Integrated activity-based demand modeling and traffic assignment on micro-level for very large scenarios

Michael Balmer
Marcel Rieser
Kay W. Axhausen

ITM 2010, Tempe, AZ

May 2010
May 2010

Abstract

Activity-based demand models as well as traffic assignment models have both gained in complexity over the last years. This leads to problems if the complex models should be applied to large, detailed scenarios. A common response to this dilemma are two types of systems, one calculating traffic flow characteristics for large regions on a simple model, the other model much more details although for smaller regions. For all these years, the interface between demand and assignment remained (time dependent) origin-destination matrices. This robust and easy to handle interface has become more and more a burden, as many details of the demand get lost. Especially, it is nearly impossible to feed back information from the assignment to the demand in order to achieve (individual) demand responsiveness.

This paper presents the newest development of the open source software MATSim (Multi-Agent Transport Simulation, [http://matsim.org](http://matsim.org), MATSim-T, 2008), developed jointly at TU Berlin and ETH Zurich. MATSim includes both a demand optimization part and an agent-based traffic simulator which are based on individual day plans instead of matrices, and thus enables researchers to feed back detailed information from the simulation to the demand and back again. In recent years, MATSim was trimmed to support the simulation of very large scenarios, consisting of several millions of agents, on very detailed navigation networks with up to one million (directed) links. At the same time, the complexity of the demand modeling part was enhanced, e.g. with the inclusion of mode choice next to departure time and route choice, which were already available for quite some time. This paper describes the technical aspects of MATSim that make the simulation of such huge scenarios possible and presents a few results from such a simulation which includes mode-choice in the demand optimization part.

Keywords

transport planning, demand modeling, dynamic traffic assignment, multi-agent micro-simulation, MATSim, Switzerland
1 Introduction

Activity-based demand modeling nowadays is a well known, widely accepted approach for transport planners. It was first suggested by Jones, Dix, Clarke and Heggie (Jones et al., 1983) more than 35 years ago. Since then an enormous development of that research area delivered a large variety of different, ready-to-use applications to model activity-based demand (implementations and discussions can be found in Hensher, 2001; Gärling and Axhausen, 2004; Bowman et al., 1999; Vovsha et al., 2002; Jonnalagadda et al., 2001; Bhat et al., 2004; Axhausen, 1990; Lohse et al., 1997; Kutter, 1983). During the same period traffic assignment models evolved almost at the same speed, from simple static assignment to very detailed and realistic dynamic micro-simulation models. All these applications have in common that the more complex the models are the less suitable they are for large detailed scenarios. The common response to this dilemma are two types of systems, one calculating static or dynamic traffic flow characteristics for large regions, the other modeling detailed physical dynamics of movements of different transport modes often used to design small areas like intersections or roundabouts. During all these years the interface between demand and assignment remained the same: (time dependent) origin-destination matrices. This very robust and easy to handle interface has become more and more a burden when researchers try to feed the outcome of the assignment step back to the demand modeling part. This is of crucial importance, especially when it comes to changes of the (individual) demand responsive to changes in the infrastructure which is the usual case in ITS applications, road pricing strategies, adaptive route guidance systems, etc. Another disadvantage of OD-matrices is the quadratic scalability which does not allow one to increase spatial detail on the demand side. Also this leads to enormous problems when modeling multi-modal trips, parking search, walk or bike mode.

To summarize, transport planning needs an integrated, dynamic, microscopic, activity-based demand and assignment model where individuals’ activity chains in space and time stay consistent during the whole feedback process and last but not least-can be computed for very large scale scenarios in reasonable time.

Agent based micro-simulation technique is a promising solution for that need. This paper presents newest development of the open source software MATSim (Multi-Agent Transport Simulation, http://matsim.org [MATSim-T] 2008) developed jointly at TU Berlin and ETH Zurich. In the recent years the MATSim project matured from an experimental playground for researchers to a robust and extendable software framework which is applied to real scenarios with promising results. The basic evolutionary (relaxation) strategy has been retained throughout: an initial travel demand is executed with the mobility simulation, which returns improved and more consistent generalized cost estimates. These are used to update the previous schedules. This is repeated until there are no chances for further unilateral improvement anymore.
This agent-based concept has an important advantage: It allows one to explicitly model individual reactions to changes in the infrastructure without losing any kind of chaining information and it gives the transport planners the chance to trace tradeoffs between utility gains and losses within the complete individual schedule, called plan in MATSim. The content of such a plan are the activities that the agent performs during the day, their precise location at geo-coordinate level of detail, the activity start and end time, the activity type, also the activity chain itself, the trips connecting two sequential activities including departure time, travel time, arrival time, transport mode and—most important—the route. This stays in contrast to the classical approaches, where route choice is part of the assignment process. In MATSim, the route is an explicit part of the demand modeling process and has to be adapted (relaxed) during the feedback loop. Therefore, the mobility simulation just executes—based on a queue model—the given complete demand for all agents in the system. Only in the mobility simulation physical interactions (i.e. congestion) between the agents happen. This has the advantage that the adaptation (“learning”) of the agents’ plans can be handled completely independent from all the other agents in the system and—even more important—adaptation can be completely personalized in any aspect of the agents’ plan.

