# VirtualBelgium: a simulation platform for the Belgian population

**Johan Barthelemy** **Laurie Hollaert**
**Philippe Toint***
Namur Research Center for Complex Systems (NAXYS),
FUNDP-University of Namur, B-5000 Namur, Belgium
*Email: {johan.barthelemy}{laurie.hollaert}{philippe.toint}@fundp.ac.be

The VirtualBelgium project aims at developing understanding of the evolution of the Belgian population using simulation and considers various aspects of this evolution (demographics, residential choices, activity patterns, mobility, ...). This simulation is based on a validated synthetic population consisting of approximately 10,000,000 individuals and 4,350,000 households localized in the 589 municipalities (NUTS-5 level) of Belgium ([1]). The individuals' and households' basic attributes are respectively described in Tables 1 and 2.

| **Attribute** | **Values** |
|---|---|
| Gender | male; female |
| Age class | 0-5; 6-17; 19-39; 40-59; 60+ |
| Activity | student; active; inactive |
| Education level | primary; high school; higher education; none |
| Driving license ownership | yes; no |

Table 1: Individuals' characteristics

VirtualBelgium uses an agent-based methodology in order to simulate the mobility behavior and the evolution of the synthetic population. These 2 aspects are briefly introduced in the next sections.

| Attribute | Values |
|---|---|
| Type | single man alone<br>single woman alone<br>single man with children (and other adults)<br>single woman with children (and other adults)<br>couple without children (and other adults)<br>couple with children (and other adults) |
| Number of children | 0 to 5 |
| Number of other adults | 0 to 2 (mate not included) |

Table 2: Households' characteristics

# 1 Mobility behavior: activity-based models

The mobility behavior is modelised using an activity-based approach in which the travel demand is derived from the activities that the individuals need to perform. This approach has the advantage of reflecting the scheduling process of activities in time and space.

Consequently, an agenda (or *activity chain*) is assigned to the individual agents in addition to their basic attributes. This agenda consist of sequence of activities. Each activity is defined by three attributes described in table 3.

About 10.000 different activity chains patterns could be extracted from the Mobel mobility survey ([2]). The duration and distance travelled for each activity are also derived from the same data source. For instance, Figure 1 represent the cumulative distribution function of distance for the leisure activity. It is assumed that every activity chains begins and end at the individual's home. Moreover the total duration of an activity chains must be less than 24 hours. These assumptions seems fairly acceptable for a large majority of the population of interest. The algorithm implementing the activity chains assignments localisation consist then of a 3-steps procedures described hereunder.

## 1.1 Step 1: Generation of activity chains pattern

Let's denote by $I$ and $n_i$ the set of all individual type by $I$ and the number activity chains that could be extracted from the data $\forall i \in I$ respectively. Depending on the data a subset $J \subset I$, $n_j$ may be lower than a given threshold $t$ $\forall j \in J$. The first step is

| Attribute | Values |
|---|---|
| Purpose | pick up or drop someone<br>staying home<br>work related visit<br>work<br>school<br>eating outside<br>shopping<br>personnal reason<br>visiting relatives<br>walk-around<br>leisure activity<br>other activity |
| Duration | duration of the activity (min) |
| Distance | distance travelled to reach the activity |

Table 3: Activities' characteristics

then to add activity chains patterns to these problematic individual types. The missing chains are drawn from the ones associated with other individual type's belonging to $\mathcal{N}_j$, a neighbourhood of $j$.

### 1.2 Step 2: Activity chains assigments

Now that a pool of activity chains pattern is available for each individual type in $I$, the next step is to assign a chain to every individual in the population of interest. This is done by randomly drawing an activity chain using their empirical cumulative distribution function derived from the Mobel data.

### 1.3 Step 3: Duration and localisation of each activity

For each activity $a$ of every individual, its duration and localisation are randomly draw accordingly to their respective empirical distribution function for this particular activity.

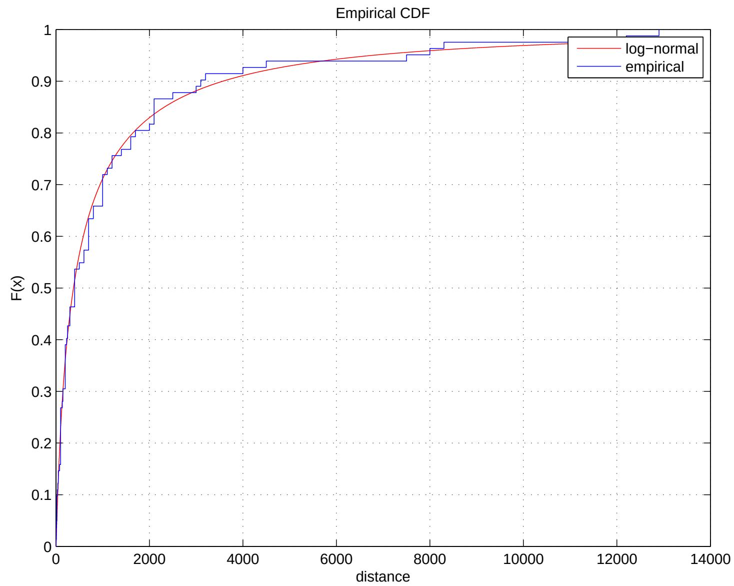


Figure 1: Cumulative distribution function of the distance (in meters) for leisure activities

# 2 Evolution of the synthetic population

The second major aspect of the simulation consist of the temporal evolution of the population. This aspect is a crucial step for dynamic micro-simulations as individuals and households change over time. For instance, individuals are aging, can get married or divorced, have children, and eventually die. All these changes can potentially have a significant influence on their behavior. In this paper we show how this population can be endogenously evolved as the model is running.

Since any modification of an individual agents attribute potentially affect his/her household, the dynamic evolution of the population is done at the household level. For every households in the simulation, the following steps are performed:

1. household individuals ages are incremented;
2. possibly add new babies to the household;
3. removing the dead individuals;
4. individuals activity status and education level are updated;
5. remove individuals from the household that are wishing to leave it, and create a new household for each of them;

6. splitting the household in two in case of the household divorcing;

7. getting the individuals married.

Note that the evolution time-step corresponds to one year and each sub-step corresponds to a single model. The results and the details of these models will be discussed in the paper.

One can easily see that the difficulty is to merge these various models into a consistent and modular *super-model*. The modularity requirement is crucial in the sense that the agents in VirtualBelgium may receive new attributes as new data becomes available, such as the one from the Beldam (the latest Belgian mobility survey conducted in 2010). Other models must also be easily added to the simulation, for instance the one developped in [3].

**Keywords** Synthetic population, agent-based simulation, mobility behavior, activity-based models, evolution of a synthetic population, microsimulation.

## References

[1] Barthelemy, J., Toint Ph.L., "Synthetic population generation without a sample", *Transportation Science*, to be published.

[2] Hubert J.P., Toint Ph.L., "La mobilité quotidienne des Belges", No. 1. *Mobilité et Transports* (Presses Universitaires de Namur, Belgium), 2002.

[3] Cirillo C., Cornelis E., Legrain L., Toint Ph. L., "A model of weekly labor participation for a Belgian synthetic population", *Network and Spatial Economics*, to be published.