Integration of Activity-Based with Agent-Based Models: an Example from the Tel Aviv Model and MATSim

Shlomo Bekhor
Faculty of Civil and Environmental Engineering
Technion - Israel Institute of Technology
Haifa, 32000, Israel
Phone: 972-4-829 24 60, Fax: 972-4-829 57 08
sbekhor@technion.ac.il

Christoph Dobler
Institute for Transport Planning and Systems, ETH Zurich
HIL F 34.1, Wolfgang-Pauli-Str. 15, CH-8093 Zurich, Switzerland
Phone: 41-44-633 65 29, Fax: 41-44-633 10 57
dobler@ivt.baug.ethz.ch

Kay W. Axhausen
Institute for Transport Planning and Systems, ETH Zurich
HIL F 31.1, Wolfgang-Pauli-Str. 15, CH-8093 Zurich, Switzerland
Phone: 41-44-633 39 43, Fax: 41-44-633 10 57
axhausen@ivt.baug.ethz.ch

TRB Paper number: 11-0997
Revised version: November 12, 2010
ABSTRACT

This paper explores the possibility of integrating an operational activity based model with a dynamic traffic assignment framework. For this purpose, the Tel Aviv activity-based model and parts of the functionality of the MATSim agent-based framework are used, in an attempt to use the best features of both approaches: on one side, the disaggregate demand representation from the activity-based model, and on the other side, the disaggregate supply representation of the agent-based framework.

The paper uses the person activity schedule produced by the activity based model directly, thus eliminating the need to aggregate them into origin-destination matrices. The paper compares results produced by this combination against a static assignment of the Tel Aviv model that shows a very good fit at the aggregate level. The purpose of the paper is to further advance the fully disaggregate implementation of activity-based models. In this aspect, this paper represents a further step towards this general goal.
INTRODUCTION

Activity scheduling models based on a jointly modeled set of choices have become increasingly popular in the last decade, and have been implemented for various policy applications. Examples of such choice-based models in the U.S. can be found in Portland (1), San Francisco (2), New York (3), Dallas-Fort Worth (4), Florida (5), to name a few; outside the US there are also several implementations such as Jakarta (6), Tel Aviv (7). The rule-based approach in the Netherlands (8) and integrated activity scheduling driven demand and traffic flow applications in Switzerland (9) follow different paths to activity-based modeling.

Although these models differ in their structure, input data and aggregation level, it is possible to depict a common structure for entire activity based model applications, which is presented in Figure 1.

In most activity-based models, the "raw" output is the individual's list of activities and trips that include detailed information about departure time, destination and mode for each trip and maybe stage. This detailed output is aggregated into origin-destination (OD) matrices needed for the highway and transit assignments. These assignments can be either static (as in most models) or dynamic (as recently in Lin et al. (10)). The assignment outputs are traffic volumes and travel times, which in turn are used as inputs to the activity based models.

The TRANSIMS software (11) was originally designed to account for the full disaggregate representation of individual travel behavior. However, in a recent paper by Lawe et al. (12), the model was applied using OD matrices. The authors used TRANSIMS as router and micro-simulator using OD matrices for a given area.

The main reason for decoupling the demand side of the problem (the activity based models) from the supply side of the problem (either assignment or simulation models) is that the activity models typically compute probabilities for a large number of alternatives, which demands an explicit choice set. To account for such alternative sets in assignment or simulation procedures for real size networks would result in very long computation times.

The possibility of integrating an operational activity based model with a dynamic traffic assignment framework has been recently explored. Castiglione et al. (13) describes the integration of DaySim, an activity-based travel demand forecast model developed for the Sacramento region, with the TRANSIMS Router, a disaggregate dynamic network assignment tool. Gao et al. (14) undertook a similar exercise, but this time using the Toronto activity based-model. Their comparison indicates that the results produced by agent-based toolkit are not only compatible to those by static assignment but are more realistic from a temporal point of view. Hatzopoulou et al. (15) explored the potential integration of an existing activity-based travel demand model (TASHA) with the agent-based toolkit for emission modeling. In both studies, the starting points are the OD matrices produced by the activity-based models.