While in the past 10 years, the focus was mainly set to private motorized transport and while the relaxation of the individual demand is set to travel routes, departure times and activity durations only (examples are Beuck et al., 2007; Raney, 2005; Charypar et al., 2006) this paper shows the extension of the search space of the relaxation process to mode choice on sub-tour basis and location choice for secondary activities (shopping and leisure) at building level of detail.

At least one other choice dimension for a transport planning application is still missing, the activity chaining. While MATSim includes (experimental) functionality to optimize the number and sequence of the activities as presented by Feil et al. (2010), this paper will focus on route, time, mode and secondary location choice only. Home, work and the location of education is kept fixed during the relaxation process. MATSim’s demand relaxation is a utility maximization process. It is shown how a simple but fairly robust utility function introduced by Charypar and Nagel (2005) based on the Vickrey model (Vickrey, 1969) has been extended to capture the new choice dimensions.

Last but not least, MATSim is a modular approach. That means that

1. the micro-simulation can be exchanged by another one (in MATSim are already two different implementations available),

2. the dimensions of the search space can be defined (e.g. Gao et al., 2010, presented that route choice only used to directly compare MATSim assignment with the assignment process of EMME/2),

3. the model of adaptation for each choice dimension can be chosen (e.g. time adaption can
be done via a random mutation process or via a so called “best-response module” like the “planomat” module introduced by [Meister et al.] (2006), and

4. the utility function can be replaced.

The paper will present the overall approach, but will focus on the—to our knowledge—largest scenario addressed so far by any agent-based micro-simulation model in transport. The system is applied to modeling a 24 hour, completely time dynamic average work day for whole of Switzerland with many millions of individuals performing over twenty million trips. The traffic simulation runs on a high resolution navigation network with ca. one million directed links while the activity locations are chosen among ca. 1.7 million facilities (homes, commercial and industrial buildings). Compared to typical planning scenarios the size of the network is about 15 times bigger than average planning networks, and the planning demand is about 300 times more detailed (each link in MATSim is a “zone”, so the model contains of about one million zones), and the land use data has the resolution of single buildings. We will report on the computational performance on a shared memory machine using 90GB RAM and 16 cores and the quality of the model.

2 Scenario Description

The goal is simulate the whole population of Switzerland on a high-resolution street navigation network. The network used in this study is based on the proprietary TeleAtlas MultiNet Specification 4.3.2.1 (Tele Atlas MultiNet, 2010), although also other data sources like the open-access alternative OpenStreetMap (OpenStreetMap, 2010) could be used. The network contains the road traffic infrastructure of all of Switzerland over the period from autumn 2008 to spring 2009. While some attributes (like topology and maximum speeds) are very precise, information on the number of lanes per driving direction and flow capacities had to be imputed based on the TeleAtlas road classification. As the number of road segments strongly influences the run time of MATSim, the detailed geometry of roads was not maintained, resulting in a network description with 472’819 junctions (called nodes in MATSim) and 1’035’305 road segments (directed links). As trips can start or end on nearly any link, this is comparable to a system with about one million available zones.

While no detailed simulation of public transport was available at the time of this study in MATSim, there was still some data required for the mode choice model to work. In this study, public transport travel times were estimated based on a travel time matrix between the centers of the 3114 Swiss municipalities (Vrtic et al., 2005) and a list of about 15’000 operated public transport stops.
Besides the traffic network, the modeled region is mainly characterized by the land use, which determines what types of activities can be performed in which places. In this study, the five major activity types home, work, education, shopping and leisure were used. Land use information for home was derived from the Swiss National Census (Vrtic and Axhausen 2003), which had a spatial resolution of 100x100 meters. Locations for work were generated from the Swiss National Enterprize Census (Swiss Federal Statistical Office 2001), which contains the exact number of full-time equivalents in industry and services, again in a resolution of 100x100 meters. In addition, the data set contained information which allowed to create locations for education, shop and leisure. The locations were enriched with opening times based on publicly available data for shops and simple assumptions about working hours. For all of Switzerland, 382,979 work locations were generated. Each activity location is finally mapped to a single link from the network.

