Iteration-free microassignment

Michael Wegener


1. Introduction

Integrated models of urban land-use and transport capture the two-way interaction between location and mobility decisions of households and firms over time. Because of the slowness by which the physical stock of cities, such as residences and commercial and industrial buildings, change, these models typically cover a twenty- or thirty-year period. To implement feedback between land use and transport, they have to run their land-use parts and their transport parts once in each simulation period.

This puts high demands on the speed by which the transport models embedded in land-use transport models are executed. Execution times of several hours, which may be acceptable if the transport model is applied only once, are prohibitive if it is to be executed once very year in a thirty-year simulation. This constraint is in conflict with the current tendency to make urban travel models more disaggregate or even entirely microscopic down to the individual traveller, which typically leads to even longer execution times even with fast parallel computers. A significant part of the computing time requirements of highly disaggregate transport models is due to the large number of iterations required to achieve user-optimal equilibrium in trip assignment.

One way out of this dilemma is to review the rationale underlying these iterations. Obviously, reality does not iterate but produces a consistent sequence of trip patterns over the twenty-four hours of each day without trials. Why is it not possible to follow reality and produce consistent travel flows without iteration?

This paper outlines a methodology to model activity patterns, trips and trip chains, destination, mode and route choice of individual travellers in urban regions by time of day, including within-day and period-to-period adjustment of behaviour, by microsimulation without iteration. The presentation is illustrated by a first simulation experiment using the urban region of Dortmund as a study region.

2. Problem statement

Mathematical models for forecasting urban travel flows originated in the 1950s in the United States pioneered in the Chicago Area Transportation Study (CATS). The paradigmatic urban travel model consisted of four steps: (i) In the trip generation step the volume of trips originating in each travel analysis zone was estimated from socio-economic zonal data using statistically derived trip rates. (ii) In the trip distribution step these trips were allocated to possible trip destinations as a function of socio-economic characteristics of destination zones, or trip attractions, and the travel times or generalized cost between them. (iii) In the modal split step these origin-destination flows were allocated to available travel modes as a function of the relative attractiveness of these modes, mostly expressed by their travel-time ratio. (iv) In the trip assignment step these model flows were assigned to the links of the modal networks.
For the trip distribution step, the gravity model was used as the first spatial interaction (or in short SIA) model. Its straightforward physical analogy has later been replaced by better founded formulations derived from statistical mechanics (Wilson 1967) or information theory (Snickars and Weibull 1976), yet even after these substitutions the SIA model did not provide an explanation for the spatial behaviour modelled. Only later it became possible (Anas 1983) to link it via random utility theory (Domencich and McFadden 1975) to psychological models of human decision behaviour (Luce 1959).

It was soon becoming apparent that it was not sufficient to apply the four steps of the paradigmatic model sequentially. Depending on the flows assigned to the road network in the trip assignment step, travel times on congested links increased and became inconsistent with those used in the trip distribution and modal split steps. This inconsistency led to the definition of user-optimal network equilibrium (Beckmann et al. 1956), a state in which the pattern of flows in the network reflects the generalized costs on its links, which is equivalent to Wardrop's (1952) condition that each used route between each origin and each destination has the same generalized travel cost and no unused route a lower cost.

There exist essentially three methods to achieve user-optimal network equilibrium through multiple iteration of trip distribution, modal split and trip assignment and averaging after each iteration (Boyce et al. 1994): the method of successive averages (MSA) over multiple all-or-nothing assignments, user-optimal assignment using Frank-Wolfe linearisation and all-or-nothing assignment using partial linearisation following Evans (1976). In all three methods travel times or generalized travel costs of each link are adjusted using speed-flow relationships (capacity restraint). The weights used for each iteration in the averaging are chosen to be the best in each iteration or pre-determined as $1/n$ in the $n$th iteration (Powell and Sheffi 1982). The iterations are largely responsible for the generally long computing times of state-of-the-art travel forecasting models. The problem prevails despite recent advances in assignment algorithms, such as the origin-based assignment by Bar-Gera (1999).