The present paper further extends this line of research, by using the person activity schedule produced by the module directly, thus eliminating the need to aggregate OD matrices. The purpose of the paper is to further advance the full disaggregate implementation of activity-based models. In this aspect, this paper represents a first step towards this goal. In this paper, we use the Tel Aviv Activity Model and the MATSim agent-based toolkit (16), in an attempt to use the best features of both frameworks: on one side, the disaggregate demand representation from the activity-based model, and on the other side, the disaggregate supply representation from the agent-based toolkit.

The rest of this paper is organized as follows. The next sections briefly describe the activity based and the agent based models. The methodological section shows the first steps performed and the subsequent section compares results of the existing Tel Aviv model with the new combined model. The final section of the paper discusses challenges of future steps.

THE TEL AVIV ACTIVITY-BASED MODEL
The Tel Aviv activity-based model system comprises a hierarchy of logit and nested logit models for the main stop of the tour - namely, activity type, time of day, destination, and mode. The intermediate stops are then modeled conditional on the main stop models. The Israel National Travel Habits Survey (NTHS) conducted in 1996 is the primary data source for model development. The Tel Aviv model accounts for up to two tours for each person: the most important tour of the day, referred to as the primary tour, and the second most important tour of the day, or the secondary tour.

The application of the Tel Aviv model contains four main functional units: Population Generator, Activity Generator, Trip Generator, and Trip Assignment. The distribution of sub-models between these units was based primarily on the computational efficiency and programming convenience, rather than on the logical interrelations between sub-models. A detailed discussion of model development and application can be found in (17).

The Activity Generator Unit applies the activity-based model of travel behavior to each person included in the synthetic population sample. As a result, each person’s daily travel is fully described and includes the types and number of daily tours, the number of intermediate stops, the destinations of each activity, the modes used in different parts of each tour, and the time of day that the individual travels. Table 1 shows the attributes that are produced by the Activity Generator Unit. The structure of the file is similar to a typical household travel survey, and each record corresponds to a person in the study area. Note that the number of vehicles in household is part of the Activity Generator model.

The Trip Generator Unit summarizes individual trips into mode-specific demand matrices for different periods of day. These matrices are then used as input to the Trip Assignment Unit, which performs all necessary auto and transit assignments by periods of day. In the present paper, we will use directly the activity-based file produced by the Activity Generator Unit.

THE MATSIM AGENT-BASED TOOLKIT

MATSim is an agent-based travel demand modeling framework that operates on the basis of individual agent plans; a plan being a schedule of activities, their locations and the travel connecting them.

Figure 2 illustrates its basic principle. An initial demand of full day plans of activities for each agent is generated (i.e. including the routes chosen) and executed in a mobility simulation. Plans are scored after the simulation step and, based on the score, agents adapt their plans in response to conditions that arose during the simulation.

A MATSim simulation converges to a state analogous to the user equilibrium through a process of systematic relaxation (18). Such convergence is achieved through adapting and deriving a set of feasible plans for each agent from their original initial plan. As these sets of plans grow to a limiting number, badly performing plans are discarded. Consequently, each agent’s set of feasible plans improves with increasing iterations. Feasible new plans can be derived from existing ones by changing activity timings, locations, re-routing travel between activities, changing transport modes connecting activities or dropping activities from the activity schedule altogether.

METHODOLOGY

This section describes the initial steps performed to integrate the Tel Aviv Activity Model in MATSim. Using the same general framework presented in Figure 1, two possible frameworks are outlined in Figure 3.

The first approach is simpler, in which the toolkit is used as a replacement to the static assignment block. In the specific case of the Tel Aviv model, the departure time choice could be directly integrated in this block, because the current model application uses aggregate time periods, as explained below. At this stage the feedback between the two
models (Tel Aviv and MATSim) is performed manually, so it is not possible to compare run times.

The second approach involves the adaptation of the utility functions estimated in the activity-based models to the dynamic assignment environment. The current MATSim implementation contains heuristic utility functions with limited set of variables, most of them related to the network level of service. In addition, there is a challenge involved in calculating the individual plans based on explicit choice sets. In the current version, the individual plans are changed according to the scoring function based on genetic algorithms (19), which are difficult to be adapted to accommodate explicit large choice sets. This logic follows the logic of assignment modeling, where the choice set is built incrementally over the iterations of the equilibrium search.

