MICROSCOPIC MODELING OF TRAVEL-DEMAND: APPROACHING THE HOME-TO-WORK PROBLEM

Peter Wagner, "NMI", German Aerospace Research Establishment
Kai "NMI" Nagel, TSASA

Transportation Research Board Annual Meeting
Washington, DC
January 1999
DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.
DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.
Microscopic Modeling of Travel-Demand: Approaching the Home-To-Work Problem

Peter Wagner
German Aerospace Center (DLR)
Porz-Wahnheide, Linder Höhe,
D-51147 Köln, Germany
eMail: peter.wagner@dlr.de

Kai Nagel,
Los Alamos National Laboratory (LANL),
P.O. Box 1663, MS M 997,
Los Alamos, NM 87545
eMail: kai@lanl.gov

Abstract

In this article the results are described that have been found when tackling the problem of the assignment of employees to their working places (destination assignment) by using a truly microscopic approach, whose output is suitable for a microscopic traffic simulation. This problem is dealt with a microscopic stochastic analogue of the gravity ansatz of transportation planning, described in this article.

However, the computation of the travel destinations is only the first step in a sequence of simulation steps. Its output will be used to compute a simulation-based dynamic traffic assignment (route assignment), resulting in travel times needed from home to work for any traveler. Those travel times will be used in a further re-assignment step, where any traveler whose travel-time has exceeded a certain limit is subject to re-assigning a new working place (destination). This creates a sequence of re-assignment and re-routing processes, whose results will be reported in this article. The results obtained show that approach presented in this article is capable of describing the destination and route choices microscopically.
1. STATEMENT OF THE PROBLEM

The basic object used in traffic planning is the origin/destination (OD-) matrix. When using a dynamical microscopic simulation model for computing the time-dependent traffic loads in a network, the OD-matrix has to be decomposed into individual travelers. Therefore it can be asked whether it is more efficient and conceptually simpler to generate in the first steps of the traditional four-step algorithm (1) not OD-matrices, but individual trips (described by an origin, a destination, and a time interval during which the activity associated with the trip starts). The trips may be used directly in the subsequent steps of a dynamical traffic assignment and a microscopic simulation, respectively.

Although promising approaches exists (see (2-7) and references therein), the modeling of the generation of activity chains from household data is not solved satisfactorily. In order to gain experience with those approaches, in the following the simpler problem of the assignment of employees to their respective working places is analyzed. In order to distinguish this assignment from the route assignment of the dynamic traffic assignment algorithm, the term “destination assignment” is used in the following. The route assignment is a subsequent step, leading to travel times that are then again input into the destination choice algorithm. The work described here depends on the availability of an excellent data-basis, being the case for Portland, which is the current focus of the TRANSIMS project (8). However, it can be expected, that similar data are available for other cities in the near future as well.

This study has used the following input data:

- reduced road network of Portland with 8,564 nodes and 20,024 links,
- data from the census in 1990 which were input to an algorithm (9) to generate synthetic households with 1,415,900 household members, of which roughly 660,000 are employees,
- a detailed data-set containing the locations of the employers of Portland with the approximate number of working places.
The working places as well as the employees (given by the synthetic households) are associated with the links in the network. The Portland metropolitan area stretches across two states, Oregon and Washington. For the part in Washington State (Vancouver), employer information was not available. The missing information was generated by counting the numbers of all necessary working-places from the demographic data (as given by the synthetic household data) and subtracting the number of working places in Oregon. The remaining working places were randomly assigned to links in the area of Washington State in proportion to the population density. The synthetic household data additionally contain the information in which PUMA (Public-Use Micro-census Area, see Figure 1) the employees are working. Although the spatial decomposition of Portland into the PUMA’s is a very rough one, this information will be used in order to judge the quality of the assignment that is to be described next.

