Truly agent-oriented coupling of an activity-based demand generation with a multi-agent traffic simulation

Rieser, Marcel; Nagel, Kai; Beuck, Ulrike (1)
Balmer, Michael (2)
Rümenapp, Jens (3)

(1)
VSP, TU-Berlin
Salzufer 17-19
D-10587 Berlin

(2)
IVT, ETH Zürich
ETH Hönggerberg
CH-8093 Zürich

(3)
Gertz Gutsche Rümenapp
Stadtentwicklung und Mobilität
Büro Berlin
Postfach 120361
10593 Berlin

Abstract

The typical method to couple activity-based demand generation (ABDG) and dynamic traffic assignment (DTA) are time-dependent origin destination matrices. With that coupling method, the individual traveler's information gets lost. Delays at one trip do not affect later trips.

It is, however, possible to retain the full agent information from the ABDG by writing out all agents' "plans", instead of the OD matrix. A plan is a sequence of activities, connected by trips. Since that information is typically already available inside the ABDG, this is fairly easy to achieve.

MATSIM takes such plans as input. It iterates between the traffic flow simulation (sometimes called network loading) and the behavioral modules. The currently implemented behavioral modules are route finding, and time adjustment. Activity re-sequencing or activity dropping are conceptually clear but not yet implemented. Such a system will react to a time-dependent toll by possibly re-arranging the complete day; in consequence, it goes far beyond DTA (which just does route adaptation).

Our paper will report the status of our current Berlin implementation. The initial plans are taken from an ABDG, originally developed by Kutter; to our knowledge, this is the first time that traveler-based information (and not just OD matrices) is taken from an ABDG and used in a multi-agent simulation. The simulation results are compared against real world traffic counts from about 100 measurement stations.

Introduction

In today's transportation planning, many tools for dynamic traffic assignment (DTA) take origin-destination (OD) matrices as input. These OD matrices are usually generated by some sort of activity-based demand generation (ABDG). Many implementations for ABDG (Arentze and Timmermans, 2000; Arentze and Timmermans, 2005; Bhat et al., forthcoming; Bowman et al., 1999; Miller and Roorda, 2003; Pendyala and Kitamura, In press) use individual travelers for internal calculation. However, individuals' information is lost when only OD matrices are written out at the end of the demand generation process. This may not lead to problems when traffic assignments are calculated for a whole day. However for cases where the temporal relationship between two journeys is of importance, OD matrices are not suitable.

On the contrary, multi-agent traffic simulations need information about individual travelers. Each modeled agent is assigned at least one "plan", a series of activities and connecting trips. Thus, a multi-agent traffic simulation requires more detailed data than OD matrices can offer. This renders many current tools for ABDG useless. Therefore new tools are being developed especially for the demand generation of multi-agent traffic simulations.

Multi-agent traffic simulations are getting more and more important in today's transportation planning, as they are more flexible when researching individuals' reactions in the case of tolls or other non-physical influences on the road network. But instead of creating new tools for the demand generation process, it is also interesting to see if existing ABDG tools could be adapted for multi-agent traffic simulation. Generating agents' plans from OD matrices has already been tried in the past, but is a complex and cumbersome process (Balmer et al., 2005). However, many ABDG already make use of "individuals" or "agents" internally to calculate the OD matrices. So if these internal states could be retained, they could make a better input for multi-agent simulations than the traditional output, OD matrices, makes.

In our paper we describe the steps taken to transform the individuals' data from an existing ABDG into agents' plans. The ABDG is based on the Kutter-Model (also known as Berliner Personenverkehrs-Modell; Kutter, 1984; Kutter and Mikota, 1990; Kutter et al, 2002) for the region of Berlin, Germany. The plans are then used as input for our multi-agent traffic simulation MATSim (Raney and Nagel, 2005). We will then compare the results of the simulation to real world data.

Scenario Setup

The chosen study area of Berlin and its surrounding—namely the federal state of Brandenburg—covers an area of 150 x 250 km and has a population of about 6 million inhabitants. We focus on the urban area of Berlin, therefore this part of the region is represented with a much higher level of detail and accuracy than Brandenburg regarding network and demand.

