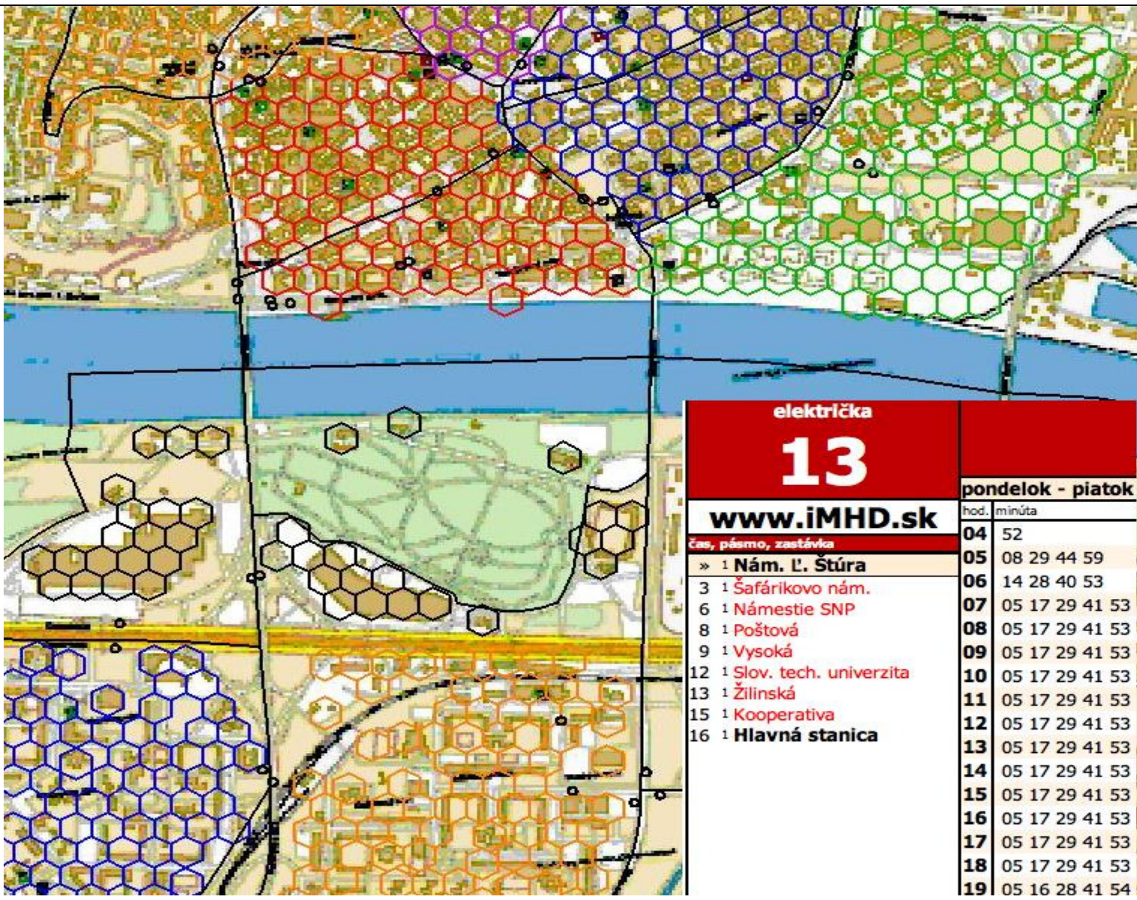


Influence of Spatial Preciseness and Stop Choice Behaviour on Total Travel Time Values in Urban Public Transport Systems

Peter Ľos, TRANSPORTNET / TU Delft

Conference paper STRC 2008

STRC

8th Swiss Transport Research Conference
Monte Verità / Ascona, October 15-17, 2008

Title of paper

Peter Ľos
TRANSPORTNET / TU Delft
Delft

Phone: +31-15-278 9397
Fax:
email: los.peter@gmail.com

October 2008

Abstract

Modelling of public transport in urban environments is quite problematic. It is required to deal with many difficulties by calculation of travel time properties due to multiple choice sets of stops, lines and routes available. Also, users' stop choice preferences are to be considered, which may even differ by trip circumstances. And indeed, contrary to private modes, time-table dependency and network captivity must be tackled as well.

