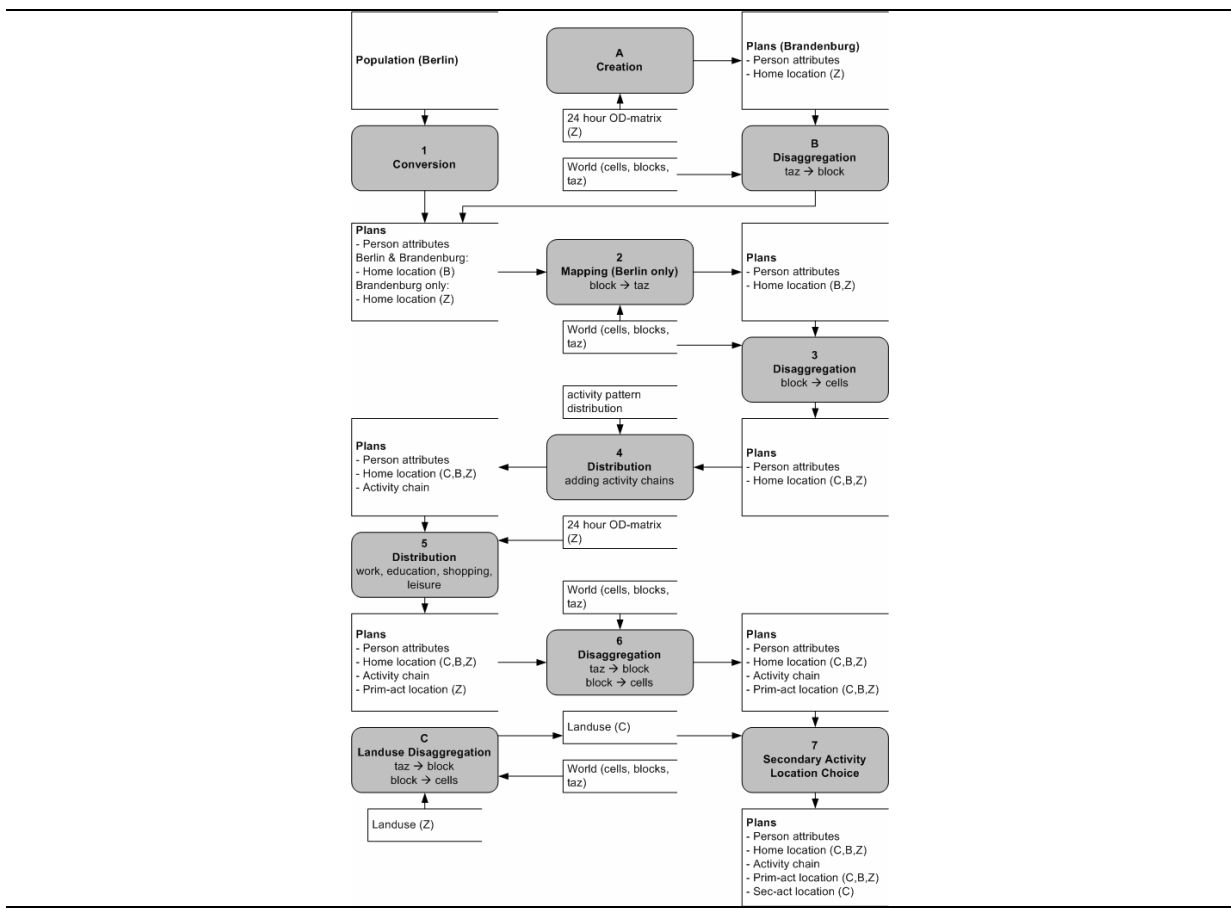


1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52



A Demand Generation Framework for Large Scale Micro Simulations

Words: 6250
 Figures: 4 (= 1000 words)
 Total: 7250

Michael Balmer, IVT, ETH Zurich
 Kay W. Axhausen, IVT, ETH Zurich
 Kai Nagel, VSP, TU-Berlin

Conference paper submitted to TRB 2006

A Demand Generation Framework for Large Scale Micro Simulations

Michael Balmer
IVT
ETH Zurich
Switzerland

Kay W. Axhausen
IVT
ETH Zurich
Switzerland

Kai Nagel
VSP
TU-Berlin
Germany

Phone: +41-1-633 27 80

Fax: +41-1-633 10 57

email:

balmer@ivt.baug.ethz.ch

Phone: +41-1-633 39 43

Fax: +41-1-633 10 57

email:

axhausen@ivt.baug.ethz.ch

Phone : +49-30-314-23308

Fax: +49-30-314-26269

email:

nagel@vsp.tu-berlin.de

July 2005

Abstract

Microsimulation tools are becoming increasingly important in traffic demand modeling. The major advantage over traditional assignment models lies in the fact that each traveler is simulated individually. This means, for example, that decision making processes can be included for each individual. The traffic demand is a result of different decisions made by each individual. Those decisions lead to plans which the individuals try to optimize. Therefore, such microsimulation models need appropriate initial demand for all the given individuals.

The challenge is now to create individual demand out of general input data. In practice there is a large variety of input data. They can differ in quality, spatial resolution, purpose, and so on. The challenge for a flexible demand generation framework is to combine this variety to produce individual demand. Even more, it has to define precise interfaces to provide portability to other models, programs and frameworks and it should be suitable for large scale scenarios holding many millions of individuals. Since also the model has to be adaptable to the given input data, the framework needs to be easily extensible with new algorithms and models.

The demand generation framework for large scale scenarios presented in this paper fulfils all these requirements. We also present the use of the framework on the basis of two different scenarios (greater area of Zurich - Switzerland and Berlin/Brandenburg - Germany).

Keywords

agent simulation methods – demand generation – individuals – framework – Transportation Research Board – TRB 2006 – Washington D.C.

Citation

Balmer, M., K.W. Axhausen and K. Nagel (2005) A demand generation framework for large scale micro-simulations, paper for the 85th annual meeting of the *Transportation Research Board, TRB*, Washington D.C., January 2006..

1. Motivation

Microsimulation is becoming increasingly important in traffic simulation, traffic analysis, and traffic forecasting (Vovsha *et al.*, 2002; Bowman *et al.*, 1999; Bhat *et al.*, 2004). Some advantages over conventional models are:

- Computational savings in the calculation and storage of large multidimensional probability arrays.
- Larger range of output options, from overall statistics to precise information about each specific synthetic traveller in the simulation.
- Explicit modelling of the decision making processes of the individuals.