2. TRAVEL TIME DISTRIBUTIONS

The synthetic household data contain additional information about when people start their work, and their actual travel duration time. The appropriate description of this duration is a frequency distribution, shown in Figure 2. There the number of persons \( p(\tau) \) that need a certain time \( \tau \) to go to work is plotted. In order to utilize this distribution in the assignment process below, a parameterization has been obtained by a non-linear least-squares fit of the integrated distribution to the integral of a beta-distribution:

\[
P(\tau) = \int_0^\tau p_\beta(\tau') d\tau' \propto \int_0^\tau \tau'^\alpha (T_{\text{max}} - \tau')^\beta d\tau' \tag{1}
\]

The parameters from the fit are \( \alpha = 0.83, \beta = 6.03, \) and \( T_{\text{max}} = 100 [\text{min}] \). Worth mentioning is that it is very unlikely that very short travel times occur. In most gravity-type approaches a distance function that is monotonously decreasing is used (although in most cases a cut-off is utilized to eliminate short travel-times). This would probably result in much too localized traffic flow patterns (see text below, and Figure 3).

3. MATHEMATICAL FORMULATION OF THE PROBLEM
As already described, the data-set available for Portland contains information about the number of working places \( n_i \) at each spatial index \( i \). The spatial indices are associated with the links of the network. The synthetic household data contain the number of employees \( E_i \), where again the index \( i \) refers to a spatial location. What has to be computed is, for any of the employees given, her or his respective working place. It is assumed in the following that the probability for choosing a certain working place for a certain employee is dependent on the travel costs, with higher costs implying a lower probability of working there. However, as can be seen from the travel time distribution above, the actual dependence follows a more complicated law (which is not monotinous). Independent on how this assignment is finally done, the result has to reflect this travel time function in order to generate a reasonable traffic flow pattern, and in order to be consistent with the data. The problem becomes more complicated by the fact that the travel-times in the loaded network, not the ones of the empty network have to be utilized for the assignment. We solve this by using an iterative approach. This approach consists of two nested iterative loops, one for the destination (activities) assignment, and one for the route assignment once the activity locations are given (see Figure 4). The route assignment is a standard microscopic dynamic assignment, where micro-simulation and router iterate until individual routes cannot be improved any more (10,11). Once these routes have been found, the resulting link delays are used to reassign a part or all of the employees. The whole process has to be repeated until convergence is reached. The next section will report on those simulations. Note, that there exists a different approach to this problem, since it can be cast into the form of a classic transportation science problem, i.e. Hitchcock's problem. The results obtained by using this approach will be described elsewhere.

3.1. Approach Used: Stochastic Gravity Assignment

The approach described in the following is better adapted to the microscopic nature of the problem and is simpler than solving Hitchcock's problem exactly. Before going into the very details, an important observation has to be made. The distribution \( p(x) \) of all travel-times for all possible OD-combinations in the Portland network, which can be estimated from a computation of all shortest paths, is already similar to the travel time distribution \( p(t) \). The \( p(x) \)-distribution is well parameterized by the product of a Gaussian and a power law:
The power-law is due to the local structure of a city, while the Gaussian part of this distribution describes the large-scale structure. The parameter \( \mu \) gives a typical distance, while the parameter \( \gamma = 1.7 \) describes the simple fact that the probability to find a point in distance \( x \) is proportional to \( x^\gamma \) if the link-density in the city is constant. (Except, of course, in the border regions of a city.) This fact will be exploited to simplify the destination assignment procedure considerably, see below.

As already mentioned, a number of households as well as a number of working places is associated with each link in the Portland network. Labeling these links by an arbitrary index \( i \), the assignment proceeds by performing the following steps for any \( i \):

1. Attractions \( a_j \) are computed for any of the other links with \( a_j = W_j q(t_j) =: \nu_j \) where \( W_j \) is the number of working places associated with link \( j \), \( t_j \) are the travel times based on the shortest paths in the network and \( q(\cdot) \) is any travel time function.

2. The \( a_j \) are interpreted as probabilities that this link \( j \) is chosen, so any of the employees \( E_j \) in link \( i \) is associated according to these probabilities to one of the working places.

3. After assigning a person from \( i \) with a working place in \( j \), the number of working places still available for assignment is reduced by one.

4. After having assigned any of the employees in \( i \), the next link to process is drawn at random and the algorithm continues with step 1.

This assignment leads to a simulated travel time to work distribution \( p(\tau) \) which can be compared with the measured one. By testing different \( q(\cdot) \) as is shown in Figure 3, it can be seen that the simplest idea works reasonably well: the usage of \( q(\tau) = const \) (flat attraction in Figure 3) already gives a very good approximation to the measured travel-time distribution. The reason why this simple approach works so well can be traced back to the general shape of the distribution function of all distances described in equation 2. The function in

\[
p(x) \propto x^\gamma \exp\left( -\frac{1}{2\sigma^2} \left( \frac{x-\mu}{\sigma} \right)^2 \right)
\]
equation 2 is very similar to the travel time distribution, and this similarity carries over to the travel time to
work distribution.