Given the land use information, the population for this project was generated using data from an industry partner and the Swiss micro census on travel behavior (Swiss Federal Statistical Office 2006). The demand modeling comprises data fusion between the two data sources as well as the generated land use information. The resulting population consists of 5,986,051 syntectic persons or agents (about 88% of the Swiss population), each with one plan describing their activities and trips during one day, with an average of 3.7 trips per agent. More details about the initial demand modeling can be found in Meister et al. (2010).

3 Simulation Framework

The key concept of MATSim is an iterative loop between the traffic flow simulation and the demand optimization. The latter is often called the replanning phase of MATSim. In the following, a very brief description of the simulation framework will be given.

The traffic flow simulation executes each agent’s day plan simultaneously with the plans of all other agents. It is therefor also called the plan execution phase. A more in-depth description of the traffic flow simulation is be given in Section 4.4. During the replanning, agents can make copies of their plans and modify the copies, which will be afterwards executed again by the traffic flow simulation. By comparing the modified copy with the original plan using a generalized cost function that evaluates the performance of a plan during the simulation, agents can distinguish between worse and better plans. By duplicating plans and modifying them, agents collect a number of different plans, some of which perform better than others. Due to memory constraints, agents will have to delete some plans after a certain number of replanning. By deleting plans with a bad performance and keeping plans with a good performance, the agents optimize their plans according to the principles of an evolutionary algorithm.
In this project, the modifications during the replanning include route choice, departure time and activity duration choice, secondary location choice, and mode choice on a sub-tour level.

4 Technical Aspects

To compute the given scenario in a useful amount of time, several advances to the algorithms used had to be implemented. In addition, the large amount of memory required optimizations as well in order to keep the memory requirements within a feasible range. This section describes the used algorithms and applied optimizations that made it possible to simulate this very large scenario.

4.1 Router

Calculating the shortest route in a high-resolution network can be quite time consuming, even more if it has to be done for several hundred thousands of agents. Dijkstra’s algorithm for finding the shortest path in a network \cite{Dijkstra1959} is a commonly used algorithm for the routing problem. MATSim uses an enhanced version of Dijkstra’s algorithm, which supports time-dependent travel times on the links, and also performs much better than the original algorithm \cite{LefebvreBalmer2007}. The performance improvement is achieved by using an A*-algorithm that uses Landmarks to further improve the applied heuristics.

The actual cost of a link, as it is used by the shortest path algorithm, is defined as a combination of the link’s length and an estimation of the time-dependent travel times thereon. Thus, the algorithm actually becomes a least-cost path algorithm, being consistent with the generalized cost function used to score plans. The estimated travel times used to calculate the cost of a link are based on events from the previous run of the traffic flow simulation. As the estimated travel times are the same for all agents, it means that all the agents have the same global knowledge of the system.

For the public transit mode \((pt)\), no network is available that could be used for routing. Instead, the travel time is estimated based on the location of stops and a given travel time matrix. In a first step, the nearest stop is searched for the location of departure and arrival of the agent. In order to make this look-up very efficient for an expected number of \(10^4\) stops, the stops are sorted in a Quad-Tree data structure \cite{FinkelBentley1974} that allows for quick spatial queries. Averaged public transit travel times are only available between zones, but not between stops directly. Thus, each stop is matched to a zone. In a second step, the travel time between the zones containing the departure and arrival stop is looked-up. To this travel time, access and egress times are added based on an assumption of walking speed and the direct distance from
activity location to the stop location. While this calculation is not depending on the time of day, it cannot be pre-calculated in the case where agents can change their activity locations (see Section 4.3).

4.2 Combined Time and Mode Choice

planomat is a replanning module for activity durations, departure times and mode choice on sub-tour level. This module replaces the random time allocation mutation module used in earlier studies with a genetic algorithm, leading to an enormously reduced number of iterations needed for convergence of the MATSim learning framework (usually several dozens of iterations versus several hundreds of iterations, see Meister et al., 2006). This reduction in the number of iterations more than compensates the increased computing time of the module due to its complexity.

