Dat.mobility

ADVANCING ANALYTICS

VACANCY: aggregation of strategic transport model systems

INTRODUCTION

Strategic transport model systems are important decision tools to determine the impact of measures and to forecast the future usage of the mobility system. These model systems consist of separate choice models for the most influential behavioral choices that lead to mobility. Typically, choice models for trip/tour frequency per activity type, destination choice, mode choice and departure time choice are included.

Usually, discrete choice models are estimated on survey data (containing observed choices and person and household characteristics) and geospatial data using log likelihood minimization (see e.g.: Train, 2009). Explanatory variables consist of variables related to the origin and/or destination (e.g.: level of urbanization at the origin, the travel time between origin and destination, the number of jobs at the destination, etcetera) and/or variables related to the traveler (e.g.: age, household size, car availability, etcetera).

For each choice model, different person and household characteristics may be used as explanatory variables, depending on existence and significance of the relationship with the observed choices and data availability in both estimation and application contexts.

For each modelled area (typically the 13000 "CBS buurten" covering the Netherlands are used), data on each separate traveler related variable (called 'margins'; e.g.: the number of persons in a specific age class) are available as model input, but data on the combined variables (called 'person/household segments'; e.g.: the distribution across age/householdsize combinations) required for the choice models are not (mainly due to privacy concerns). Therefore, a population synthesizer (see figure

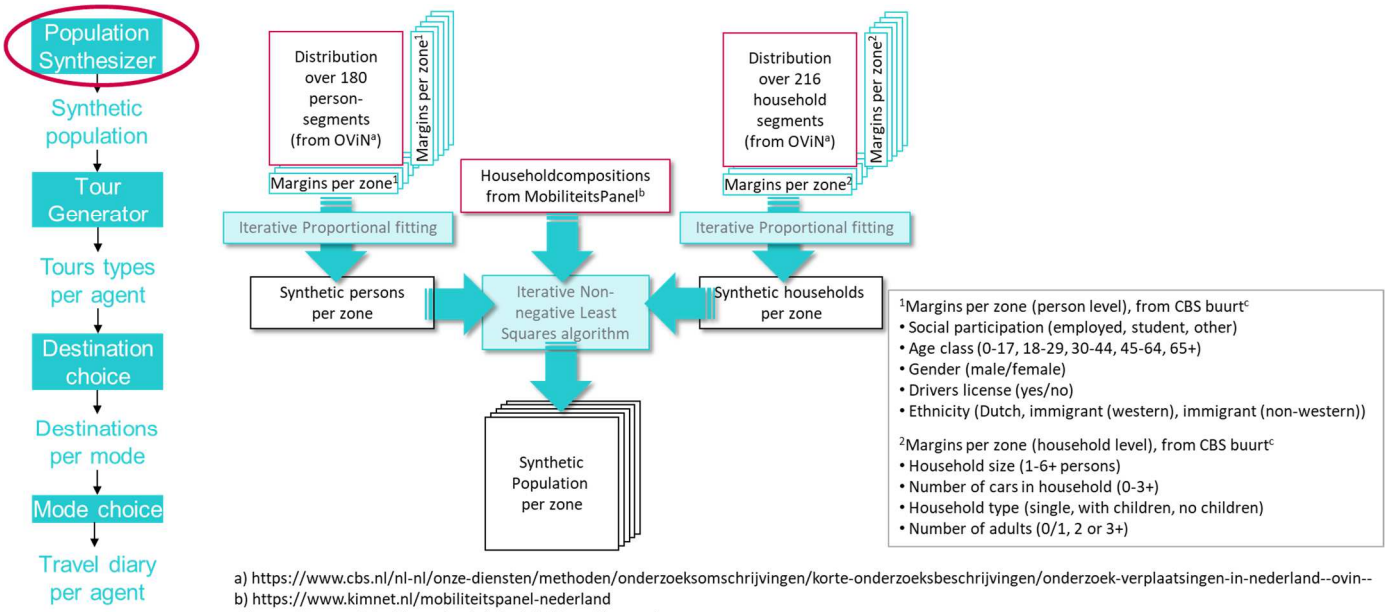
Snipperlingsdijk 4
7417 BJ Deventer
The Netherlands

P.O. Box 161
7400 AD Deventer
The Netherlands

+31(0) 570 666 222
info@datmobility.nl
www.datmobility.nl

VAT NL 0062 45 079 B01
CC. 2710 3813
IBAN NL59 INGB 0701 2168 08

below) is prepended to the choice models to generate the most likely distribution across segments for each modelled area. To do so, the population synthesizer matches margins for each area whilst adhering as much as possible to a known distribution across all person/household segments known from a representative sample.



a) https://www.cbs.nl/nl-nl/onze-diensten/methoden/onderzoeksomschrijvingen/korte-onderzoeksbeschrijvingen/onderzoek-verplaatsingen-in-nederland--ovin--
b) https://www.kimnet.nl/mobiliteitspanel-nederland
c) https://www.pdok.nl/introductie/-/article/cbs-wijken-en-buurten

Because most of the traveler related variables have several classes, the number of segments is relatively large. For example, in the latest travel demand models of Goudappel, there are 180 person-segments (derived from 5 age-, 2 gender-, 2 driver's license-, 3 ethnicity- and 3 social participation-classes) and 216 household segments (derived from 6 household size-, 4 car availability-, 3 household composition classes and 3 classes describing the number of adults in the household).

PROBLEM DESCRIPTION

Although the use of person and household segments yields more accurate models, their large numbers prohibit analysis and reporting on the segment level. Furthermore, the large number of segments causes relatively small numbers of travelers per segment per zone, which leads to high computational requirements for model system evaluation and, when the choice models are applied in microsimulation context, causes relatively large effects of statistical noise.

RESULT / OBJECTIVE

The objective is to adopt or develop a methodology that clusters the segmentation in such a way that the descriptive value of the aggregated model outcomes remains as close to the non-aggregated outputs as possible.
Initially, focus should be on adoption or development of a methodology that only clusters the output of the strategic transport model system, hence increase the suitability for analysis and reporting at the cost of reduction of the descriptive value due to the output aggregations.

If time allows, the methodology could be extended to allow for application on the model system itself (instead of only on its outcomes), which would decrease the computational requirements and statistical noise when applied in microsimulation context on top of the increase in suitability for analysis and reporting. This extension would come at an increased cost of reduction of the descriptive value, as the entire model system (not only its outputs) would be aggregated.

ASSIGNMENT

The assignment entails the adoption or development of an output clustering algorithm on the population synthesizer of strategic transport model systems. This requires literature research into clustering methods based on already estimated choice model parameters, choosing (a) suitable clustering method(s) or developing a new method and implementing it in prototypical form within the context of a strategic transport model system.

If time allows, the output clustering method could be extended into a model aggregation method that aggregates the population synthesizers input as well as the subsequent choice models within the strategic transport model system.
The added value of the output clustering method (increase in suitability for analysis and reporting) at the expense of the decrease of descriptive value should be demonstrated by comparing outcomes from a real world strategic transport model on the aggregated and non-aggregated levels.

Similarly, if time allows, for the model aggregation method, the additional reduction in descriptive value and the decrease in computational requirements and statistical

noise could be demonstrated by comparison to the outcomes of the output clustering method.

INFORMATION

When interested in this internship assignment please contact: Luuk Brederode at lbrederode@dat.nl. More information on Goudappel and Dat.mobility can be found via www.goudappel.nl and www.goudappel.nl/nl/dat-mobility.

Reference

Train, K., 2009. Discrete choice methods with simulation, 2nd ed. ed. Cambridge University Press, Cambridge ; New York.