Urban data fusion for the generation of an activity-based weekly mobility demand

Author(s):
Ordoñez Medina, Sergio A.

Publication Date:
2018-07

Permanent Link:
https://doi.org/10.3929/ethz-b-000268854

Rights / License:
In Copyright - Non-Commercial Use Permitted
Activity-based models have been developed in recent years as a response to the assumptions and simplifications of the of the mobility demand in classic aggregated models. With the increasing availability of disaggregated mobility information, large-scale activity-based demand can be generated with high accuracy. The main assumption is that people travel because they want to perform different activities at different locations and times. In these models, a virtual population represents people from an area of interest, and the objective is to estimate realistic sequences of activities for each of these persons. Hence, the complexity of this task depends on the period of time in which activities are scheduled. Moreover, the complexity grows exponentially with the length of the period of time. This is one of the reasons why the majority of modellers focus on daily activity patterns [1], [2], [3], [4].

Recently, multiday datasets have become more accessible [5], [6], and some researchers have turned their attention to multiday activity-based models [7], [8], [9], [10]. In this work, a methodology to generate an activity-based weekly mobility demand is proposed. The approach is based on the mental map structure proposed by [8]. In that model, activities are categorized as fixed or flexible, and two information structures are used to schedule activities: (i) a fixed activity skeleton, and (ii) a flexible activity agenda.

*A fixed activity skeleton* is a sequence of activities with a fixed start time, duration and location. As common people don’t plan all the activities of the week in this manner, leaving time for improvisation, these fixed activities only represent the structure of the schedule. In other words, after scheduling fixed activities, there will be periods of free time between them.

An *activity agenda* is a set of option activity types which depend on sociodemographic and spatial characteristics. Different types of people prefer to perform different types of activities. For each type, information necessary for the scheduling process is also included (i.e. when was the last time an activity of that type was performed, how often people perform that type of activities, when does these activities commonly starts and which is its typical duration).

To include location choice in the model, a third structure is proposed based on the definitions introduced by [11]: (iii) a set of evoked places. In a similar way than activity agendas, the evoked places are optional locations which depend on sociodemographic and spatial characteristics. This structure is very helpful to reduce the complexity of the location choice problem.

In conclusion, using their activity agenda and set of evoked places, agents schedule flexible activities during free time periods in-between fixed activities of the activity skeletons. With this approach, the complexity of the activity scheduling task doesn’t depend anymore on the total scheduling time (one week), but on the longest period of free time (it is normally less than a day).

To evaluate the model, this proposed methodology is applied to a synthetic population from Singapore. These five million agents were generated using census data, the national household interview travel survey of Singapore (HITS) and spatial datasets of the city. Details of this procedure can be found in [12]. Figure 1 illustrates the flow of information to assign weekly activity sequences to the agents of this population.
Figure 1: Information flows from urban datasets to weekly activity demand.

First, fixed activity patterns are extracted from HITS. These results are used to detect primary activities of frequent public transport users from public transport smart card data (CEPAS), and subsequently to recognize weekly patterns of fixed activities. Details of these processes can be found in [13]. After assigning sociodemographic characteristics to each public transport user according to zonal distributions, random forest classifiers are trained to assign weekly fixed activity skeletons to agents of the Singapore scenario. Figure 2 shows the comparison of start time and durations of primary activities by time of the day.

After assigning activity skeletons, flexible activity type and place type models are estimated using HITS. These models determine how likely is an agent to perform a certain flexible activity, and to travel to a certain type of place, according to sociodemographic and spatial characteristics (e.g. accessibility of the work location to shops). These results are employed to assign a flexible activity agenda and a set of evoked places to each agent of the synthetic population. More information of this step can be found in [14].

Finally, agents are put together in a large-scale simulation platform: MATSim [15]. Table 1 presents the characteristics of the simulated scenario. Agents start the simulation with incomplete plans containing fixed activities only. During a first weekly simulation, agents use the scheduling algorithm explained in [16] to plan flexible activities and generate trips between them. For a second iteration, a proportion of the agents are selected to start the simulation with incomplete plans and schedule flexible activities and trips again, while others follow the ones generated in the first iteration. By running the MATSim mobility simulation a hundred times, agents optimize their weekly activity schedules and trips by interacting with others in a limited transport supply.
Figure 2: Start time and Duration validation of assigned fixed activities.
Table 1: Characteristics of simulated scenario

Because of the assignment of activity agendas and sets of evoked places, flexible activities scheduled on-the-fly during the mobility simulation follow temporal and spatial distributions of activities reported in HITS as shown in Figure 3 and Figure 4.

Figure 3: Comparison of the spatial distributions of shopping and eating activities.
Figure 4: Duration and Start time validation of assigned shopping and eating activities.

On the computation side, the proposed model also presents good statistics. The iterative process needed a total 60GB of RAM to manage information of almost 400,000 agents. It includes fixed activity skeletons, one weekly plan in memory per agent, flexible activity agendas, sets of evoked places and experienced travel times, among others. The simulation took 1.5 hours for the first iteration and 30 minutes in average for the others. In 2 days and 3 hours, 100 iterations were executed with agents closed to reach User Equilibrium. In other word, the process is tractable and very competitive comparing other large-scale multi-day models.

In conclusion, a weekly activity-based model was developed using (i) a household travel survey with observations of different days of the week, (ii) multi-day public transport smart card data, and (iii)
urban spatial information (activity location facilities, sociodemographic characteristics of the population by zone). Thus, activity patterns are extracted from commonly available datasets and produces realistic weekly activity sequences. The method includes MATSim as final step to optimize the activity demand and simulate realistic trips. Accurate spatio-temporal distributions of fixed and flexible activities are generated with the model. The full paper will include travel time validations of trips by activity purpose and by day of the weeks. Besides, comparisons of the distribution of the number of public transport users by time of the week will also be included. The process is also tractable computationally, and it can be used for transport or land use planning.

Bibliography