In the present paper, we limit ourselves to the iterative approach, and the next subsections describe the steps needed to convert the static parts of the Tel Aviv model to the dynamic environment.

**Departure Time Disaggregation**
In the Tel Aviv model system there are two separate models that determine the activity starting and duration. The main model is a logit model accounting for the joint choice of time period of departure from home and from the main activity stop of the tour. The time periods modeled include early morning (MO – 03:00 to 06:30), a.m. peak (AM – 06:30 to 08:30), midday (MD – 08:30 to 15:00), p.m. peak (PM – 15:00 to 20:00), and late evening (EV – 20:00 to 03:00). A total of fifteen time period combinations cover a full day, as defined in Table 2.

[Table 2 here]

The second model is a detailed time of day model developed by Popuri et al. (20). This model is a post-processor to the travel demand model system and is designed to be a policy analysis tool. This model looks at the time-of-day choices at a greater level of temporal resolution for automobile users. In the current version, we only made use of the first model, since the second one was not implemented yet.

Since the agent-based toolkit allocates individual trips dynamically, it is necessary to provide precise departure times. This was achieved by disaggregating the 15 periods considering activity duration constraints. These constraints are the feasible time windows of an activity. For primary activities, the following assumptions were made, based on average values found in the NTHS survey: 8 hours and 4 hours for full-time and part-time workers, respectively; 5 hours for education activities; 3 hours for shopping and other activities. For secondary activities, it was assumed that each activity purpose last two hours shorter than the primary activity. The remaining hours are either spent traveling, which is calculated using the travel flow simulator, or at home. Assuming initial departure times at the half-points of each time period, the model allocates the best departure times, accounting for the constraints mentioned above.

**Traffic Analysis Zones Disaggregation**
The Tel Aviv Metropolitan Area is divided into 1,219 traffic analysis zones (TAZ). The agent-based model represents each facility (dwelling unit, workplace, etc.) along the links that compose the network. To keep consistency between the approaches, but also to take advantage of the more disaggregate supply representation, the aggregate zonal values were distributed proportionally in the links that form each TAZ. Only local and collector streets were included, to avoid placing facilities onto freeways. The allocation was proportional to link length, so longer links will have more facilities along them.

This procedure was performed with purpose-built code using the GIS database of the Tel Aviv model that includes zone boundaries, zonal characteristics and highway network. First, each road link is assigned to one and only one zone. If a link crosses multiple zones, it is assigned to that zone where it is predominantly located. Since the planning network from the Tel Aviv model already contains connector links for each traffic zone, they are assigned to their corresponding zone, regardless of their link length. This ensures that each zone at least has one assigned link.
The facilities within a zone are equally distributed among the links of a zone. When doing so, some links are ignored because on some link types (e.g. freeways) no facilities are expected. The network coding of the Tel Aviv model was used to identify the link types. Since there is no specific information about facility locations, it is assumed that the facility distribution respects the length of the links, in order to keep the facility density uniform along a link. This procedure ensures that the trips departing or arriving at a certain zone do not converge to a single entry/exit point, as in planning networks.

The Tel Aviv Metropolitan Area has 3.2 million inhabitants. The Tel Aviv model currently runs with 10% of the full population to avoid long run times. To keep consistency, the same sample was used for the integrated runs.

**Network Conversion and Level of Service Computation**

The Assignment Unit of the Tel Aviv model uses EMME/2 software for performing traffic and transit assignments. There are no major issues related to the conversion of EMME/2 formats to MATSim formats, and details can be found in Gao et al. (14). To maintain consistency, the same link free-flow travel times and capacities of the static assignment were employed. Turn penalties are converted in the dynamic network by expanding the nodes containing turn restrictions to a set of nodes that are interconnected according to the given restrictions. The converted network has 7,879 nodes and 17,118 links. The model parameters related to the dynamic assignment were calibrated using aggregate traffic counts and data from the static assignment network in the Tel Aviv model (free-flow travel times and capacities).