The last point is important since it is not a vehicle which produces traffic; it is the person who drives it. Persons do not just produce traffic; instead each of them tries to manage his day (week, life) in a profitable way. They go to work to gain money, they go hiking for their health and pleasure, they visit their relatives for pleasure or because they feel obliged to do so, they shop to cook a nice dinner at home, and so on. Since not all of this can be done at the same location they have to travel, which produces traffic. To plan a day efficiently, many decisions are made by each person:

- Which route should I take to get to work? - *Route choice decision*
- Which mode should I use to go to the lake? - *Mode choice decision*
- Should I drink another beer before going home? - *Activity duration choice decision*
- Should I go shopping near my home or at the mall? - *Location choice decision*
- When should I do sports today? - *Activity starting time choice decision*
- Should I go to visit my friend? - *Activity choice decision*
- Whom should I take along? - *Group composition decision*
- Should I go swimming before or after work? - *Activity chain choice decision*

There are many more decisions to make; some of them are made hours (days, months) in advance while others are made as spontaneous reactions to specific circumstances. Many decisions induce other decisions. For example, if I am late for work, I am supposed to work longer, so there is no time left to go shopping today, so I need some time tomorrow to do the shopping. This example shows the importance of describing schedules for each individual in a simulation model, because it is the schedule and the decisions made by the person who adhere to this schedule that produces the traffic.

1 To simulate a typical day in an urban area, microsimulation tools need precise information
2 about the schedule of each individual and some knowledge about people's decision making
3 process. The challenge is to create this individual demand out of general input data. In
4 practice there is a large variety of input data. They can differ in quality, spatial resolution,
5 purpose, and so on. The challenge for a flexible demand generation framework is to combine
6 this variety to produce individual demand. Even more, it has to define precise interfaces to
7 provide portability to other models, programs and frameworks and it should be suitable for
8 large scale scenarios including many million of individuals. Since the model should be
9 adaptable to the given input data, the framework needs to be easily extensible with new
10 packages, algorithms and models.
11
12
13

14 This paper presents such a demand generation framework for large scale scenarios. After a
15 summary about other work (Section 2) and describing the program structure (Section 3) the
16 framework is used to generate daily demand for two different scenarios: One—medium
17 resolution scenario—takes place in the greater Zurich area (Kanton Zurich) of Switzerland
18 and consists of about 1.3 million individuals (Section 4). The second—large scale scenario—
19 is defined for Berlin and Brandenburg of Germany for about 7 million inhabitants (Section 5).
20 The two scenarios differ in the amount of available information, in spatial resolution, in size
21 and—last but not least—also in the demand generation process itself.
22
23
24

25 Section 6 gives a short overview about the further use of the generated daily demand. For that,
26 an iterative, large scale microsimulation model is used (MATSIM, 2005). The paper finishes
27 with a summary and outlook to future work.
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52

2. Other work

The work presented in this paper falls into the area of activity-based demand generation (ABDG). There exist already a fair number of AGDB packages (e.g. Vovsha *et al.*, 2002, Bowman *et al.*, 1999, VISEM, 2005, Pendyala, 2004, Bhat *et al.*, 2005, Arentze *et al.*, 2000). Internally, those packages are structurally somewhat comparable to what is described in this paper, and in terms of methods they are in fact considerably more sophisticated. Yet, their output is typically in terms of (time-dependent) origin-destination matrices, to be fed into dynamic traffic assignment (DTA) models (e.g. DYNAMIT, 2005, DYNASMART, 2005, Friedrich *et al.*, 2000), an important exception being TRANSIMS (2004), which is, however, not easily obtainable. In addition, the input data requirements of most of those packages are geared towards typical data availabilities in the U.S.

In contrast, in the work presented in this paper, the intent is to present an approach that is fully traveller (agent) oriented on all levels, and adapted to data availability in Switzerland and Germany. Although most of the above-mentioned packages could in principle be adapted towards those needs, in practice this would require source code change, meaning that these changes cannot be implemented by users of the package alone. Therefore, at this point we design and implement placeholder modules that do something plausible based on the available input data, until better approaches become available. Those better approaches could be in the form of modified versions of the above-mentioned or other packages, or in the form of enhancements of our own code.

Other approaches are the agent-based land use models, for example URBANSIM (Waddell *et al.*, 2003), ILUTE (Salvini and Miller, 2003), or the models by Abraham/Hunt (e.g. Hunt *et al.*, 2001). These models face similar difficulties as the approach described in our paper, in particular the need to assemble a consistent agent-based view of the world from diverse data sources (Abraham *et al.*, 2005).

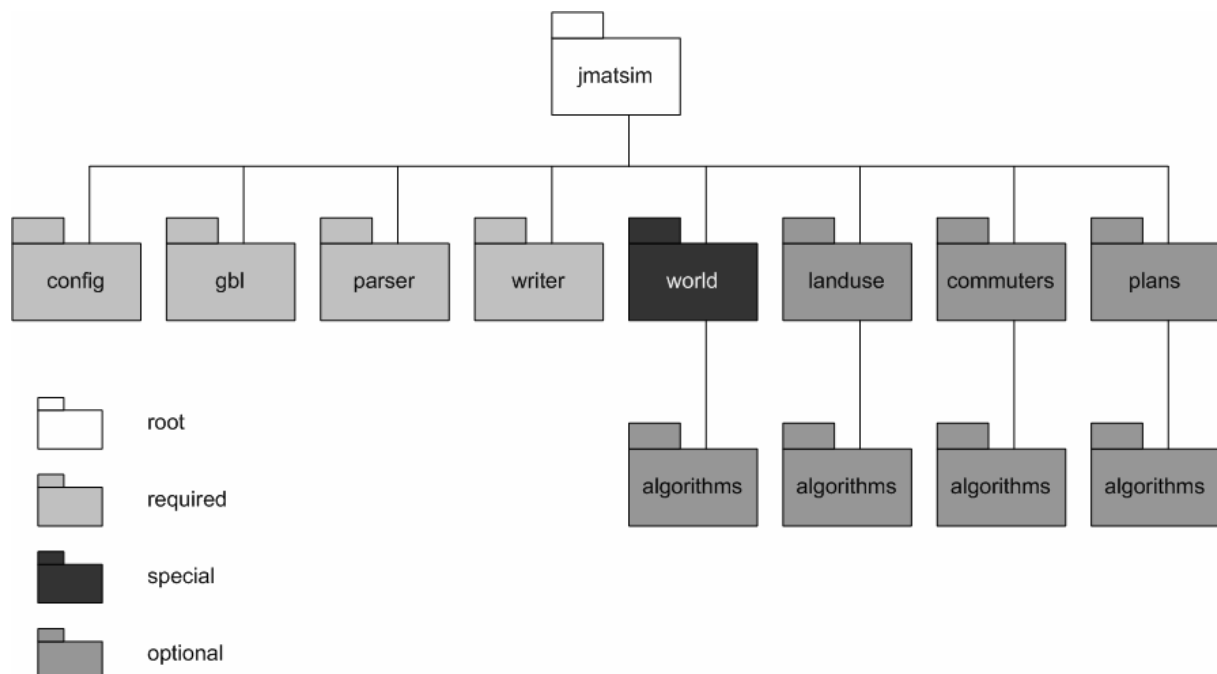
In the longer run, it would be useful to have a “plug-and-play” approach between these different models, i.e. where modules between different modelling systems can be interchanged and coupled in arbitrary ways. In fact, at least one such effort is under way (Waddell *et al.*, 2005). However, the software engineering challenges for such a project, despite considerable progress in the last decade, are still enormous, in particular for large scale scenarios (Nagel and Marchal, 2003). Therefore, for the time being it seems more productive to make progress in smaller steps, which in our case means the implementation of a general framework, demonstrating the overall feasibility and highlighting potential problems. This is what is done in this paper.