Things are getting worse with the current tendency to make urban travel models more disaggregate or even entirely microscopic down to the individual traveller in order to models multipurpose unimodal and intermodal trip chains and time of day of trips, the interaction between activity and mobility patterns of household members, new lifestyles and work patterns, such as part-time work, telework and teleshopping, the interaction between travel demand, car ownership and residential and firm location, and environmental impacts of transport such as traffic noise and exposure to air pollution. Disaggregate travel models aim at a one-to-one reproduction of spatial behaviour by which individuals choose between mobility options in their pursuit of activities during a day (Axhausen and Gärling 1992; Ben-Akiva et al. 1996). Activity-based travel models start from interdependent 'activity programmes' of household members of a 'synthetic population' (Beckman et al. 1995) and translate these into home-based 'tours' consisting of one or more trips. Activity-based travel models do not model peak-hour or all-day travel but disaggregate travel behaviour by time of day, which permits the modelling of choice of departure time. There are also disaggregate traffic assignment models based on queueing or cellular automata approaches, e.g. in the TRANSIMS project (Barrett et al. 1999; Nagel et al. 1999), which reproduce the movement of vehicles in the road network with a level of detail not known before.

However, microscopic disaggregation typically leads to even longer execution times even with fast parallel computers. As with aggregate models, a significant part of the computing time requirements of highly disaggregate transport models is due to the large number of iterations required to achieve user-optimal equilibrium in trip assignment. There are approaches to model within-day and day-to-day adjustment of behaviour by modelling dynamic network equilibrium (e.g. Bernstein and Friesz 1998; Nagurney and Zhang 1998). These
highly sophisticated methods, however, suffer from even larger computing time problems through iteration. Long computing times have also been a serious problem for TRANSIMS.

These long computing times become even more of a problem if the travel forecasting model is combined with an integrated urban land-use transport model. These models capture the two-way interaction between location and mobility decisions of households and firms over time. Because of the slowness by which the physical stock of cities, such as residences and commercial and industrial buildings, change, they typically cover a twenty- or thirty-year period. To implement feedback between land use and transport, they have to run their land-use parts and their transport parts once in each simulation period. Moreover, these models are increasingly becoming more disaggregate to deal with aspects of urban form, travel demand management and environmental impacts (Spiekermann and Wegener 2002). This puts high demands on the speed by which the transport models are executed. Execution times of several hours, which may be acceptable if the transport model is applied only once, are prohibitive if it is to be executed once very year in a thirty-year simulation.

One solution to this problem would be to develop a travel forecasting model that does not require iteration. Obviously, reality does not iterate but produces a consistent sequence of trip patterns over the twenty-four hours of each day without trials. Why is it not possible to follow reality and produce consistent travel flows without iteration?

A first step towards this goal is to review the rationale behind the concept of user-optimal network equilibrium. There are good reasons to put doubts on the half a century old proposition that user equilibrium is the best representation of travel behaviour. After all, the basic assumption underlying user equilibrium, complete rationality and complete information of all travellers, is highly unrealistic. Instead, many travellers often find themselves trapped in no-return situations, such as traffic jams, wrong lanes, no-turn intersections, train delays or missed connections they would have avoided if they had had prior complete and timely information. Modelling travel behaviour then becomes the art of modelling decision making under uncertainty with incomplete information, short-term adjustment and trial-and error with a significant proportion of routine and habitual behaviour. However, it can be assumed that travellers apply knowledge from previous experience. This can be exploited in a modelling environment in which the transport model is applied recursively in each simulation period.

3. Model framework

The urban travel model envisaged is part of long-term effort to develop a microsimulation model of urban land use, transport and environment (Wegener and Spiekermann 1996; Salomon et al. 2002; Moeckel et al. 2002) based on the existing land-use transport model of the urban region of Dortmund (Wegener 1998). Parts of the planned model are presently being implemented in the project ILUMASS (Integrated Land-Use Modelling and Transportation System Simulation) funded by the German Federal Ministry of Education and Research. The study region for tests and first applications of the model is the urban region of Dortmund.