The Activity Generator unit of the Tel Aviv model, as with other activity-based models, needs level of service data from network assignment models. In the current implementation, transit travel times were still provided from the EMME/2 assignments, therefore only car travel times were provided from the dynamic assignment model. In order to feed the Activity Generator unit, a procedure that computes travel time matrices from MATSim was added. To keep consistency between the two model components, the travel times on connector links to the traffic analysis zones were kept constant.

**RESULTS**

This section presents selected results from the combined models. The first result is obtained with the original Tel Aviv model. Figure 4 compares the distribution of primary tours by time period obtained from the Tel Aviv model and the survey. As expected, most tours correspond to AM-MD (morning peak to mid-day) and AM-PM (morning peak to afternoon peak) periods, which are respectively related to the Education and Work activities. The labels on the x-axis correspond to the 15 period combinations.

The results of the discretization process are displayed in Figure 5, which shows the distribution of the agents (individuals) for all trip modes by time of day. As indicated in the methodology section, the model starts the run assuming the departure time equal to the half-point of each time period, and tries to find the best score for each individual. For example, it is assumed that the first departure in the early morning period occurs at 04:30. After several iterations, the dynamic iterative model reaches a condition in which the combined departure time and route choice for each individual cannot be improved further, given the constraints imposed on activity duration and starting activity times. After 24 hours (that is, hour 28:30 in Figure 5) all agents should be at home.

The traffic flows obtained from dynamic assignment were compared against the original static assignment Tel Aviv model results. For ease of comparison, the dynamic assignment results were aggregated by hour. Figure 6 provides a qualitative view of the flow difference between the two models on the main roads of the metropolitan area, for the AM peak hour. The green signs are dominant, indicating that the differences between the two models are relatively small (up to 20% deviation).
Figure 7 provides a quantitative comparison of the link flows on main roads for the morning peak hour. The regression line is a bit skewed, meaning that the dynamic assignment flows on these roads are smaller compared to the static model.

To investigate the reason for the differences in traffic flows, we first compared the flows produced by the models with respect to traffic counts. Table 3 presents a summary of the comparison for 2 main screen lines of the metropolitan area: Highway 4, which runs in the North-South direction, and Highway 5, which runs East-West. The traffic counts were obtained from regular counts performed at the junctions that cross these highways (22). The overall results show a very good match against traffic counts.

To investigate the reason for the differences in traffic flows, we first compared the flows produced by the models with respect to traffic counts. Table 3 presents a summary of the comparison for 2 main screen lines of the metropolitan area: Highway 4, which runs in the North-South direction, and Highway 5, which runs East-West. The traffic counts were obtained from regular counts performed at the junctions that cross these highways (22). The overall results show a very good match against traffic counts.

Table 3 here

Figure 8 shows that the regression slope for the link flow comparison on screen lines is close to 1, compared to the results presented in Figure 7.

The MATSim dynamic assignment provides different link flows compared to the static assignment. This is expected given the high congestion at the AM peak hour, and consequently several links in the static assignment are oversaturated. The dynamic model includes a queuing representation that better reflects the congestion effect, compared to the simple volume-delay functions used in the static assignment.

Further inspection of the simulation results reveals the explanation for the link flow differences. Figures 9 and 10 respectively display two simulation snapshots for 7:00 AM and 7:30 AM. Because of space limits, only these two time slots are shown. The green dots represent agents traveling on non-congested links, and the red dots represent agents traveling on congested links. A congested link is defined when the speed falls below 10% of the free-flow speed. The highway crossing the area in the East-West direction is Highway 5, used in the screen line comparisons.

As expected, there are more trips on the network at 7:30 AM compared to 7:00 AM. In contrast to the 7:00 snapshot, which has few congested links, the 7:30 snapshot shows that in certain links the flow is close to capacity, which causes some spillback. Further inspection on the hourly results obtained from the simulation model shows a good match with the traffic counts. The overall result is closer to traffic counts in comparison to the static assignment model, because the link flows obtained in the static assignment are higher than the capacity, a well-known deficiency of any static assignment model.