The road network was originally developed by the planning department of the city of Berlin (Senatsverwaltung für Stadtentwicklung). It has been used as part of the city's forecast model representing supply side of the year 2015. Manual changes were necessary in order to exclude modifications of the road network planed until 2015.

Nodes described by their coordinates and links connecting these nodes form the network. The network links are described by its major attributes like length, free flow speed, number of lanes, and capacity. In general, these network attributes are sufficient for our queue simulation, but their values did not always meet our requirements.

The network has been used for mid term or long term transportation planning in Berlin with a scope of 24 hours. The demand is described as 24 h OD-matrices, which are based on defined traffic analysis zones (TAZ). Such demand is assigned to the network using static assignment according to defined capacity speed functions.

As stated above, the official transport model has a scope of 24 hours, no further level of detail like time-dependent matrices is used. Thus, further network modifications were necessary in order to use the network in our multi-agent simulation.

The most crucial attribute of a network link is its capacity, which is interpreted very differently by the aggregated model used by the planning department of Berlin and the multi-agent simulation used by us. In our simulation capacity is understood as maximum outflow of a link in a given time period, while the model of the planning department does not treat capacity values as hard constraints The traffic assignment method applied uses suitable functions to relate capacity and flow with the resulting cost in terms of travel times. Thus, we had to adapt the theoretical capacity values that were the basis for a 24 hours static assignment. A factor was derived according to the fact that the daily traffic basically occurs in 12 hours of a day. In a second step, we converted the resulting theoretical 1-hour values used in static assignment into maximum values of outflow of a link in 1 hour. The maximum outflow of a link used in our simulation is double the 1-hour values used in traffic assignment. Additionally, the storage of a link is constrained. The storage of a link can be calculated as length times the number of lanes divided by the space a vehicle occupies in a jam (7.5 m). Unfortunately the number of lanes attribute is set to be one on all link of the original network. Since the number of lanes is not necessary for static assignment such a network error is tolerated in the aggregated model. In our simulation we currently set the number of lanes to two in order to calculate maximum storage, but a better solution has to be found in the near future.

The final road network representation consists of more than 10,000 nodes and almost 30,000 links connecting those nodes. As stated above, the demand data that is available for Berlin does only consist of 24h OD-matrices. This kind of demand data is only sufficient for strategic long term transportation planning. Multi-agent based simulations require the demand to be given on the individual level as day-plans containing the activities with the scheduling and location information and the trips between them.

Day-plans could be derived from the ABDG by the Kutter-Model. The Kutter-Model is a disaggregated, actvitiy- and behavior-oriented traffic demand generation model. It simulates the traffic demand of person groups with homogenous behavior with the help of expectancy values (Kutter, 1984; Kutter and Mikota, 1990; Kutter et al, 2002). This model is currently used to calculate 24 h OD-matrices for strategic planning and was modified to output activity chains (Rümenapp and Steinmeyer, 2006). The remaining of the paper describes how the ABDG data is used in our multi-agent simulation.

In order to speed up the Berlin scenario, the demand and the network capacities where scaled down to 10 % of the actual values.

Creating Plans from Activity Chains

Each activity chain contains information about the start location (reference to a zone), up to four activities, and the frequency of occurrence of the activity chain. The four activities are each described by their type, their location, and the transportation mode used to reach that location. Additionally, the activity chains are grouped, each group corresponding to one of 72 person groups with similar demographic attributes. The activity chains always describe trips that end at the same place as where they start (round-trips), while all activities between the first and the last activity take place somewhere else. The sum of all frequencies corresponds to the total number of round-trips accomplished by the people in the study area. There is no special provision for people who perform more than one round-trip per day; this leads to problems as we will see later.

As one can see, much information needed for agents' plans like activities and locations is available, with only the time information missing. In traditional DTA, the demand for 24 hours is distributed over the day using typical hourly volumes. This way, a temporally differentiated assignment can be calculated. In our simulation, time information is generated and optimized over several iterations by a special module (Meister et al., 2006). Initially, all activities are assigned a random activity duration within a range, where the range depends on the type of the activity. These random durations are replaced by more convenient durations during the iterations by the mentioned module.