Since the results shown here are based on the empty network, it can be expected that after correcting for a
loaded network, the simulated $q(t)$ would fit even better. However, the detour of formulating the model with
an arbitrary $q(t)$ (and testing different assignments) now gives some confidence that it is sufficient to use $q(t) = \text{const.}$

3.3. Re-Scheduling

Assuming that a reasonable trip for any traveler has been generated, the trips are fed into the router, which
assigns a shortest route to any of the trips. These shortest routes are computed by using the simulated time-
dependent travel times. (The first iteration uses the empty net travel times, of course.) The routes are now in-
put into a micro-simulation that executes these routes in the network, giving “real” travel-times for all the
trips. These real travel times are then used to re-route a certain part of the travelers, which gives a new set of
routes which are executed, until this process has converged. Then, the destination assignment has to be done
again based on the loaded network, therefore two types of iterations have to be performed. Figure 4 gives a
graphical account on this procedure.

The dynamic traffic assignment has been done by using a simple micro-simulation to compute the travel times.
It describes the main effects relevant for the determination of travel times in a dynamically driven network: (i)
the spill-back phenomenon of a loaded link, and (ii) the FIFO-condition (First-In, First-Out, meaning that a car
which enters a link first leaves this link first), and (iii) that the travel times that depend on the load of a link.
Any link in the network is modeled as a waiting queue (10,11) with an exit flow constraint and the constraint
that cars are allowed to enter the next link only if there is enough free space. This kind of simulation gives an
approximation to the traffic load distribution in the network and has the advantage that it is numerically very
fast. Running the queue-model for the morning period, lasting from 4 a.m. to 1 p.m. with 520,000 cars needs
25 min on a SUN Ultra 250 MHz.
4. SIMULATION RESULTS

After performing a certain number of iterations, the simulation was stopped and travel times for the initial empty network and the final loaded network were computed by a Dijkstra algorithm for time-dependent link-costs. These two travel times, available for each traveler, were used for re-assigning a number of employees. For details on the Dijkstra algorithm used, see for instance (13); for numerical performance details, see (14). Only the travelers whose travel-time in the loaded network exceeds the travel-time in the empty-network by a certain amount \( \theta \) were subject to re-assignment. The difference between the two travel times was measured by computing \( \eta = (t_{\text{full}} - t_{\text{empty}})/(t_{\text{full}} + t_{\text{empty}}) \) which maps the difference in travel times to the interval \([-1,1]\). The re-assignment algorithm is very simple: the working-place of any traveler who matches the condition \( \eta > \theta \) above was given back to the pool of working places, and then the procedure described above that assigns the employees to the working places was repeated for this subset of all travelers. This yields a new set which was then fed into the iteration that determines the dynamic equilibrium. The whole process has to be repeated until all travelers are satisfied with their choice, or until there are no more changes between subsequent iteration steps.

After the first route-assignment, re-assignment of the destinations was done with \( \theta = 0.5 \), leading to 150,000 out of 520,000 new destinations. Then the route-assignment was restarted, again with an initially empty network and the corresponding travel times. After the second route-assignment, with \( \theta = 0.4 \), roughly 100,000 travelers got new destinations. After the third outer iteration loop was completed, again roughly 20\% of all travelers have travel times that are larger than 40\% of their expected travel times, so the iteration has met a natural (albeit not very satisfactory) termination criteria.

The results of the simulation are summarized as follows. Table 1 shows the traffic flows between Vancouver (Washington, W) and Portland, Oregon (O), compared to the data obtained from the census data. The assignment has an average error of about 15\%. The simulation results basically assign too much traffic between O and W, and they only partially catch the asymmetry between O and W. Different realizations (different random number seeds) of this stochastic assignment algorithm lead to similar results, where the statistical errors are of the order of what has to be expected naively (i.e. any number in this table has an error of the order of the square-root of this number). Therefore, the algorithm is robust against small changes in the parameters. In
Figure 5 a more detailed comparison between the measured OD-matrices (between the Portland PUMA’s) is shown. In order to facilitate the comparison, the difference between the empirical OD-matrix and the simulated ones is shown, both for the initial destination assignment as well as for the third re-assignment. It seems, that the re-assignment did not change much. In Figure 6, the measured travel times are compared to the simulated ones. Nevertheless, in both cases, reasonable agreement between the simulation data and the empirical data can be seen, however the agreement is not completely satisfactory. This shows, that the proposed algorithm can be used to compute the location choice of a large number of individuals. Additional research is needed in order to understand the interplay between destination and route choice better.