Practical models usually calculate total travel times between rather large pre-defined zones and their affiliation to the network through virtual connectors is also problematic. Component times, especially walk times to and from stops are mostly only estimated, since whole area (often not completely spatially consistent) is represented by a single centroid. In most cases this is sufficient; nevertheless, because of spatial detail deficiency there is always some error, which may not be sometimes negligible. Moreover, traditional approaches hardly detect minor variance in total travel times based on differentiated stop choice types.

Two research questions are to be clarified in this paper: (i) to what extent are computed total travel time values dependent on level of spatial detail?; and (ii) how does different stop choice behaviour affect average total travel time values? A special modelling technique and calculation procedure must have been developed to analyse these issues in a real system. It is based on precise calculation of total travel time on all feasible routes simultaneously. Six types of stop choice behaviour are applied, simulated minute-by-minute and averaged over whole peak and off-peak hour. Spatial preciseness for calculation of access and egress times is altering from one-minute to zone-size cells in couple of levels and is again averaged over whole zones for comparison.

Keywords

urban public transport networks modelling – travel time computation – stop choice behaviour

1. Introduction

Determination of the total travel time and shares of its components is one of the most important tasks in the transportation network and accessibility modelling. Factual and (by travellers) perceived travel time, which is necessary to reach the destinations, stands for a main accessibility indicator. It is also the essential utility attribute that is encompassed in three of four stages of classical modelling methods (trip distribution, modal choice and network assignment). Hence definitely, it must critically influence the outcomes of the models, and thus their validity and explanatory value.

A model is a virtual mathematical representation that attempts to describe the features of real objects and their interactions within the system. In more complex systems, and transport is indeed suchlike, this is only possible through various simplifications and reductions of intricacy. In case of transport network and accessibility models are these simplifications related especially with (i) the spatial preciseness of the study area and network representation, and (ii) travel behaviour of users, which is indeed not completely predictable. Both these factors result in inaccuracy of the travel time calculations.

Space-time continuum is for purposes of transport modelling normally discretised in a limited number of elementary units in both dimensions. Spatial units are usually called transport analysis zones (TAZ), or shortly just zones. A lot of various methods for zoning design can be found in literature (Openshaw and Taylor, 1981; Xie, 1995; Ortuzar & Willumsen, 2001; Duckham et al., 2001; Eagleson et al., 2002 and many others).

The results obtained from spatial data are not independent of the scale and the delineation of TAZ directly impacts the modelling results (Viegas et al., 2007). This phenomenon is called Modifiable area unit problem (MAUP) and has been also already studied from various perspectives (Openshaw, 1977; Openshaw & Rao, 1995, Alvanides, 2000). Many other authors investigated impacts of MAUP on (real) transport planning models. (Chou, 1991) and (Zhang & Kukadia, 2005) assessed the spatial aggregation effects, confirming the presence of MAUP. (Ding, 1998) has found that scale effects play an important role on quality, validity and reliability of model results. (Chang et al., 2001) confirmed that higher level of detail of the networks decreases the modelling error. (Viegas et al., 2007) concluded that zoning (grid cell) size has significant effects on some important modelling variables and indicators, such as percentage of intra-zonal trips and number of cells with not statistically significant O-D flows.

Surely, the level of necessary spatial preciseness is dependent on the purposes of the particular study and required preciseness of results. However, the number and boundaries of the zones are most often modifiable and are thus set by modellers just arbitrarily. Since the computation power has incredibly increased over last years, the number of TAZ is now in fact only limited through feasible amount of data that can be collected, for instance to build O-D matrices. The amount of data usually rises exponentially with number of zones. Other restrictions come out e.g. from administrative division that often need to be somehow adopted. The number of zones in practical models ranges from tens (e.g. Bratislava – 420 000 inhabitants / 43 zones in General Municipality Masterplan) to several thousands (e.g. Rotterdam – cca 500 000 inhabitants / over 1500 zones in the model by Goudappel Coffeng).