Based on the given description it seems possible to use all the information from the activity chains and transform them into agents' plans. This would work flawlessly if the frequency of occurrence of each

activity chain were an integer value. But the methods used in traditional DTA allow any value as frequency, including floating-point numbers. This may work in assignments where frequencies are summed up on each link to get the total volume. However, in agent-simulations the smallest unit is an agent that cannot be split into two or more parts. Thus, a way had to be found to deal with the fractional frequencies of activity chains.

The input data from ABDG contains more than 7 million round-trips per day in our study area. More than 250 million different activity chains are used to describe the round-trips. There are many times more activity chains than there are round-trips, resulting in an average frequency per activity chain of far less than 1. This leads to problems when generating agents' plans from the activity chains: not every activity chain can be converted into a plan, but the frequencies of the activity chains must be considered to decide whether to use the data or not.

A simple approach would be to just round the frequencies to the next integer value and then decide if (and if so, how many) agents should be created based on that activity chain. Because of the special distribution of the frequencies (values range from less than 0.0001 up to the lower tens with an average value below 0.1), creating agents based on the rounded frequencies results in too few agents. Multiplying the frequencies with a fixed factor does not help much either: In that case, activity chains with a high frequency of occurrence would be chosen even more often, while many other activity chains still would not make it over the threshold of 0.5. This results in an overvaluation of already more often occurring activity chains, while low-frequency activity chains in general would rarely or never be used for an agent's plan.

It was then decided to sum up one frequency after the other. Every time the sum reaches 1.0 or any higher value, an agent with a plan based on the current activity chain is generated and the sum is reduced by 1.0. This way, no rounding occurs which could favor higher frequencies over lower ones. If the activity chains are in a random sequence (which we assume), then our method corresponds to a weighted random draw without replacement. Finally, over 7 million agents with each one plan assigned were generated, representing the 7 million round-trips undertaken by the people in the study area. As we will see later, it is quite important to note that the number of agents created does not represent the number of inhabitants in the area, but the number of round-trips performed in one day by the inhabitants. This means we have more agents in the simulation than there are in the real world, but the agents have shorter day plans than their real world counterparts.

MATSim is currently only able to simulate individual car traffic. Therefore, only agents using the car for transportation were considered for the simulation. Every trip starts and ends at a link in MATSim. However, many traditional DTA use demand on the level of traffic zones, and the corresponding ABDG, including the Kutter-Model, only generate demand on the level of traffic zones, too. Thus we had to assign links to activity locations, respectively to the location where trips start and end.

The following procedure was used to assign links to activity locations. In a first step, each activity is assigned a coordinate. In a second step, the nearest link to this coordinate is searched for and assigned to the activity. This 2-step procedure allows reusing the plans with different networks, as long as the coordinates are maintained. The coordinate is drawn randomly around the centroid of the traffic cell using polar coordinates. By randomly choosing an angle and randomly choosing the distance, the density around the center of the circle is higher than it is on the border of the circle. This corresponds to the different population densities in the center of a village and the rural areas. The maximal distance a point can be located away from the centroid is 0.7 times the distance from the centroid to the nearest neighbor. The factor 0.7 has proved to lead to a good coverage of the circle areas while keeping overlaps low.

Traffic Simulation

Once the agents' plans are available, the simulation process can start. MATSim (Raney and Nagel, 2005) iterates between the traffic flow simulation (sometimes called network loading) and the behavioral modules. The traffic flow simulation moves the agents through the network according to their plans and generates events from which travel times, travel speeds, link densities, and other characteristics can be calculated. At the end of an iteration, each plan is evaluated regarding how successful the agent was performing the (planned) activities, resulting in a score for the plan.