DISCUSSION
This paper reports the first step towards the integration of an activity based model in an agent-based dynamic framework. The paper benefits from two existing applications, the Tel Aviv model and the MATSim toolkit. Although at this stage is not possible to compare run times, the dynamic assignment in MATSim is quite fast for the medium size network considered: it takes about 2 hours to perform 100 iterations. The Activity Generator Unit of the Tel Aviv model needs about 1 hour per iteration to create the activity schedules. This means that the combined model runs in a reasonable amount of time.

The paper focused on the comparison of the aggregate results, to show that the more detailed behavioral representation can be also used for general planning as well as more detailed case studies. Similar to Gao et al. (14) we found a very good match at the aggregate level. Note however, that in their paper the only re-planning strategy allowed was re-routing. In the current paper, both re-routing and activity timings were adapted by the agent-based toolkit.

This paper shows that the full activity list can be directly employed, without the need to create origin-destination matrices. This feature is one of the reasons for improved run times, as indicated in Rieser et al. (23).

The next step towards a full integration is the inclusion of the utility functions developed for the activity-based models in the agent-based toolkit. For the Tel Aviv model,
we are currently working on the destination choice problem, since there is a location choice module available in MATSim (24). The idea is to replace the search space by a pre-calculated destination probability distribution, which would be computed in accordance with the activity based model. The initial results are promising and will be reported in a subsequent paper.

REFERENCES

12. Lawe, S., J. Lobb, A. W. Sadek, S. Huang, and C. Xie. TRANSIMS Implementation in Chittenden County, Vermont: Development, Calibration, and Preliminary


LIST OF TABLES AND FIGURES

TABLE 1 Activity Based Items in the Tel Aviv Model
TABLE 2 Definition of Time Period Combinations
TABLE 3 Screen Line Comparison – AM Peak Hour

FIGURE 1 Overall activity-based model application structure
FIGURE 2 MATSim basic structure
FIGURE 3 Proposed modeling frameworks
FIGURE 4 Distribution of trips by time period
FIGURE 5 Distribution of agents by time of day
FIGURE 6 Comparison between dynamic and static assignment results – AM peak hour
FIGURE 7 Link flow comparison on main roads – AM peak hour
FIGURE 8 Link flow comparison on screen lines – AM peak hour
FIGURE 9 Snapshot of the simulation at 7:00
FIGURE 10 Snapshot of the simulation at 7:30
**TABLE 1 Activity Based Items in the Tel Aviv Model**

<table>
<thead>
<tr>
<th>Variable type</th>
<th>Variable description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Personal and household data from the Population Generator</strong></td>
<td>Person ID</td>
</tr>
<tr>
<td></td>
<td>Household ID</td>
</tr>
<tr>
<td></td>
<td>Expansion factor for given record</td>
</tr>
<tr>
<td></td>
<td>Age of person</td>
</tr>
<tr>
<td></td>
<td>Gender Indicator</td>
</tr>
<tr>
<td></td>
<td>Student Indicator</td>
</tr>
<tr>
<td></td>
<td>Years of study</td>
</tr>
<tr>
<td></td>
<td>Industry of employment</td>
</tr>
<tr>
<td></td>
<td>License holder Indicator</td>
</tr>
<tr>
<td></td>
<td>Household size</td>
</tr>
<tr>
<td></td>
<td>Number of employed persons in household</td>
</tr>
<tr>
<td></td>
<td>Number of licensed drivers in household</td>
</tr>
<tr>
<td></td>
<td>TAZ of residence</td>
</tr>
<tr>
<td></td>
<td>Employment status</td>
</tr>
<tr>
<td></td>
<td>Number of children in household</td>
</tr>
<tr>
<td><strong>Variables calculated by the Activity Generator</strong></td>
<td>Number of vehicles in household</td>
</tr>
<tr>
<td></td>
<td>Main activity, primary tour</td>
</tr>
<tr>
<td></td>
<td>Primary tour combined Time of Day</td>
</tr>
<tr>
<td></td>
<td>TAZ of main destination, primary tour</td>
</tr>
<tr>
<td></td>
<td>Main mode, primary tour</td>
</tr>
<tr>
<td></td>
<td>Intermediate stops for primary tour</td>
</tr>
<tr>
<td></td>
<td>Activity at stop before main destination</td>
</tr>
<tr>
<td></td>
<td>Activity at stop after main destination</td>
</tr>
<tr>
<td></td>
<td>TAZ of intermediate stop before main destination</td>
</tr>
<tr>
<td></td>
<td>TAZ of intermediate stop after main destination</td>
</tr>
<tr>
<td></td>
<td>Mode switch at main destination</td>
</tr>
<tr>
<td></td>
<td>Main activity, secondary tour</td>
</tr>
<tr>
<td></td>
<td>Main mode, secondary tour</td>
</tr>
<tr>
<td></td>
<td>TAZ of main destination, secondary tour</td>
</tr>
<tr>
<td></td>
<td>Secondary tour combined TOD</td>
</tr>
<tr>
<td></td>
<td>Intermediate stops for secondary tour</td>
</tr>
<tr>
<td></td>
<td>TAZ of intermediate stop before main destination, secondary tour</td>
</tr>
<tr>
<td></td>
<td>TAZ of intermediate stop after main destination, secondary tour</td>
</tr>
</tbody>
</table>
### TABLE 2 Definition of Time Period Combinations