But the calculation of travel times is not a trivial task; and especially this is true for urban public transport systems, because of several multiple choice problems, briefly discussed in chapter 2 of this paper. The intensity of these issues is increasing with the size of the zone. Apparently, different parts of larger zones may be served by public transport differently; hence, the calculation of average travel times is a bit doubtful when common modelling techniques are applied.

Nevertheless, travel times that users of public transport systems do actually undergo, are clearly also dependent on their route and stop choice behaviour. The actual choice is then indeed affected by different utility perception of various groups of users and level of information (that they have available or are aware) about the system performance.

The general objective could be the minimisation of total travel time, factually irrespective of other factors (because they are mostly already incorporated in this condition). This would necessitate complete changing of the route alignment in every moment, however. So usually, majority of people (familiar with the system) simply go to one particular stop (or "knot" as explained later), providing the highest probability of a good connection. Such knots are usually those, which are served by most frequent and fastest services, and laying on routes with low number of transfers; hence fulfilling other choice characteristics.

Nonetheless, some specific groups (e.g. elderly, handicapped or users with temporary mobility constraints) would probably prefer the stops, which are the closest as possible to both, origin and final destination. Furthermore, daily commuters would often choose other routes and connections than visitors unfamiliar with the system; they may opt to walk to another stop, from which the lines are more frequent and/or reliable, or wait few minutes longer to take a speed line instead of first arrived.

In the first point of analysis in this paper, the discussion on MAUP effects is contributed with the analysis of the influence of the spatial preciseness on results of travel time computations in a (part of a) real public transport system in the city of Bratislava. In the second point it is shown, how are the total travel times and shares of its components affected by different types of route and stop choice behaviour of passengers.

2. Calculation of travel time in urban public transport

If the routes and their itineraries are determined, the total travel time by public transport T^{PT} can be calculated by a well-known formula as a sum of access t_a and egress t_e, waiting t_w in-vehicle journey t_j and all transfer t_t times (if applicable).

$$T^{PT} = c_1.t_a + c_2.t_w + (\Sigma c_3.t_j + \Sigma c_4.t_t).C^T + c_5.t_e \ [min]$$

In some models, specific coefficients are used to express different perception of flow of time by separate travel activities (e.g. waiting is perceived longer than in-vehicle time). This is reflected through application of the perception coefficients (c_i). However, there is an additional problem with the transfer penalties: for example, three-minute transfer has certainly bigger deterrence effects by 10-minute then by 30-minute total trip time. This can be considered by a transfer necessity penalty (C^T) that virtually prolongs the total journey time.

Public transport supply varies during the day (peaks and off-peaks) and also between the various days of a week or year (work-days, week-ends, school holidays, etc.). The big advantage of urban systems is that they are quite regular, and thus temporal discretisation into only few different periods is rather easy. Regretfully, urban public transport systems exhibit several other specific features that make proper determination of the travel times more complicated then in inter-urban relations. A special attention should be devoted to this issue.

Most of commercial software packages for (public transport) modelling allow two basic ways of travel times calculation: (i) headway-based and (ii) timetable-based. The former is more popular, because it requires much less detail input data; the average waiting times are just approximated as a half of the headway and even in-vehicle times can be just estimated from the known distance and average vehicle speed (during the particular period). When the latter approach is used, it is possible to compute in-vehicle, transfer and waiting times precisely; however, in larger networks that implies extreme computation time and memory requirements (PTV), as well as incomparably more laborious input and network definitions.