A fixed percentage of agents will re-plan its day plan with one of the behavioral modules. The currently implemented behavioral modules are route finding and time adjustment. Using route finding, agents try to find better routes, but do not change their departure times or the duration of activities. To find better routes, they make use of the events to calculate actual travel times and thus recognize jammed links. Using time adjustment, the departure times and activity durations are modified with the goal to optimize the individuals' plans score. Additional behavioral modules are conceptually clear, but not yet implemented: Activity re-sequencing would change the order of activities (e.g. shopping after work instead of before work), while activity dropping would remove certain activities in an overloaded plan.

Such a system with several different behavioral modules and adjusted scoring algorithms will react to a time-dependent toll by possibly re-arranging the complete day; in consequence, it goes far beyond DTA, which just does route adaptation.

The simulation is stopped when the agents' average score does no longer significantly improve.

Results

The average score of all agents' plans gives usually a good overview of the iterations' progress. In the first iterations, the average score is very low as the system is far away from being relaxed. With ongoing iterations, the agents learn how to avoid traffic jams by choosing different routes or starting their trips at different times of day. Figure 1 shows the average score of the first 100 iterations. Up to iteration 20, the score improves enormously. From iteration 30 to 60, no real improvement can be recognized. But it can be observed that the score jumps up and down. This can be explained by a too large amount of agents re-planning per iteration, searching for the best route in an otherwise relaxed network. When the re-planning agents all choose their best route, they initiate new traffic jams, as many of them will select the same link for similar route-sections. At iterations 60 and 70, configuration settings were modified to reduce the number of re-planning agents and generally reduce the fluctuations. The score improves around iterations 60 and 70 after the settings were changes (see figure 1). As can be seen by the increased average score, a more stable traffic distribution results with fewer agents stuck in jams. But the fluctuations cannot be completely suppressed, which is okay as in real world the non-intuitive behavior of some drivers may lead to similar, seemingly unforeseeable traffic jams.
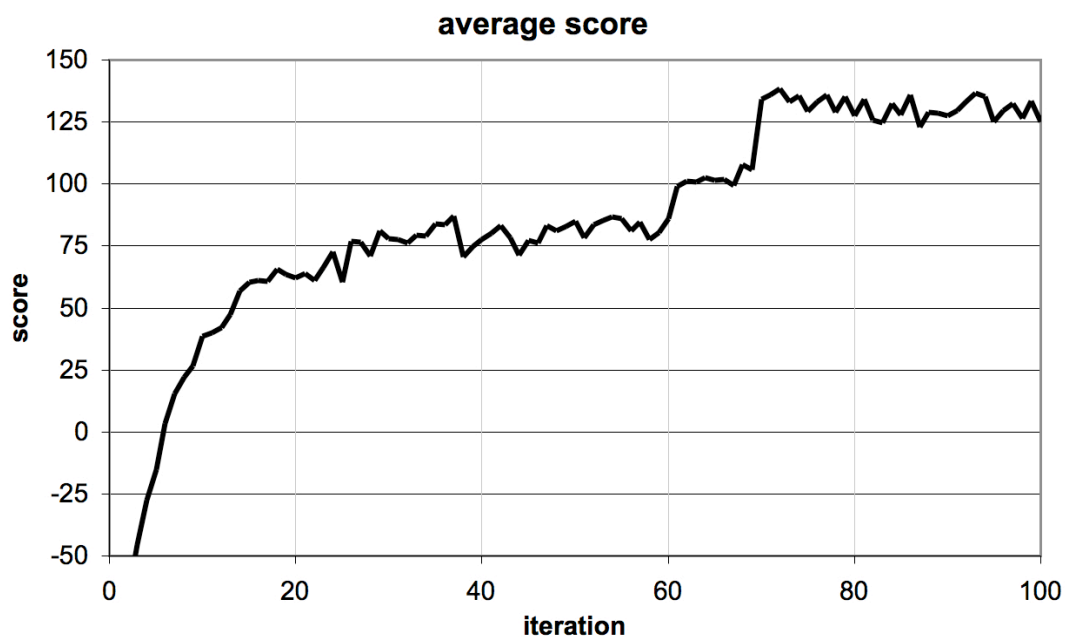

Figure 1: The agents' average score of the first 100 iterations.