3. Framework

The demand generation process for each individual modelled in a given scenario is highly dependent on the available input data. The more precise the data is, the fewer methods or algorithms are needed to aggregate parts of the demand of an individual; i.e. if we have the complete information about what people are doing, it is just a simple conversion to describe it in an appropriate data format. On the other hand, we do not have access to all needed information because of data privacy, imprecise or aggregated data, costs of the required surveys and limitation of census data.

Therefore, a demand generation framework (see Figure 1) has to be flexible enough to handle this variety of input data. In other words, it has to provide interfaces for the different data in an appropriate data format; it has to allocate an appropriate data structure in memory such that each data point can be accessed in a feasible amount of time and it needs to provide uniqueness of data points while creating, manipulating and deleting those.

Figure 1 package structure of the demand generation framework



3.1 General framework structure

The framework consists of necessary, optional and a special “world” package (see Figure 1). The necessary packages provide general XML parsing and XML writing interfaces, global constants and a global and unique data structure for accessing input parameters. Optional packages are always based on a specified input written in XML format (MATSIM, 2005). The base functionality of each optional package is to read in the defined XML data, store them in an appropriate data structure and write it again enriched, reduced or even unchanged in the XML data format.

3.2 The framework packages in details

3.2.1 The “parser”, “writer” and “gbl” package

The “parser” package provides a general base class for parsing XML data via the SAX parser (SAX, 2005). Respectively, the “writer” package is the base for writing XML files. The purpose of the “gbl” package is to hold global constants and globally accessible functions.

3.2.2 The “config” package

In the whole framework exactly one configuration data structure exists which holds all required input parameter for a specific demand generation process. Typical parameters are locations of input data, different file formats, special function parameters, and so on. The information stored in this package can be accessed from every part of the framework (global singleton design pattern).

3.2.3 The “world” package

The “world” has a special functionality. It describes the region for which the demand is to be generated. Therefore, it is required to guarantee that only one world exist during the demand generation process. This package holds all information about cells, blocks, zones, municipalities, etc. which are modelled as primitive shapes. If the other optional packages refer to such shapes, they have to point to them in the “world”. With this, uniqueness of each shape object is guaranteed. Even more, it is also a control mechanism to verify if other input data is consistent with the given scenario.

Another, quite important functionality of the world package has to be mentioned, too. During the demand generation process, the resolution of the world can differ. For example, the landuse information used might be based on a raster of 100x100 meter cells, while a given

1 commuter matrix is based on traffic analysis zones. Therefore, it is necessary to
2 (dis)aggregate one data set into another world resolution. The world holds this mapping for an
3 arbitrary amount of resolutions of the same region. With it, on the fly (dis)aggregation is
4 provided.
5

6 While this package plays a central role in the framework, it is not a required package. The
7 reason for that lies in the fact that there can be demand generation steps where no spatial
8 information is needed (i.e. generating departure times of activities, assigning activity chains to
9 individuals, etc.).
10

11 **3.2.4 The “landuse” and “commuters” package**

12 Those optional packages store information for the schedule generation process. Dependent on
13 the input data those packages are used i.e. to find locations of work activities, to create sets of
14 leisure facilities, and so on.
15

16 **3.2.5 The “plans” package**

17 This is—in principle—the core of the demand generation. The data structure is based on the
18 MATSIM schedule DTD (in MATSIM called a plans-DTD) which also can be accessed via
19 MATSIM (2005) and it is used as a “working file”. The reason for this choice lies in the high
20 flexibility of the MATSIM schedule file (MATSIM plans file). In the minimum version the
21 file only holds the identity number (id) of all simulated persons. On the other hand, it is
22 possible to add a large amount of information about each person, such as age, sex, car
23 ownership, home and work location (at different world resolutions), etc.
24

25 Additionally each person can hold one or many different individual schedules, describing
26 when a person wants to start an activity, where it will be performed, which route and mode to
27 take to go from one location to the next, and so on.
28

29 The internal data structure of the “plans” package provides exactly the same flexibility as the
30 XML file format. Therefore, it is possible to add sequentially more and more schedule details
31 to a given incomplete MATSIM file.
32

33 **3.3 Demand generation algorithms**

34 Algorithms can be added to each optional package (and the world package) to verify,
35 manipulate, add or delete data items according to the purpose of the algorithm. Since for each
36 new scenario different algorithms have to be used or implemented, it is of great interest that
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52

1 algorithms should be clearly separated from the data structure. They also should be easily
2 exchangeable by others. The order in which algorithms are called should be flexible as well.

3
4 The algorithms are collected into a sub package of that data structure which they manipulate.
5 Therefore, an algorithm at “jmatsim.commuters.algorithms” works on data in the
6 “commuters” data structure (see below for examples).
7

10 **3.4 External models**

11
12 At any point in time during the demand generation process, the framework allows one to store
13 all data into well defined XML data files. They do—per definition—respect the format as
14 described in the underlying DTDs. Therefore, a clean and stable interface to third party
15 programs and models is available even when those models are not part of the demand
16 generation framework.
17

18
19 For the two following case studies described in Section 4 and Section 5 we will use for
20 example the *Secondary Location Choice Module* (Marchal and Nagel, 2004) (in its adapted
21 version described in Balmer *et al.* (2005), which chooses locations for secondary activities in
22 such a way that each agent improves its daily chain. Primary and secondary activities are
23 defined such that in each given activity chain there is exactly one (well defined) primary
24 activity while the others are defined as secondary activities (details in Balmer *et al.*, 2005).
25 The module uses the same DTD, allowing simple information interchange.
26
27
28
29
30

31 **3.5 Computational Issues**

32
33 One important issue for demand generation is the amount of input information needed.
34 Because of the variety of possible demand generation algorithms and input sources it is
35 necessary to have fast access to that data. One simple and very fast solution is to load all
36 information into memory and to provide a hierarchical data structure (tree structure) to access
37 any item from any other location in $O(2 * \log(n))$, where n is the depth of the data tree.
38
39