The model consists of a number of microsimulation modules. A microsimulation module is a programme unit that executes one elementary process (a choice, a transition or a policy) and stores the result in the common micro database. Each microsimulation module has defined input and output interfaces. Co-ordination between the modules is facilitated by a co-ordinator or scheduler programme. The rows and columns of Figure 1 represent microsimulation modules ordered by increasing speed of change:

Transport infrastructure and buildings represent the slowest kind of change; their construction takes many years, and their lifecycle is counted in decades. Firms and households have also lifecycles of several years but are more easily established or dissolved. Firms and
households change their location several times during their life yet even more frequently adjust their vehicle fleets to changing needs. Whereas all these changes are counted in years, logistics and household activities change from hour to hour during a single day. The fastest urban processes are goods transport and travel. They adjust in response to events in a matter of minutes. Environmental processes partly reflect the effects of human activities without delay but some have long-term consequences.

The microsimulation modules interact in various ways with each other. Figure 1 shows the direct interactions between microsimulation modules represented in the model.

![Figure 1. Interactions between microsimulation modules](image)

Some of these interactions are highly delayed, i.e. take their time to work their way through the system. For instance, increasing demand for office space or housing will result in new office space or new housing only after several years because of long planning and construction periods. Other impacts are much faster. For instance, dwellings vacated by households enter the supply of available housing after a few weeks. Still other impacts are almost immediate, such as driver response to congestion. This variety of response speeds requires that the exchange of information between the microsimulation modules is very efficient. This is achieved by the common micro database.
3.1 The travel microsimulation module

The travel microsimulation module presented here is an attempt to combine several partially conflicting objectives:

- to model mobility decisions in a microscopic perspective in order to capture aspects of behaviour that are that are crucial for achieving sustainable urban transport, such as multipurpose unimodal and intermodal trip chains and time of day of trips, the interaction between activity and mobility patterns of household members, new lifestyles and work patterns, such as part-time work, telework and teleshopping, the interaction between travel demand, car ownership and residential and firm location, and environmental impacts of transport such as traffic noise and exposure to air pollution;

- to take into account that travellers have different travel preferences and perceptions of the transport system based on incomplete information about its current state, that they are uncertain about unexpected events, such as accidents, that they frequently find themselves trapped in situations they would have avoided if they had had prior information and that they base their travel decisions on previous experience that my be outdated or on habits and routines that are insensitive to current information;

- to reproduce the dispersion of mode and route choice that result from that diversity of preferences, incomplete information, uncertainty, trap situations and habits,

- to model both short-term adjustment, such as change of departure time or en-route change of destination, mode or route as well long-term learning based on prior experience;

- to develop efficient algorithms for activity generation, journey and trip generation, destination, mode and route choice and assignment that does not require extensive iterations.

To achieve these objectives, the stochastic microsimulation already proposed by Burrell in 1968 is applied. However, unlike the procedure proposed by Burrell, congestion is taken into account by using generalized link travel costs based on the network link flows of the previous simulation period. The travel microsimulation module models for each member of each household the selection of an activity programme and, following that, a departure time for each tour and a departure time, destination, mode and route for each trip (see Figure 2):

- Select household. In the first step a household is selected for processing from the list of households. Each selected household is defined by its household attributes and the personal attributes of its members. The household attributes include its residential location. A location in the model is a micro location, i.e. street address, geographical co-ordinates or raster cell of 100 x 100 m size.

- Select person. Next the first household member is selected. For each working person in the household the location of the workplace is known. For school children and university students the location of the school or university is known.

- Select activity programme. Depending on the personal attributes of the household member, i.e. age, sex and occupation, a daily activity pattern is selected from a catalogue of activity patterns. A daily activity pattern is defined as a schedule of tours.

- Select car ownership and availability. Depending on household and personal attributes it is determined whether the person has a car at his or her disposal.