Adjusting trip departure times and activity durations are the most efficient ways to get a relaxed system. Initially, all agents are assigned random start time for the first activity and random activity

durations, each within a certain range of time. The range depends on the type of the activity. Each time an agent re-plans, the agent tries to optimize its possible score by re-allocating activity-durations (Meister et al., 2006). This leads to a differentiated distribution of trip departure times. Figure 2 shows the number of trip departures over the course of a day, in the upper part for the initially assigned times in iteration 0 and in the lower part for iteration 100, where the times were optimized during the iterations to nearly reach a relaxed system. The numbers of trip departures are furthermore differentiated between trips of plans having work or education as primary activity, and plans having other primary activities like shopping or leisure. It can be seen that the agents try to avoid traffic jams in the morning by leaving home earlier than initially assigned. Additionally, agents that do not have to work or go to school and thus are more flexible, try to avoid the evening rush hour by performing activities before or after the peak hour.
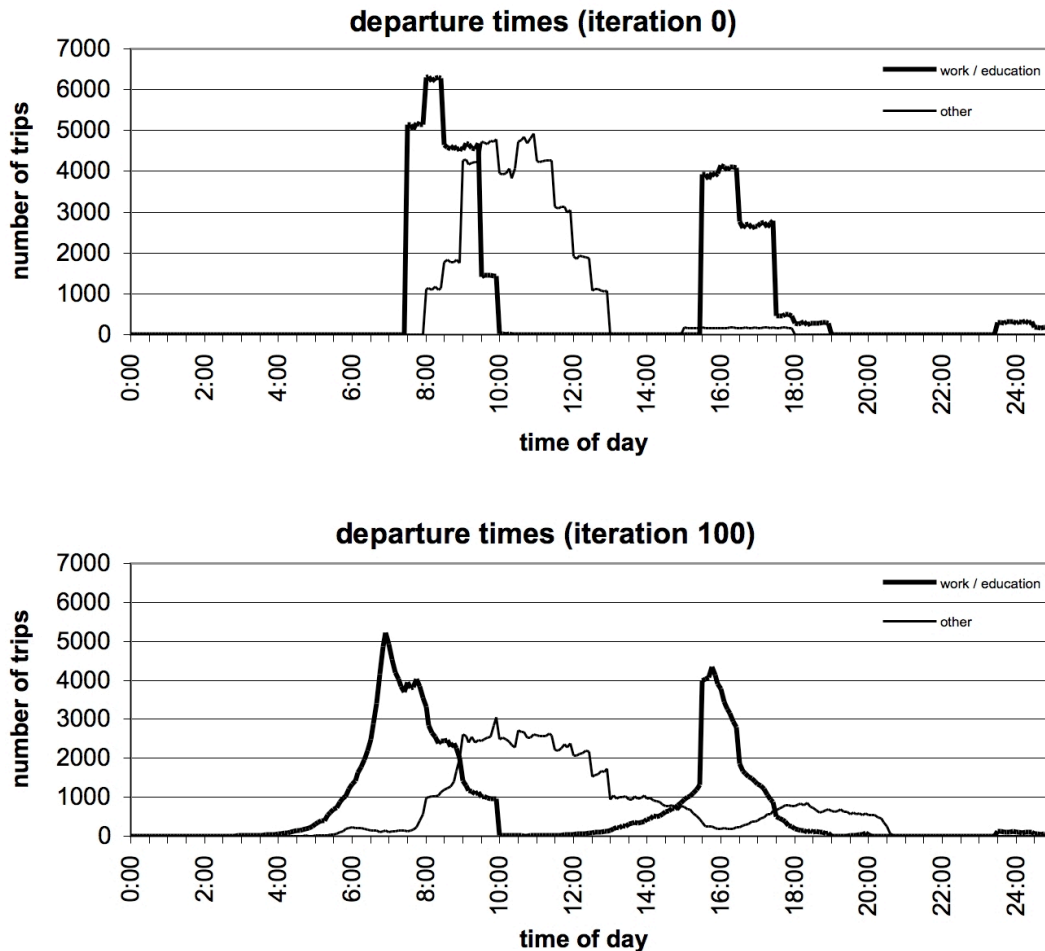

Figure 2: The number of trip departures over the course of a day in iteration 0 (top) and iteration 100 (bottom), differentiated by the type of the primary activity of the corresponding plan.