In addition, the choice of a genetic algorithm makes it possible to include additional plan variables into the optimization process. In this study, mode choice on a sub-tour level was also replanned by planomat, with the given choice set of car, public transit – pt, bike, and walk for each sub-tour.

4.3 Secondary Location Choice

Some activity locations do usually not change, like the locations for the activities home, work, and education. As these locations are based on empirical data, the agents cannot change them. For other activity types although, like shopping and leisure, the locations can usually be freely chosen among some alternatives. The secondary location choice module first estimates a travel time budget, based on the fixed locations as boundary conditions (Horni et al., 2009). This estimation helps to reduce the search space, in which afterwards a randomly new activity location is determined given the available locations. The routes to and from the newly placed activity is then re-calculated using the same routing algorithm as described above. As a random replanning module, it executes very fast.

4.4 Traffic Flow Simulation

The traffic flow simulation is responsible for executing the agents’ plans simultaneously, while respecting physical boundary conditions of the infrastructure (e.g. maximal flow capacities, physical space on a link, etc). In previous studies, different implementations of a time-step based queue model was used to simulate the traffic flow (Cetin, 2005; Raney and Nagel, 2005;
This simulation model advanced the simulation time in fixed increments, checking each link and node in each time step for possible movements. In order to speed up the traffic flow simulation, the concepts of the event-driven queue model of Charypar et al. (2007) were re-implemented in Java. Instead of advancing the time in pre-defined time-steps, the event-driven model maintains a chronologically ordered list of events. Such events are for example when an agent wants to depart, or when it will be able to leave a link based on the free flow travel time. The simulation can thus always advance as far as possible in time, skipping time steps in which no vehicle is about to leave a link.

The traffic flow simulation only supports the movement of private cars on the network. Transport modes other than car are not simulated in detail. Instead, agents using another mode are teleported: They are removed at their current location at the time of departure, and placed at their target location after the travel time specified in their leg-description. This teleportation feature makes it possible to support additional transport modes beside car, even though without additional benefits (e.g. no influence of congestion, no quality-measurements in over-crowded public transit vehicles).

### 4.5 Further Optimizations

Besides the above-mentioned algorithmic changes, a number of other optimizations were applied to further speed-up the execution of the simulation. One important aspect is the utilization of multiple CPU cores available in modern computers for the parallel execution of some tasks. During the replanning phase, each agent is processed independently from the others. This makes it easy to process several agents at the same time, using multiple threads. The duration of the replanning phase can thus be scaled nearly linearly to the number of available CPU cores.

The traffic flow simulation cannot that easily be sped-up, as each agent may be dependent on other agents that are in front of it on the road, and the simulation time must increase sequentially for all agents. But the simulation can still make run faster by moving all non-critical code (that is, code whose results are not needed immediately during the simulation) to another thread, which works on those items independently from the simulation. In previous versions of the traffic flow simulation, the generated events were processed immediately when they were created, interrupting the flow of the simulation during that time. Instead of processing the events immediately, the newly generated events are now just stored in a queue. Thus, the simulation can continue with its calculation without interruption. An additional thread accesses the queue and processes the available events independently from the traffic flow simulation. This helps to bring the duration of the execution phase nearer to the runtime of the actual traffic flow simulation, reducing the overhead by processing it in parallel.

Not only speed is a concern when running large scenarios, but also memory consumption. Stor-
Routes proved to occupy the biggest amount of memory, as each route was stored as a simple list of links. With navigation networks, it is not uncommon that an average route consists of more than one hundred links. Storing that amount of information for a few million agents, each having multiple plans with multiple routes in it, easily takes up several tens of gigabytes of memory, by far the largest part of the memory used. By analyzing the network and calculating a designated follow-up link for each link of the network, the amount of data needed to store routes can be massively reduced. The follow-up link of a link should be the one that is most likely to be taken by agents traveling on a link. Once all follow-up links are calculated, the routes only need to store the first link of a route and then each subsequent link of the route that is not the follow-up link of the previous link in the route. Compressing the data that needs to be stored for routes in this way yielded in massive memory improvements. It must be noted though that a small overhead is added to the runtime of the simulation by this memory optimization, as the route data needs to be specially processed on each access.