- Select tour departure time. The first tour of the activity programme is selected. The departure time is determined as a random variation of the scheduled departure time.
Figure 2. Microsimulation of travel behaviour
- **Select trip departure time.** The first trip of the tour is selected. The departure time is determined as a random variation of the scheduled departure time.

- **Select destination.** The destination of the trip is selected by logit choice. The locations of destinations are micro locations as above. Generalized costs of travel to the destinations are calculated as the logsum of stochastic shortest routes (see below) of relevant modes. Relevant modes are walk, cycling, public transport and car (if available, see above). For work, school and university trips the destinations are already known.

- **Select mode.** For the selected destination, mode choice is performed by logit choice based on the generalized costs of stochastic shortest routes (see below).

- **Select route.** For the selected mode the stochastic shortest route is selected as route. Stochastic shortest route is the shortest route with a random disturbance term added to each link generalized cost and each waiting/transfer time in the public transport network.

- **Move person through network.** Each person travelling through the network is recorded on each traversed link by 10-minute time interval.

After each trip the next trip of the route, if any, is selected. After each route, the next route, if any, is selected. After each person, the next person, if any, is selected. After each household, the next household, if any, is selected.

There are two ways of selecting the next trip. One intuitively appealing way is to start in the early morning hours with an empty network, process trips in the order of their departure time, that is spatially randomly, and after each trip update the generalized travel cost of all traversed links. In this way the gradual filling up of the network over the day is reproduced. This would be a microscopic version of the incremental loading assignment in use prior to the development of user-equilibrium assignment algorithms. If, however, it is assumed that travellers use their prior experience about network conditions when making travel decisions, an even simpler procedure can be applied. In this case the assignment does not start with free-flow generalized link costs but with the higher link travel costs of the loaded network of the previous simulation period. It is then not necessary to process trips in the order of their departure time. Only after all tours and trips have been executed, the travel times of all traversed road links in each ten-minute interval of the day are updated to account for congestion; however, this information will be used only in the next simulation period. Representative travel times and generalized costs between zones (required for accessibility calculations in the land use model) are calculated on the basis of shortest routes with updated travel times.

If during a trip a significant amount of congestion is encountered, short-term adjustment resulting in a postponement of the trip or a change of mode or route may occur. However, only changes of departure time and mode that can be made *en route* are implemented. Long-term adjustment of travel behaviour, such as going to work later or buying a monthly public transport pass, are based on the generalized costs of the network in the previous simulation period. Generalized costs are a combination of travel time and travel cost and can be different for each type of traveller to take account of the diversity of travel preferences.

Special provisions are necessary when no prior information is available as in the first simulation period or in the case of large infrastructure changes. In the first simulation period either one aggregate user-equilibrium assignment using the Evans algorithm or one microassignment iteration starting from medium-flow generalized link travel costs may precede the actual assignment. Similarly, large infrastructure improvements, such as new road links, may be introduced with medium-flow generalized link travel costs representing the most likely expectation of travellers.

It is hoped that microassignment without iteration will produce a similar distribution of trips across destinations, modes and routes as user-optimal assignment with iteration. Total
user benefit should be less due to the effects of uncertainty, incomplete information, trial and error and habitual behaviour. It will be an interesting task to examine the degree of sub-optimality and how the simulated travel behaviour compares to observed behaviour and the results of travel models based on user-optimal network equilibrium.

### 3.2 A first test

As a first test of the proposed method, iteration-free microassignment was applied to peak-hour car trips in the Dortmund urban region and compared with the results of an user-optimal equilibrium assignment of the aggregate transport component of the existing land-use transport model of the Dortmund urban region.

The existing transport model applies the Evans algorithm to arrive at user equilibrium of trip generation, car ownership, trip distribution, modal split (walk-cycle, public transport, car) and route choice (Wegener 1986). Normally eight iterations are performed, but for this exercise the number of iterations was increased to 20. A simplified version of the model with 36 zones and 1800 network links was used; in the final version 200 000 microlocations (grid cells) and 8000 network links will be used.