40
41 The hierarchical data structure is already provided by the input data (XML format), but the
42 available space in the memory might not be sufficient. While the description of the world, the
43 landuse data information, commuter matrices, etc. typically hold a relatively small amount of
44 data, the amount of information for the individuals goes far beyond the size of a typical
45 memory capacity (on average around 1-2 GB of memory).
46

47
48 The demand generation framework therefore uses the idea of sequential individual demand
49 generation (streaming of individuals). In other words, the framework reads in one individual
50 at a time, runs the defined algorithms for it, writes the results to file and frees the memory. In
51
52

1 this way, the number of individuals of the given scenario is unlimited. This idea will still
2 work if—instead of single persons—we will generate demand at the household level for a
3 small number of persons. But the limit of this approach will be reached if we also want to add
4 the concept of social networks (Axhausen, 2005). In this case demand generation of one
5 individual can—in principle—depend on all other individuals in the scenario and therefore the
6 whole population must be stored in memory.
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52

4. Demand Generation for the Kanton Zurich, Switzerland

Using the framework described in Section 3 we present the steps taken to generate daily demand for the Kanton Zurich, Switzerland. The following subsections describe briefly which input data were used and which algorithms were employed to generate the daily demand.

4.1 Data Resources

The *world* describes the region (Kanton Zurich of Switzerland) at two different resolutions:

- 170 municipalities and 12 additional districts inside the city of Zurich (municipality level denoted as M)
- Raster of 100x100 meter cell resolution (raster level denoted as R). In total 167881 cells are given.

The mapping between those two levels is also available. The mapping rule is non-ambiguous. Each raster cell belongs to none or one municipality while each municipality holds at least one raster cell. Even more, the municipalities hold only those cells which include urban areas, e.g. lake “cells” are not part of the mapping.

Information about *population* distribution in the municipalities is generated by Frick (2004). It holds about 1.3 million inhabitants with the following data:

- Home location (R)
- Population group (children, worker, non-worker, senior)
- Mobility (car, season ticket ownership, bike, walking)
- Age
- Sex

The Swiss Federal Statistical Office, BFS (2005) provides *landuse* data holding information about capacities of different activity types like “work”, “shopping”, “education”, etc. That information is also based on raster level (R).

The “Pendlermatrix 2000” (Vrtic and Axhausen, 2003) holds the information about work and education *commuters* at municipality level (M).

The microcensus (ARE, 2001) is a periodic survey of the travel behaviour of the Swiss population. It has been run every five years since 1974. The microcensus is carried out by the Federal Office for Spatial Development, ARE in cooperation with the Swiss Federal

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52

Statistical Office, BFS and contains detailed information about the mobility behaviour of 30'000 persons from all over Switzerland (Axhausen and Frick, 2005). About 1670 different *activity chains* can be found in Micocensus2000. Most of them appear very rarely, so only the 100 most frequently occurring activity chains are used later, which still cover more than 90% of all days. By reducing the number of activity types to “home”, “work”, “education”, “shop” and “leisure” some activity types like “service” were recoded to “work”. It ends up by 21 different activity chains (more details see Balmer *et al.*, 2005).

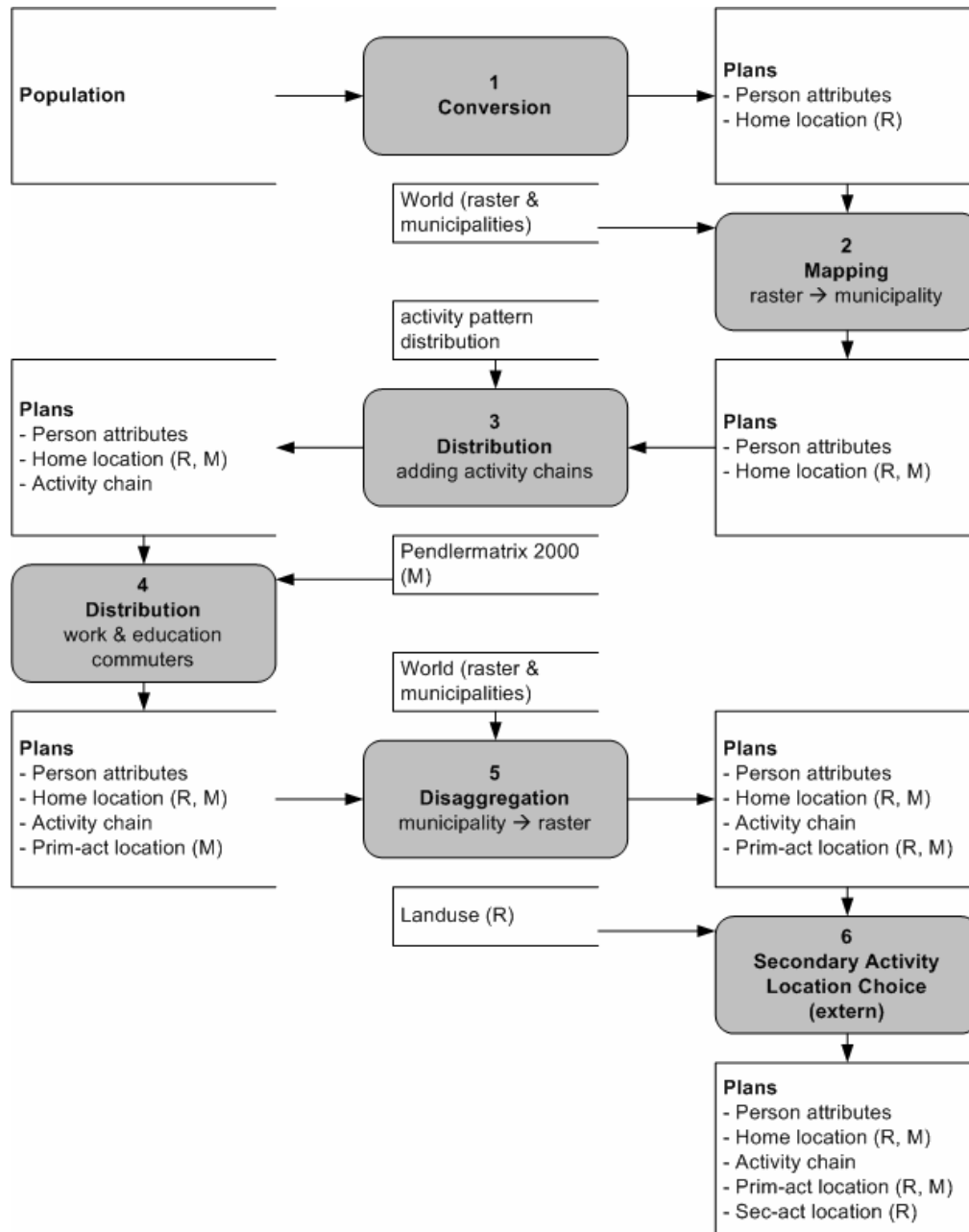
4.2 Demand generation process steps

The demand generation process is split into six sequential steps (see Figure 2 for an overview). Each process step (except the first one) uses one specific data resource to add details to each individual schedule:

1. *Conversion*: This process step just converts the input population file into the XML person description file. None of the person tags holds a plan yet. But additional attributes like age, sex, etc are included. It also holds the information in which landuse raster element (R) this person lives.
2. *Mapping*: Each raster element of the land use data belongs exactly to one municipality. By using this mapping, each agent can be assigned to the municipality of his home location.
3. *Distribution*: Given the distribution of the activity chains described above, we assign one of the chains to each person according to the given distribution. We also respect the fact that i.e. children do not go to work, therefore persons of young age are not allowed to hold an activity chain including a “work” activity.
4. *Distribution*: The “Pendlermatrix 2000” (see above) holds the information about work and education commuters at municipality level of detail (M). With the assumption about primary activities given in Balmer *et al.* (2005) we are able to add the locations of the primary activities “work” and “education”. Unfortunately we do not have similar data for the primary activities “shop” and “leisure”. As long as no better data or no appropriate algorithm is available we just assume that those activities are undertaken in the same municipality (district) where the person lives.
5. *Disaggregation*: The process step 6 uses the location choice (see Section 3.4) which is working at the raster level of detail (R). Therefore, we need to disaggregate the locations of the secondary activities to that level. It is done by uniformly picking one of the raster elements of the given municipality (district).

6. *Secondary activity location choice (extern)*: The module of Marchal and Nagel (2004) is used as an example of an external program (Section 3.4) to add the missing location of the secondary activities.

Figure 2 Process plan for generating demand of Kanton Zurich



The final schedule of each person in the scenario now describes the following:

- Person attributes
- Which activities will be performed
- In which order the activities will be performed
- Where the activities will be performed

There are some final algorithmic steps missing in that generation process: There is no information about the times when an activity should be performed and how long it should take. We also still do not know which mode of transportation will be used and which route to take to go from one location to the next. These steps are left out willingly because we will use the MATSIM iterative optimisation process (Raney and Nagel, 2004) to create that information. A short overview about this process will be given in Section 6.

Note that, it is also possible to write appropriate algorithms to create that information as a part of the framework.

5. Demand Generation for the Länder Berlin and Brandenburg, Germany

The scenario Berlin and Brandenburg does have some small differences compared with the Kanton Zurich scenario, due to the fact that only other less precise information is available.

5.1 Data resources

The *world* describes the region of Berlin and Brandenburg (RBS, 2000). In this scenario tree different spatial resolutions are used:

- The region is split up into 1008 traffic analysis zones TAZ (zone level denoted as Z)
- Each zone consists of several Blocks. 12260 blocks are available for this scenario (block level denoted as B)
- A raster of 169420 cells is also available. Each cell has an extent of 500x500 meters (raster level denoted as R)

The mapping rules are the following: Each traffic analysis zone consists of at least one block. Each block belongs exactly to one traffic analysis zone. A cell refers to none, one or many blocks while each block refers to at least one cell.

The raster is synthetically generated and therefore, there are no specialities encoded like in the Kanton Zurich scenario (i.e. only urban cells). Note that the generation of the raster and the mapping is also done by the above described framework using a special “world algorithm” in the “*jmatsim.world.algorithms*” package.

Only the *population* of Berlin itself was available at that time consisting of about 3.3 million individuals (RBS, 2000). The data was not that extensive as for the population of Kanton Zurich. The only two available items are:

- Home locations (B)
- Age

The *landuse* of Berlin and Brandenburg is based on traffic analysis zone level Z. It consists of capacities for “home”, “work”, “education”, “shopping” and “leisure” activities.

For this scenario there is no *commuter* matrix available. On the other hand a 24 hour origin-destination matrix of Berlin and Brandenburg is given (SfSB, 1998). We use this matrix in two different ways. First, for all trips starting in Berlin we use this information to guess where individuals of the corresponding zone are performing their primary activity. In other words,

1 we use it in a similar way as the commuter matrix of the Kanton Zurich scenario. The
2 remaining trips starting in the traffic analysis zones of Brandenburg are used to create the
3 missing part of the population and also to define the location of their primary activity. This
4 is—of course—just a very vague way to interpret an origin-destination matrix, but still
5 feasible to close the information gap.
6

7 We use the same *activity chain distribution* as in the Kanton Zurich scenario. Those chains
8 surely reflect average Swiss behaviour and not necessarily the one from Berlin or
9 Brandenburg, but it is at least a feasible start as long as no other data is available.
10

11 5.2 Demand generation process steps

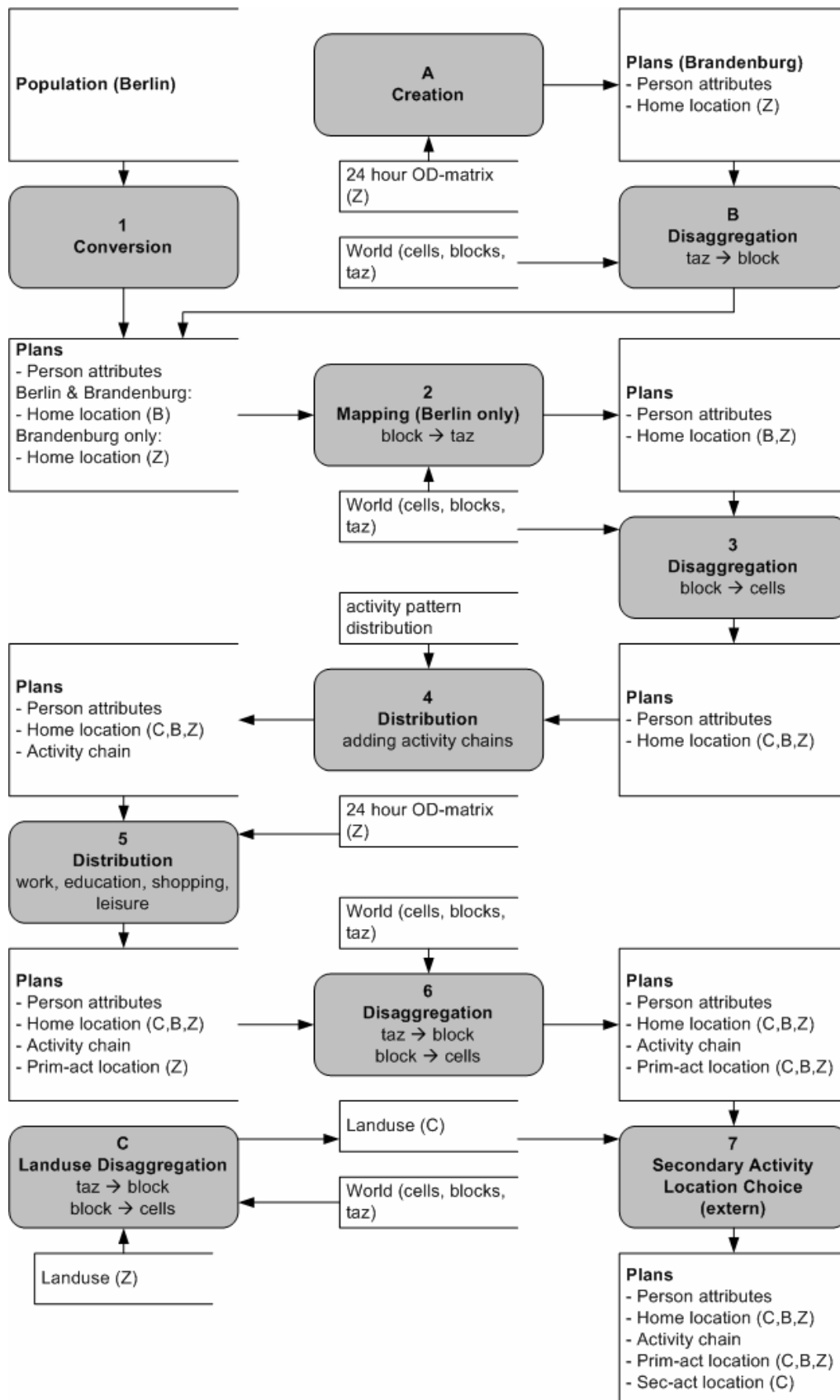
12 In comparison with the process of the Kanton Zurich scenario (Section 4) there is one step
13 more because of the fact, that three different levels of spatial resolution are used instead of
14 only two. Further more, three additional steps are included (Figure 3):
15