We have data from about 100 measurement stations in Berlin where the traffic passing by was measured. The counts are available in hourly slices, but not all measurement stations have values for every hour of a day. Only a few stations offer counts for the whole day, while many stations only have counts for the morning rush hour.

We can compare these counts for each measurement station to the number of vehicles that travel across the corresponding links in our simulation during the time period of interest. Alternatively, we can calculate an average volume capacity ratio of all measurement stations based on the links' capacities. For this, we sum up the capacities of all links we have counts for in a specified hour. Next, we sum up the counts for those links. With these two sums, an average volume capacity ratio can be calculated for the specified hour. The same can be done for the number of vehicles on those links in the simulation. Figure 3 shows the two volume capacity ratios in comparison over the course of a day.

It can be seen that the average volume capacity ratio is generally lower in the simulation than it is in the real world, except during the early morning hours. The missing traffic between morning and evening can be explained with the lack of commercial traffic in the Kutter-Model. The overestimation of traffic during the early morning hours may be explained with the input-data only containing round-trips but not complete day plans. Therefore there is no temporal relationship between two round-trips of one person on the same day. It is possible and likely that the times the two activities are performed will overlap. Because of the missing information that one activity can only take place after another, there is currently no need to perform activities late in the day. Instead, the agents try to accomplish the activities during regular work-hours, starting in the morning.
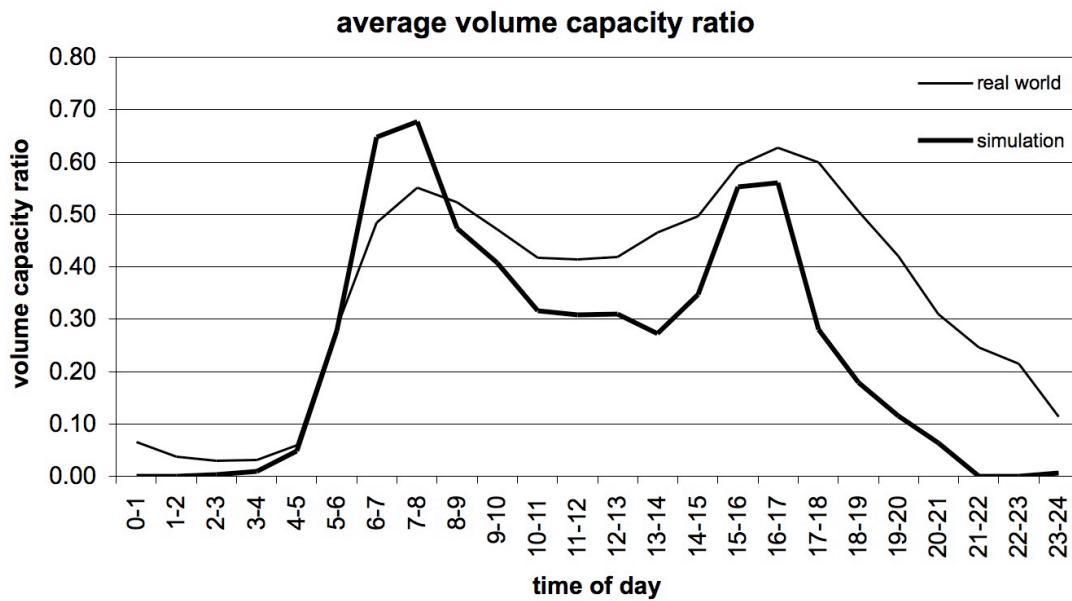

Figure 3: a comparison of average volume capacity ratios over the course of a day.