Note finally that the algorithm that computes the destinations runs very fast even when using the exact distances in the network as computed by Dijkstra. The assignment of 600,000 individuals took about 20 minutes. When using the flat travel time function the main time is spent with reading and writing the results from/to disk.

5. SUMMARY

It has been shown that very simple approaches lead to reasonable dynamical distributions of traffic flow in a city, using a combined microscopic destination and route assignment algorithm. This article has discussed only the problem of assigning the employees to their respective working places, which is a small part of the trip distribution problem. Further work into this direction is currently being performed at TRANSIMS, in collaboration with Marc Bradley.

The results obtained up to now by such an approach compare well with empirical data as far as they are available. Even the simplest approach seems to work well, provided the underlying data are as detailed as in the case of Portland. Unsatisfactory is the somewhat hand-waving way how the criterion for re-assignment of travelers is done; it has to await future work to find more reasonable criteria. It may be necessary to understand the interactions between the destination and the route assignment better. What remains to be done is to compare the results of the micro-simulation to real traffic flow data. However, this leads to a sensible comparison only together with a full set of activities as mentioned above.
ACKNOWLEDGEMENTS

It is a pleasure for one of us (PW) to thank LANL for the kind hospitality during a stay in Los Alamos, during which part of this work was conducted. Also thanks to the FVU–people for numerous discussions about the topic presented here. Special thanks to Guido Rindsfusser from the RWTH Aachen, who has helped a lot in understanding the classical approach on this theme.

REFERENCES

8. TRANSIMS–Project, see http://studguppy.tsasa.lanl.gov/


FIGURE CAPTIONS

Figure 1: Spatial decomposition of Portland into PUMA's. The numbers in the legend are the PUMA-ids.

Figure 2: Empirically determined travel time distribution, together with a smooth approximation obtained by the procedure described in the text; note the peaks at multiple integer of $5 \ [min]$, due to the tendency of humans to report simple numbers only.

Figure 3: Resulting travel time function from different assignment cost functions, compared to the empirical function (fit to data). The $1/cij$-attraction function results from an destination assignment with the $t^{-1}$-function, flat attraction is the function without distance dependence mentioned in the text.

Figure 4: Flow diagram of the simulations used to compute the combined destination and route choice assignment. The inner loop is the route assignment loop, the outer loop computes the destinations based on the results of the route assignment.

Figure 5: Comparison of the results of the simulation with the empirical data. The left plot shows a comparison between empirical data and the simulation for the OD-streams between the Portland PUMA's. Note, that in this case the re-assignment of the destinations didn't improve the original assignment.

Figure 6: This plot shows the evolution of the travel time distribution during the assignment loops. Shown here is the fit to the empirical travel time distribution, the initial assignment based on the empty net, and three intermediate results during the iterations.

TABLE CAPTIONS

Table 1: Comparison between simulation and data on the basis of an aggregated OD-matrix. The flows between Vancouver (W) and Portland (O) are shown.
Figure 3

Figure 4

1. Destination (trips)
2. Routing depending on travel times
3. Routes
4. Microsimulation
5. Convergence?
6. Travel times
7. Assignment of destinations
8. No
9. Yes
TABLES

Table 1:

<table>
<thead>
<tr>
<th>O/D-pair</th>
<th>Census data</th>
<th>Assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>O → O</td>
<td>423,362</td>
<td>393,942</td>
</tr>
<tr>
<td>O → W</td>
<td>8,282</td>
<td>37,702</td>
</tr>
<tr>
<td>W → O</td>
<td>29,732</td>
<td>51,864</td>
</tr>
<tr>
<td>W → W</td>
<td>48,971</td>
<td>26,839</td>
</tr>
<tr>
<td>Total trips</td>
<td>510,347</td>
<td>510,347</td>
</tr>
</tbody>
</table>