- 16 A *Creation*: As already mentioned above, there is no information about the population of
17 Brandenburg. Because of the fact that the number of trips of the origin-destination matrix
18 starting in zones of Brandenburg fits with its total population, we simply assigned for each
19 of those trips one individual to that start zone. This is—of course—a rough approximation
20 of the distribution of the population of Brandenburg.
21
- 22 B *Disaggregation*: Since the individuals are placed at traffic analysis zone level (Z) we need
23 to map them to block level (B) we assume a uniform distribution over all blocks of a given
24 zone.
25
- 26 C *Landuse disaggregation*: To use the secondary activity location choice module (see
27 Section 3.4) the landuse information has to be given at the raster level. Therefore, the
28 TAZ based landuse data also has to be disaggregated into raster level (R). It is done in a
29 first step by uniform distribution over all blocks (B) of each TAZ. As a second step
30 another uniform distribution is used to propagate the landuse data from block level (B) to
31 raster level (R).
32
- 33 1. *Conversion*: This is the same conversion as already described in Section 4. Together with
34 the two special steps A and B (see above) the whole population is now given in the
35 appropriate MATSIM XML plans format (MATSIM, 2005).
36
 - 37 2. *Mapping*: To add the corresponding TAZ level to the population of Berlin a mapping from
38 block level to TAZ level has to be done. This mapping is non-ambiguous.
39
 - 40 3. *Disaggregation*: The last remaining spatial level (raster level R) is mapped in this step.
41 The method is again an additive uniform distribution over each cell of the given block.
42
43
44
45
46
47
48
49
50
51
52

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
4. *Distribution*: as above (Section 4.2, item 3)
5. *Distribution*: In contrast to the same step of the Kanton Zurich process we assign all different primary activity types according to the given matrix. Because we do not have any information about the purpose of a given trip of the matrix we cannot distinguish it from others. The location of the primary activity is chosen according to the distribution of the destinations of those trips starting at the home location of the current individual.
6. *Disaggregation*: In this step we again need to disaggregate the location of the secondary activity to raster level (R). It is done in the same way as described for the landuse (see special process steps).
7. *Secondary activity location choice (extern)*: The secondary location choice module by (Marchal and Nagel, 2004) is again used to find the missing locations in the same way as described in process step 6 of the Kanton Zurich scenario (Section 4).

The final demand of that scenario again holds the same data points as described at the end of Section 4.2. Also for this scenario we will use MATSIM to generate the missing information.

Figure 3 Process plan for generating demand of Berlin and Brandenburg



6. MATSIM iterative optimization process

As already mentioned, we use the generated demand as input for the MATSIM iterative optimization process. In Raney and Nagel (2004) is shown that feasible timing of given schedules can be generated by simple initial departure times and durations of activities. Therefore, it is not necessary to build in appropriate algorithms into the demand generation framework but we can use MATSIM as a final external module for that.

In contrast to Raney and Nagel (2004) we are using an enhanced “time allocation” module (the so called “planomat”) based on Meister and Balmer (2005). Using the concept of genetic algorithms this module can be used for time scheduling as well as for location or mode choice. Since location choice is already done by the demand generation framework and the traffic flow simulation of MATSIM (Cetin and Nagel, 2003) handles only the individual transport mode, the functionalities of the “planomat” module is reduced to time choice only.