For each measurement station, a relative error can be calculated for every hour data is available. The relative error is defined as the absolute difference between real world and simulated counts, divided by the real world counts. An average of the relative error over all measurement stations can be plotted. As can be seen in figure 4 the average relative error changes over the course of the day. It is relatively high during the night and improves during the course of the day.
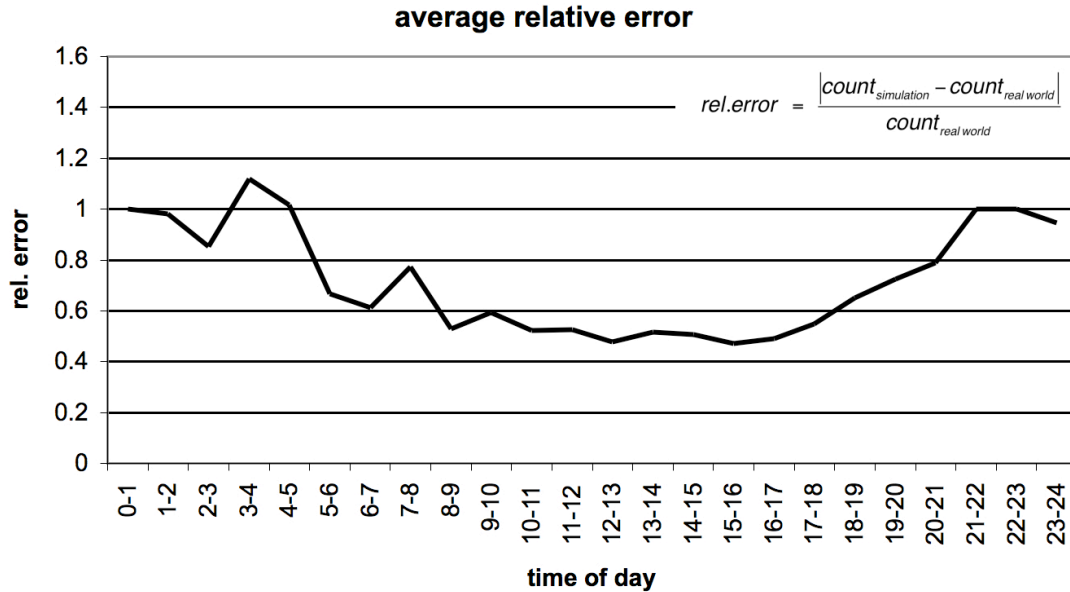
$$\text{rel.error} = \frac{|count_{simulation} - count_{real\,world}|}{count_{real\,world}}$$

Figure 4: Average relative error when comparing real world counts with simulated volumes.

Conclusions and Future Work

Our work has shown that it is possible to couple ABDG with multi-agent traffic simulations. But the results are not yet satisfying. While it is possible to re-use internal data from ABDG, the data has, at least in our case, some severe shortcomings. If data should be used from ABDG for multi-agent simulations, these shortcomings must first be recognized and then overcome. This can either be done by modifying the ABDG itself, or the internal data needs to be more thoroughly post-processed to make it suitable for multi-agent simulations.

A major shortcoming is that each agent in the simulation corresponds to a round-trip, but not to a "real" person. This can be seen in the number of agents (7m agents compared to 6m inhabitants) as well as in the fact that every agent is only at home at the start and the end of a plan, but never in-between two activities (e.g. having lunch at home). This leads to missing temporal relationships between trips and too little traffic in the evening, as shown in chapter "Results". Combining two or more activity chains into one agent reduces the number of agents, while at the same time increases the average complexity of a plan. The combination of several activity chains into one plan must be done carefully and first needs some research first, as not every combination of activity chains has the same probability. But the fewer but longer plans might help to relax the pressure in the morning hours and might lead to an increased amount of traffic later in the day.

If ABDG is modified to include complete day plans, a huge step is made towards agent-based demand generation. Additionally, traditional DTA could make use of the improved ABDG when the traffic assignment is done for a limited time frame only and not for the whole day.

Our results have to be interpreted with regard to Berlin's special history. The partition of the city into two parts by the wall and the reunification led to a city having more than one center. Additionally, the behavior of some parts of the population (mainly those of older generations) still differs based on their origin and historical background. This requires special modifications of the behavioral modules and the algorithm used for scoring. The amount of modifications must yet be figured out.