<table>
<thead>
<tr>
<th>Departure time period</th>
<th>Arrival time period</th>
</tr>
</thead>
<tbody>
<tr>
<td>MO – 03:00 to 06:30</td>
<td>MO 1, AM 2, MD 3, PM 4, EV 5</td>
</tr>
<tr>
<td>AM – 06:30 to 08:30</td>
<td>MO 6, AM 7, MD 8, PM 9</td>
</tr>
<tr>
<td>MD – 08:30 to 15:00</td>
<td>MO 10, AM 11, PM 12</td>
</tr>
<tr>
<td>PM – 15:00 to 20:00</td>
<td>MO 13, AM 14</td>
</tr>
<tr>
<td>EV – 20:00 to 03:00</td>
<td>EV 15</td>
</tr>
</tbody>
</table>
TABLE 3 Screen Line Comparison – AM Peak Hour

<table>
<thead>
<tr>
<th>Screen Line</th>
<th>Direction</th>
<th>Traffic Counts</th>
<th>Static assignment</th>
<th>Dynamic assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>volumes</td>
<td>Deviation</td>
<td>volumes</td>
</tr>
<tr>
<td>Highway 4</td>
<td>Eastbound</td>
<td>30,416</td>
<td>28,982</td>
<td>-5%</td>
</tr>
<tr>
<td></td>
<td>Westbound</td>
<td>39,507</td>
<td>42,230</td>
<td>7%</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>69,923</td>
<td>71,212</td>
<td>2%</td>
</tr>
<tr>
<td>Highway 5</td>
<td>Northbound</td>
<td>15,344</td>
<td>15,244</td>
<td>-1%</td>
</tr>
<tr>
<td></td>
<td>Southbound</td>
<td>24,814</td>
<td>23,515</td>
<td>-5%</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>40,158</td>
<td>38,758</td>
<td>-3%</td>
</tr>
<tr>
<td>Grand Total</td>
<td></td>
<td>110,081</td>
<td>109,970</td>
<td>0%</td>
</tr>
</tbody>
</table>
FIGURE 11 Overall activity-based model application structure
FIGURE 12 MATSim basic structure
FIGURE 13 Proposed modeling frameworks
FIGURE 14 Distribution of trips by time period
FIGURE 15 Distribution of agents by time of day
FIGURE 16 Comparison between dynamic and static assignment results – AM peak hour
Dynamic assignment volumes
Static assignment volumes

FIGURE 17 Link flow comparison on main roads – AM peak hour

y = 0.8187x + 142.68
R^2 = 0.8204
$$y = 0.9315x + 132.49 \quad R^2 = 0.9047$$

FIGURE 18 Link flow comparison on screen lines – AM peak hour
FIGURE 19 Snapshot of the simulation at 7:00
FIGURE 20 Snapshot of the simulation at 7:30