6.1 Time scheduling results

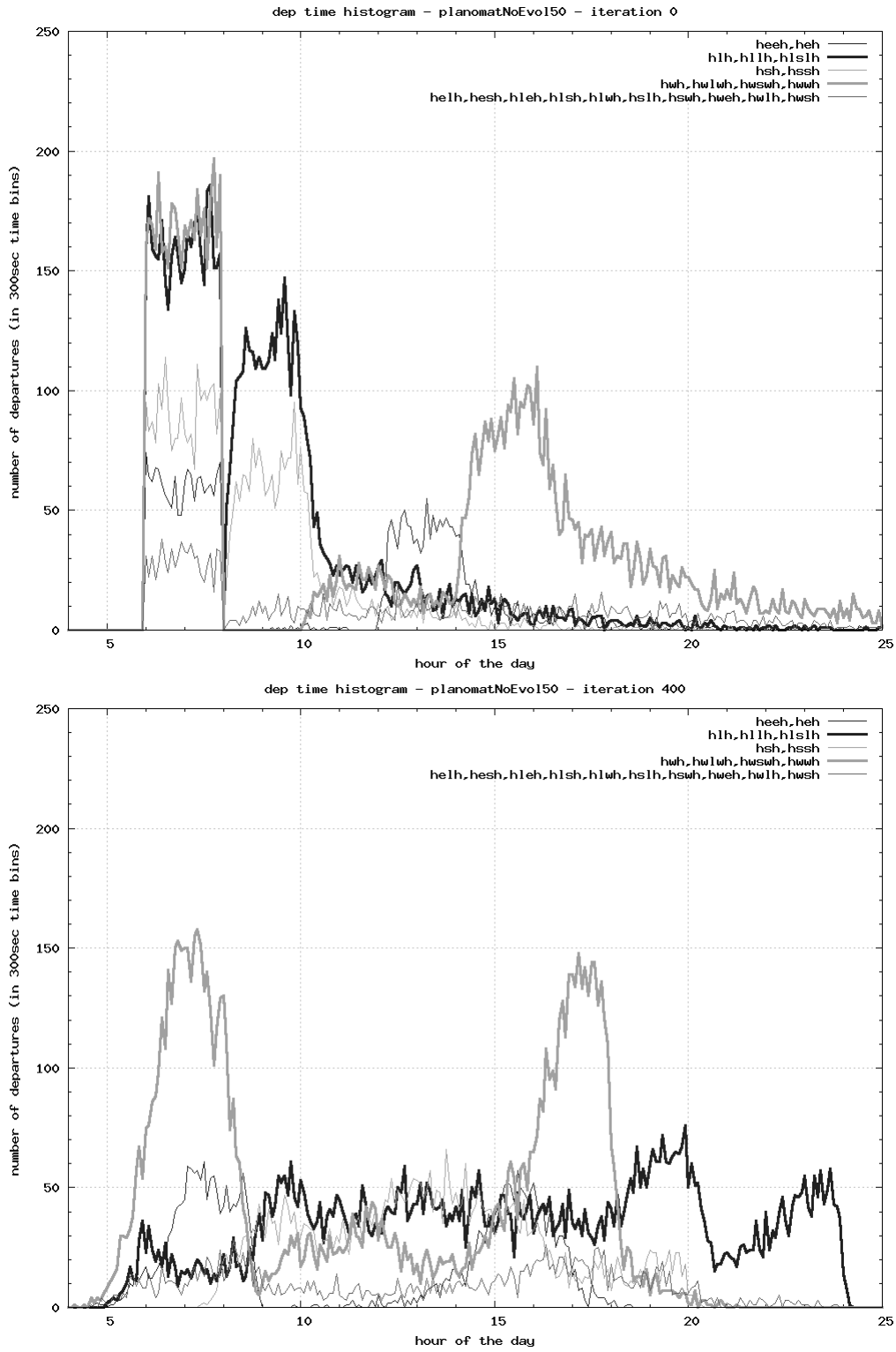
Using the same MATSIM setup as described in (Balmer *et al.*, 2005) we run a one percent sample of the population of the Kanton Zurich scenario. By an artificial reduction of the capacities of each link of the given network (Cetin and Nagel, 2003) by a factor of 50 we can produce the resulting traffic with similar effects (i.e. congested areas) as it would be produced for the real amount of persons.

Figure 4 shows the differences between the initial departure time schedule (iteration 0) and the departure time schedule after 400 iterations. The figure highlights (bold graphs) two groups of individuals from the remaining groups:

- Bold grey graph: Individuals with schedules containing “work” as the second and second to last activity in the chain
- Bold black graph: Individuals with schedules containing “leisure” as the second and second to last activity in the chain

Since the “work” activities should be performed between 7am and 6pm and the duration of performing “work” should be around 8 hours (Balmer *et al.*, 2005), they produce the expected morning and evening peak (rush hour) after 400 iterations.

Figure 4 Departure time histograms of a 1% sample population of Kanton Zurich scenario



The leisure activities do not have such a hard constraint (the setup defines it between 5am and midnight). The duration of performing one leisure activity is set around 2 hours. Therefore, there is much more flexibility of time scheduling for leisure activities. The Figure 4 reflects that fact. While the time choice at iteration 0 is quite worse, the improvement after 400 iterations is immense. Only a few persons choose the morning rush hour to travel while the major part performs their leisure activity in the evening (finishing their last leisure activity almost before midnight, when almost no traffic exists).

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52

7. Summary

This paper presented a demand generation framework for large scale micro simulations. By using the framework for two different scenarios described in Section 4 and Section 5 it is shown that

- it is flexible enough to handle the variety of input data (Zurich data differ from Berlin and Brandenburg data),
- it is flexible enough to extend or replace algorithms (different algorithms for Berlin and Brandenburg commuter data than for Zurich commuters),
- it provides (dis)aggregation to different spatial resolutions (two levels of resolution for Zurich scenario, three levels for Berlin and Brandenburg scenario),
- it provides a robust interface to third party models, programs and frameworks (the external modules: secondary activity location choice and MATSIM),
- it is suitable for unlimited number of individuals for a given scenario (about 7 million persons for the Berlin and Brandenburg scenario), and
- it is easy extensible to replace algorithms with more enhanced ones, to add new algorithms for existing packages and to add new packages to handle new input data.

8. Outlook

The “planomat” module mentioned in Section 6 provides an attractive amount of functionality. Since it is—besides the already used time choice—also a location choice and a mode choice module it is of high interest to actually use these functionalities. For that we need to provide this module with additional information.

As a first step, we need to provide an alternative to individual transport mode in the MATSIM project. Therefore, the demand generation framework will be extended such that it delivers information about travel times from one zone (block, district, TAZ or municipality) to another. With it, the “planomat” module decides which mode will be used for each person in the population. In simple words, we expect the modal split to be generated automatically.

As a second step, we will add another package to the framework which contains data about catchment areas and commuter sheds (Axhausen *et al.*, 2004). With it, we can replace primary activity location choice (commuter package) and also secondary activity location choice (external module) by adding a set of possible locations for all available activity types to each person of the population. This set will then be used by the “planomat” module to find iteratively—in the MATSIM framework—an appropriate location for the given activities of each individual.

9. References

1
2
3
4 Abraham, J. E., T. Weidner, J. Gliebe, C. Willison and J.D. Hunt (2005). Three methods for
5 synthesizing baseyear built form for use in integrated land use-transport models.
6 *Transportation Research Board Annual Meeting*, Paper 05-2325, Washington, D.C.

7
8 ARE, Bundesamt für Raumentwicklung und Bundesamt für Statistik (2001). Mobilität in der
9 Schweiz, Ergebnisse des Mikrozensus 2000 zum Verkehrsverhalten, Bern und Neuchâtel.

10
11
12 Arentze, T., F. Hofman, H. van Mourik, and H. Timmermans (2000). ALBATROSS: A multi-
13 agent rule-based model of activity pattern decisions. *Transportation Research Board Annual*
14 *Meeting*, Paper 00-0022, Washington, D.C.

15
16 Axhausen, K.W. (2005) Activity spaces, biographies, social networks and their welfare gains
17 and externalities: Some hypotheses and empirical results, Paper for the PROCESSUS
18 Colloquium, Toronto.

19
20
21 Axhausen, K.W., M. Botte and S. Schönfelder (2004). Systematic Measurement of Catchment
22 Areas. CTPP 2000 Status Report, U.S. Department of Transportation in cooperation with the
23 TRB Census Subcommittee. <http://www.fhwa.dot.gov/ctpp/sr0804.pdf> (Accessed July 2005).