Additionally, the data for the Berlin scenario must be further improved. Some attributes of the network provided by the planning department of the city of Berlin cannot be reconstructed or fully understood, while other attributes (like the number of lanes per link) are completely missing. We currently use estimations to overcome these shortcomings, but having the correct values from the source of the data would clearly help to improve the results of the simulation. The amount of work needed to get useful data from public authorities is enormous. A simplification of the process as well as relaxed conditions to use data for research purposes could boost the research of transportation planning and traffic simulation in the region of Berlin.

Acknowledgments

References

Arentze, T.A. and H.J.P. Timmermans (2000): Albatross: A Learning-Based Transportation Oriented Simulation System. EIRASS (European Institute of Retailing and Services Studies), Eindhoven, NL.

Arentze, T.A. and H.J.P. Timmermans (2005): ALBATROSS – Version 2.0 – A learning based transportation oriented simulation system. EIRASS (European Institute of Retailing and Services Studies), TU Eindhoven, NL.

Balmer, M., M. Rieser, A. Vogel, K.W. Axhausen, and K. Nagel (2005): Generating day plans using hourly origin-destination matrices. In T. Bieger, C. Laesser, and R. Maggi, editors, Jahrbuch 2004/05 Schweizerische Verkehrswirtschaft. St. Gallen. pages 5–36.

Bhat, C.R. J.Y. Guo, S. Srinivasan, and A. Sivakumar (forthcoming): A comprehensive econometric microsimulator for daily activity-travel patterns (cemdap). Transportation Research Record. Also see www.ce.utexas.edu/prof/bhat/FULL_REPORTS.htm.

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

Kutter, E. (1984): Integrierte Berechnung städtischen Personenverkehrs – Dokumentation der Entwicklung eines Verkehrsberechnungsmodells für die Verkehrsentwicklungsplanung Berlin (West). Berlin.

Kutter, E., H-J. Mikota (1990): Weiterentwicklung des Personenverkehrsmodells Berlin auf der Basis der Verkehrsentstehungsmatrix 1986 (BVG). Untersuchung im Auftrag des Senators für Arbeit, Verkehr und Betriebe. Berlin.

Kutter, E., H-J. Mikota, J. Rümenapp, and I. Steinmeyer: Untersuchung auf der Basis der Haushaltsbefragung 1998 (Berlin und Umland) zur Aktualisierung des Modells "Pers Verk Berlin / RPlan", sowie speziell der Entwicklung der Verhaltensparameter '86 - '98 im Westteil Berlins, der Validierung bisheriger Hypothesen zum Verhalten im Ostteil, der Bestimmung von Verhaltensparametern für das Umland. Entwurf des Schlussberichts im Auftrag der Senatsverwaltung für Stadtentwicklung Berlin. Berlin / Hamburg.

Meister, K., M. Balmer, K.W. Axhausen, K. Nagel (2006): planomat: A comprehensive scheduler for a large-scale multi-agent transportation simulation. In Swiss Transport Research Conference (STRC).

Miller E.J. and M.J. Roorda (2003): A prototype model of household activity/travel scheduling. Transportation Research Record, 1831:114–121.

Pendyala R.M. and R. Kitamura (In press): FAMOS: The Florida Activity Mobility Simulator. In H.J.P. Timmermans, editor, Progress in activity-based analysis. Elsevier, Oxford, UK.

Raney, B. and K. Nagel (2005): An improved framework for large-scale multi-agent simulations of travel behavior, in TRB (2005).

Rümenapp, J and I. Steinmeyer (2006): Activity-based demand generation: Anwendung des Berliner Personenverkehrsmodells zur Erzeugung von Aktivitätenketten als Input für Multi-Agenten-Simulationen. Arbeitsberichte Verkehrssystemplanung und Verkehrstelematik Nr. 06-09. TU Berlin.

TRB (Ed.) (2005): TRB 2005 Annual Meeting CD-ROM, Transportation Research Board, Washington, D.C.