5 Performance

The integrated demand modeling and agent-based and dynamic traffic assignment process produced by the iterative relaxation process is performed on a shared-memory machine of type Sun Fire X4600 M2 with 16 cores (8 Dual-Core CPUs) and 128 GB RAM. The process re-
Figure 2: Relaxation

![Graph showing relaxation process](image)

requires 90 GB of RAM on java 1.5.17 virtual machine for the given scenario (5’986’051 agents, 22’152’812 trips, 28’138’863 performed activities, 1’618’267 different facilities, 11 different activity types, 472’819 nodes and 1’035’305 directed links).

The average computation time for one iteration is about 4.5 hours (see Figure 1) whereas the major part is used by the micro-simulation. At each tenth iteration intermediate results are dumped to the hard drive. The relaxation is reach after 60 iteration shown in Figure 2. In contrast to previous optimization processes presented by [Balmer (2007); Balmer et al. (2009) and Balmer et al. (2004) best-respond replanning modules for time, mode and route optimization reduces the number of iteration needed. More insides of the mechanisms of MATSim’s relaxation can be found in [Meister et al. (2010) and Balmer et al. (2010).

6 Results

The scenario is calibrated by an experimental design varying the set of the fitness function and scaling parameters. To speed up the simulation and allow for more variations being run at the same time, the calibration is done using a 25% random sample of the complete demand. This allowed us to run 3 simulations at the same time on the machine capable to run the full 100% simulation, while reducing the runtime to about four days (compared to two weeks for the full demand) and still producing meaningful results. The quality of the simulation outcome is validated by comparing the results with the values of the Swiss micro census for travel (Swiss Federal Statistical Office 2006) and to real-world traffic counts.

As it is the first time that mode choice on sub-tour level is included in the simulation setup, the
Table 1: Modal split (trip based)

<table>
<thead>
<tr>
<th>mode</th>
<th>micro census</th>
<th>MATSim</th>
</tr>
</thead>
<tbody>
<tr>
<td>car</td>
<td>43.3 %</td>
<td>44.9 %</td>
</tr>
<tr>
<td>share a ride</td>
<td>4.4 %</td>
<td>0.9 %</td>
</tr>
<tr>
<td>public transport</td>
<td>13.4 %</td>
<td>15.3 %</td>
</tr>
<tr>
<td>bike</td>
<td>7.6 %</td>
<td>6.6 %</td>
</tr>
<tr>
<td>walk</td>
<td>31.3 %</td>
<td>32.2 %</td>
</tr>
</tbody>
</table>

Figure 3: Cumulative mode share by trip travel time

(a) Swiss micro census 2005
(b) MATSim

validation of the modal split received quite a lot of attention. Table 1 shows a comparison of the values from the official micro census 2005 and the values calculated by MATSim. As it can be seen, the values match quite well. The only bigger divergence is in the mode ride, where the value calculated by MATSim is very low. Because of a not yet validated micro model of the transport mode ride and the fact that the real share is below 5 percent it was decided to ignore this mode in the relaxation process.

As a representative for the various comparisons of the outcome with real world measurements (see Meister et al. 2010; Balmer et al. 2010) Figure 3 is shown presenting the cumulative mode share distribution by trip travel time. The values for 00:00 and values for 02:00 or larger can be ignored, as they are either artifacts of the model (in the case of 00:00) or are statistically not relevant, contributing less than 1% of all observations. For the other values, the comparison shows a satisfactory match between the results and the values observed in the micro census. Walking trips are underestimated in the simulation in the shorter distances, which might be a result of the used travel time estimation. The mode ride is again under-represented due to it not being part of the mode choice set (see above).
7 Conclusions

While in the recent years of developing models and presenting various work with the open source MATSim framework this paper proofs the ability to extend the behavioral model, the search space (replanning) and the physical simulation of MATSim to fulfill the various needs in transport planning. The concept of the integrated demand modeling and dynamic assignment process by separating the physical world from the individual planning world allows one to include more and more demand modeling parts into the system without breaking up or revamping already existing demand models. Even more, the system allows one to extend and/or replace a given implementation of the physical world by another.

The fact that a very large and fine grained scenario based on very detailed navigation network information and single facilities can be used to model completely dynamic and activity based demand for each individual of a whole country also proofs the scalability of such a system where each element only contributes linearly to the computation time. A typical planning scenario with a resolution that is about 15 times smaller than the one presented here and therefore, can be calculated within one day.

Last but not least, the comparison with real world validation data also shows that the model presented in here produces results in the same quality as macro models, while the outcome is still microscopic, individual and completely time dynamic.

References