24
25
26 Axhausen, K. W. and M. Frick (2005) Nutzungen, Strukturen, Verkehr, in G. Steierwald, H.-
27 D. Künne and W. Vogt (eds.) *Stadtverkehrsplanung: Grundlagen, Methoden, Ziele*, 61-79,
28 Springer, Heidelberg.

29
30
31 Balmer, M., M. Rieser, A. Vogel, K.W. Axhausen and K.Nagel (2005) Generating day plans
32 using hourly origin-destination matrices, in T. Bieger, C. Laesser and R. Maggi (eds.)
33 *Jahrbuch 2004/05 Schweizerische Verkehrswirtschaft*, 5-36, SVWG, St. Gallen.

34
35
36 Bhat, C. R., J. Y. Guo, S. Srinivasan, and A. Sivakumar (2005). A comprehensive
37 econometric microsimulator for daily activity-travel patterns (cemdap). *Transportation*
38 *Research Record*, forthcoming. http://www.ce.utexas.edu/prof/bhat/FULL_REPORTS.htm
39 (Accessed Feb. 2005).

40
41
42 Bowman, J. L., M. Bradley, Y. Shiftan, T. K. Lawton, and M. Ben-Akiva (1999).
43 Demonstration of an activity-based model for Portland. In *World Transport Research:*
44 *Selected Proceedings of the 8th World Conference on Transport Research 1998*, 3:171-184,
45 Elsevier, Oxford.

46
47
48 BFS, Swiss Federal Statistical Office (2005). <http://www.statistik.admin.ch/eindex.htm>
49 (Accessed Feb. 2005).

1 Cetin, N. and K. Nagel (2003). A large-scale agent-based traffic microsimulation based on
2 queue model. In *Proceedings of Swiss Transport Research Conference (STRC)*, Monte Verita,
3 Switzerland.

4
5 DYNAMIT (2005), <http://mit.edu/its> (Accessed July 2005).

6
7 DYNASMART (2005), <http://www.dynasmart.umd.edu> (Accessed July 2005).

8
9 Frick, M. A. (2004) Generating Synthetic Populations using IPF and Monte Carlo
10 Techniques: Some New Results. *Swiss Transport Research Conference (STRC)*, Monte
11 Verita, Switzerland. <http://www.strc.ch> (Accessed Feb. 2005).

12
13
14 Friedrich, M., I. Hofsäß, K. Nökel, and P. Vortisch (2000). A dynamic traffic assignment
15 method for planning and telematic applications. In *Proceedings of Seminar K, European*
16 *Transport Conference*, Cambridge, GB, 2000.

17
18
19 Hunt, J. D., R. Johnston, J.E. Abraham, C.J. Rodier, G.R. Garry, S.H. Putman, and T. de la
20 Barra (2001). Comparisons from sacramento model test bed. *Transportation Research*
21 *Record*, 1780:53–63.

22
23
24 Marchal, M. and K. Nagel (2004). Modeling location choice in activity-based models with
25 cooperative agents. *Swiss Transport Research Conference (STRC)*, Monte Verita,
26 Switzerland. <http://www.strc.ch> (Accessed Feb. 2005).

27
28
29 MATSIM, multi agent traffic simulation (2005). <http://www.matsim.org> (Accessed Feb.
30 2005).

31
32 Meister, K. and M. Balmer (2005) An improved replanning module for agent-based micro
33 simulations of travel behavior, Working Paper, Institute for Transport Planning and Systems
34 (IVT), ETH Zurich, Zurich, Jul. 2005

35
36
37 Nagel, K. and F. Marchal (2003). Computational methods for multi-agent simulations of
38 travel behavior. In *Proceedings of the meeting of the International Association for Travel*
39 *Behavior Research (IATBR)*, Lucerne, Switzerland.

40
41
42 Pendyala, R. M. (2004). Phased implementation of a multimodal activity-based travel demand
43 modeling system in florida, Project number 0510812 (BA496), [http://www.dot.state.fl.us/](http://www.dot.state.fl.us/research-center/Completed_PTO.htm)
44 [research-center/Completed_PTO.htm](http://www.dot.state.fl.us/research-center/Completed_PTO.htm) (Accessed Jan. 2004).

45
46
47 Raney, B. and K. Nagel. (2004) An improved framework for large-scale multi-agent
48 simulations of travel behaviour. *Swiss Transport Research Conference (STRC)*, Monte Verita,
49 Switzerland 2004. <http://www.strc.ch> (Accessed Feb. 2005).

1 RBS, Raumbezugssystem Berlin (2000) Bevoelkerungsverteilung, Teil-Verkehrszellen und
2 Block Geometrie / Hierarchie, Statistisches Landesamt Berlin, June.2000.

3
4 Salvini, P. A. and E.J. Miller (2003). ILUTE: An operational prototype of a comprehensive
5 microsimulation model of urban systems. *In Proceedings of the meeting of the International*
6 *Association for Travel Behavior Research (IATBR)*, Lucerne, Switzerland.

7
8 SAX, Simple API for XML (2005). <http://www.saxproject.org> (Accessed July 2005).

9
10 SfsB, Senatsverwaltung für Stadtentwicklung Berlin (1998), Datengrundlagen
11 Stadtentwicklungsplan Verkehr 1998.

12
13 TRANSIMS, TRansportation ANalysis and SIMulation System (2005). Los Alamos National
14 Laboratory, Los Alamos, NM. <http://transims.tsasa.lanl.gov> (Accessed July 2005).

15
16 VISEM, Planung Transport und Verkehr, PTV (2005). [http://www.ptv.de/cgi-
17 bin/traffic/traf_visem.pl](http://www.ptv.de/cgi-bin/traffic/traf_visem.pl) (accessed July 2005).

18
19 Vovsha, P.; E. Petersen and R. Donnelly (2002). Microsimulation in Travel Demand
20 Modeling: Lessons Learned from the New York Best Practice Model. In *Transportation*
21 *Research Record: Journal of the Transportation Research Board, No. 1805*, pp. 68-77, TRB,
22 National Research Council, Washington, D.C.

23
24 Vrtic, M. and K.W. Axhausen (2003) Experiment mit einem dynamischen
25 Umlegungsverfahren, *Strassenverkehrstechnik*, **47**(3), pp 121-126.

26
27 Waddell, P., A. Borning, M. Noth, N. Freier, M. Becke, and G. Ulfarsson (2003).
28 Microsimulation of urban development and location choices: Design and implementation of
29 UrbanSim. *Networks and Spatial Economics*, **3**(1): 43-67.

30
31 Waddell, P., H. S,Úevčíková, D. Socha, E. Miller, and K. Nagel (2005). OPUS: an
32 international collaboration to develop an open platform for urban simulation. *Presented at the*
33 *9th Conference on Computers in Urban planning and urban management (CUPUM)*,
34 University College London, UK.