The resulting origin-destination matrix of peak-hour interzonal car trips was then assigned to the links of the road network on a car-by-car basis. For each car trip, a stochastic shortest route to its destination was determined using a shortest-route algorithm. For this the generalized link travel costs of the aggregate user-optimal equilibrium assignment produced by the Evans algorithm described above were used – in the final model the travel costs of the loaded network of the previous simulation period will be used.

During the shortest-route search, these generalized link travel costs were disturbed by an uniformly disturbed random increment of decrement of up to ten per cent. A 'once-through' shortest-route algorithm, in which the nodes already reached but not further processed (the 'candidates') are temporarily preserved in the 'candidate list' in the order of their travel cost from the origin node (and are hence processed only once), made sure that each link cost was disturbed only once. Only route changes, no other behavioural responses, such as change of departure time or change of mode, were yet implemented nor were time intervals or microlocations considered – this will be left to future experiments.

In Figure 3 the resulting link flows of the microassignment are compared with the link flows generated by the aggregate user-equilibrium assignment with the Evans algorithm. It can be seen that the majority of link flows produced by the two methods are very similar, with a few significant deviations that need further investigation by comparison with observed link flows. It should be noted that a perfect fit of the two link flow distributions cannot be expected – and is not even desirable if the hypothesis holds that user equilibrium is not the best approximation of actual travel behaviour.

### 4. Conclusions

This paper outlined a methodology to assign individual trips generated in a microscopic activity-based travel forecasting model to a multimodal transport network without iteration. The rationale of the method rests on the assumption that travellers have only limited information about the current state of the network and that they base their travel decisions on prior knowledge from earlier experience. This assumption challenges the common assertion that user-optimal equilibrium represents the best approximation of travel behaviour.
The iteration-free nature of the approach makes it particularly suitable for integrated models of urban land use, transport and environment (LTE), which, as it was indicated, are increasingly becoming more disaggregate. Ideally, the transport component of a microscopic LTE model should also be microscopic but this conflicts with the need to have a very fast transport model to implement feedback between transport and land use. The proposed iteration-free microassignment method promises to be a solution to this conflict.

Nevertheless, before this becomes reality, several problems have to be solved. At a conceptual level, the question to what degree transport networks are in equilibrium needs to be investigated empirically. This will require new approaches of analysing travel choice behaviour from a cognitive-science perspective. In addition, a number of technical problems have to be addressed. Even iteration-free assignment is too slow unless efficient methods to calculate individual shortest routes between one origin and one destination (not trees) are developed. Also the short-term adjustment conceptualized in the algorithm will have to be implemented in a theoretically sound and at the same time efficient manner. Finally, new methods of calibrating and validating the model against observed travel data will need to be developed.

Acknowledgements

The author is grateful to David Boyce for a highly useful discussion on the concepts underlying this paper which led to a significant change of the original idea — bringing to mind a similarly important suggestion made by him when the author struggled with his first imple-
mentation of the Evans algorithm (Wegener 1986). The author also greatly benefited from several enlightening discussions with Britton Harris, who has developed innovative ideas about microassignment, although eventually an independent strategy was followed. Helpful comments by two anonymous referees are gratefully acknowledged.

References


Axhausen, Kay W. and Tommy Gärling (1992), 'Activity-based approaches to travel analysis: conceptual frameworks, models and research problems', Transport Reviews, 12, 324-341.


Ben-Akiva, Moshe E., John L. Bowman and Dinesh Gopinath (1996), 'Travel demand model system for the information era', Transportation, 23, 241-166.


Boyce, David E., Yanping Zhang and Mary R. Lupa (1994), 'Introducing "feedback" into four-step travel forecasting procedure vs. the equilibrium solution of a combined model', Transportation Research Record, 1443, 65-74.


Evans, Suzanne P. (1976), 'Derivation and analysis of some models for combining trip distribution and assignment', Transportation Research, 10, 37-57.


Moeckel, Rolf, Carsten Schürmann, Klaus Spiekermann and Michael Wegener (2002), 'Microsimulation of urban land use', Paper presented at the 42nd Congress of the European Regional Science Association in Dortmund.


