>
- ³Mattsson, L.G., and Weibull, J.W., "Effort-based probabilistic choice." Technical Paper, Department of Infrastructure and Planning, Royal Institute of Technology and Department of Economics, Stockholm School of Economics, Stockholm, 2001.
- ⁴Arentze, T., and Timmermans, H., "ALBATROSS: A Learning Based Transportation Oriented Simulation System", European Institute of Retailing and Service Studies, Eindhove: Eirass, 2000.
- ⁵Bush, B.W., "An algorithmic overview of TRANSIMS" [online database], URL: <http://transims.tsasa.lanl.gov> [cited 10 July 2013]
- ⁶Simon, P., and Nagel K., "Simplified Cellular Automaton Model for City Traffic", *Physical Review E*, Vol. 58, 1998, pp.1286-1295. <http://dx.doi.org/10.1103/PhysRevE.58.1286>
- ⁷Guo, J. Y., Srinivasan, S., and Sivakumar, A., "A Comprehensive Micro-Simulator for Daily Activity-Travel Patterns", Proceedings of the Conference on Progress in Activity-Based Models, Maastricht, 2004, 28-31 May (CD-ROM).
- ⁸Roorda, M., and Miller, E.J., "Strategies for Resolving Activity-Scheduling Conflicts: An Empirical Analysis", Proceedings of the Conference on Progress in Activity-Based Models, Maastricht, 2004, 28-31 May (CD-ROM).
- ⁹Vovsha, P., Bradley, M., and Bowman, J.L., "Activity-based Travel Forecasting Models in the United States: Progress since 1995 and Prospects for the Future", Proceedings of the Conference on Progress in Activity-Based Models, Maastricht, 2004, 28-31 May (CD-ROM).
- ¹⁰Pendyala, R.M., Kitamura, R., Kikuchi, A., Yamamoto, T., B.D., and Fujii, S., "FAMOS: Florida Activity Mobility Simulator", Proceedings of the 84th Annual Meeting of the Transportation Research Board, Washington, DC, 2005, 9-13 January (CD-ROM).
- ¹¹Huang, H. J., and Lam W.H.K., "A stochastic model for combined activity/destination/route choice problems", *Annals of Operations Research*, Vol. 135, 2005, pp. 111-125. <http://dx.doi.org/10.1007/s10479-005-6237-5>
- ¹²Ma, T. Y., Raux, C., Cornelis, E., and Joly, I., "Multi-State Non-Homogenous Semi-Markov Model of Daily Activity Type, Timing and Duration Sequence", Transportation Research Record: Journal of the Transportation Research Board, No. 2134, 2009, pp. 123-134. <http://dx.doi.org/10.3141/2134-15>
- ¹³Lee, B., Timmermans, H.J.P., Junyi, Z., and Fujiwara, A., "A Generalized Multi-State Hidden Markov Extension Model for Multi-State Activity Episode Duration", *Proceeding of the 11th International Conference on Travel Behaviour Research*, Kyoto, 2006.
- ¹⁴Ciari, F., Balmer, M., and Axhausen, K. W., "A New Mode Choice Model for Multi-Agent Transport Simulation", Proceedings of the 8th Swiss Transport Research Conference, Ascona, Switzerland, 2008, 15-17 October.

- ¹⁵Vaughn, K.M., Speckman, P., and Sun, D., “Identifying Relevant Socio-Demographics for Distinguishing Household Activity Travel Patterns: A Multivariate Regression Tree Approach”, NISS report, Research Triangle Park, North Carolina, 1999.
- ¹⁶Huang, Z., and Williamson, P., “A Comparison of Synthetic Reconstruction and Combinatorial Optimisation Approaches to the Creation of Small-Area Microdata”, Department of Geography, University of Liverpool, 2001, working paper.
- ¹⁷Ryan, J., Maoh, H., and Kanaroglou, P. “Population Synthesis: Comparing the Major Techniques Using A Small, Complete Population of Firms”, *Geographical Analysis*, Vol. 41, No. 2, 2009, pp. 181-203. <http://dx.doi.org/10.1111/j.1538-4632.2009.00750.x>
- ¹⁸Muller, K., and Axhausen, K.W., “Population Synthesis for Microsimulation: State of the Art”, *Proceedings of the 90th Annual Meeting of the Transportation Research Board*, Washington DC. 2011.
- ¹⁹Kurban, H., Gallagher, R., Kurban, G.A., and Persky, J., “A Beginner’s Guide to Creating Small-Area Cross-Tabulations”, *Cityscape: A Journal of Policy Development and Research*, Vol. 13, No. 3, 2011, pp. 225-235.
- ²⁰Huynh, N., Namazi-Rad, M., Perez, P., Berryman, M., Chen, Q., and Barthelemy, J., “Generating A Synthetic Population In Support of Agent Based Modelling of Transportation in Sydney”, *Proceedings of the 20th International Congress on Modelling and Simulation*, Adelaide, Australia, 2013.