And after all, the access and egress times are in fact just estimated in both cases – through arbitrary placing of the zone centroid and virtual length of the connector(s). In private car models this may not be a big issue, as average speed of cars, even on intrazonal access roads, is at least 5 times higher than average walk speed to/from a stop. But in case of public transport, off-vehicle components (access/egress walk, wait and transfer times) comprise very often significant or even major part of the total travel time. This is because the trips within the city are usually shorter compared to inter-city ones; average total travel times are normally in the order of several minutes. Thus, even a minor inaccuracy or error in the calculation/ estimation of walk and wait times may have severe impacts.

2.1 Multiple choice problems in public transport systems

Therefore, detail and preciseness of the spatial representation of the study area and of the network play an important role. The major problems occur due to multiple choice problems in the dense urban networks, briefly described in following:

A) **Multiple line choice problems:** There are usually several lines with different routing available for any studied O-D pair. Often it is the case that from one particular place in the zone of origin, there is one line appropriate for a trip to the certain place in the destination zone but the other line to go to another place, within the same zone of destination. For a particular trip, this problem has an influence on both, access and egress sides. Although it can be partially overcome with a finer zonal division, this solution would require also more data to work-out precise enough O-D matrix, as well. Traditional estimations and derivations of average access/egress and waiting times are unable to subdue these problems precisely. Although using them is possible, they may substantially bias the results.

B) **Multiple stop choice problems:** To be able to determine (average) access/egress times, we must know which stop would/could be used by what number of passengers. Due to relatively high stop density of stops, there are usually several stops within the reasonable walk distance from any place in a city. These stops are often served by different lines providing different level of service (accessibility, frequencies, transfer necessity) or impose different fare costs. Farther stops may be more appropriate for travelling to particular destinations and the closest one convenient just for few directions. Observation and surveys proved that people are usually willing to walk further at origin than at destination or to get to the faster mode (e.g. train/metro vs. bus). Since urban areas exhibit almost features of fractals, this problem cannot be overcome through finer zonal division of the study area (the problem remains, just the stops are out of the zone then).

C) **Multiple connection choice problems:** Even when the particular stop is known, the problems are not solved yet completely. In some cases there are several reasonable lines, serving the same O-D pair (of stops). An average waiting time of passengers arriving to that stop (going to that destination and knowing about the alternatives) simply cannot be calculated as usual half of headway. Any real-world example would prove higher level of service than estimated in that way. Problem is that departures are rarely harmonized, because of different frequencies. Moreover again, those lines may provide different level of service, so an experienced well-informed passenger may wait few minutes longer, knowing that he will still get to the destination before an uninformed passenger who took the first connection.

The multiple choice problems result in large sets of travel route alternatives for any O-D pair that need to be taken into consideration, both by passengers and planners.

2.2 Route choice

Hence, basic prerequisite of a proper travel time calculation is the generation of route choice sets. As highlighted in (Bovy, 2007), the population of all available routes (so-called universal choice set) is very large and often almost impossible to know. Therefore, at least a basic route choice *master set* need to be generated, from which the actual route choice alternative can be filtered afterwards. Three classes of explicit route set generation procedures can be distinguished: (i) path search (shortest route) based; (ii) constrained enumeration (sometimes called branch and bound); and (iii) probabilistic methods. For an in-depth elaboration of choice set requirements and overview of methods for different modelling applications, see (Hoogendoorn-Lanser, 2005) and (Fiorenzo-Catalano, 2007).

Mostly, route choice in common public transport models (as well as other applications, such as trip advisors) is based on Dijkstra's shortest path algorithm and its enhancements, including various forms of K-shortest path algorithms (Ramming, 2002; Cascetta et al., 2002), labelling-method (Ben-Akiva et al., 1984) and several modifications of Monte Carlo approach (Fiorenzo-Catalano & Van der Zijpp, 2001; Bliemer et al., 2004). Slightly adjusted, these algorithms allow for accommodation of other users' stop and route choice preferences, such as minimum travel time, minimal transfer and minimal walk distance (Wu & Hartley, 2004). These algorithms are fairly simple; however, if they are non-schedule-based then they require post-processing to validate the outcomes.

As argued in (Liu et al., 2001), *"...these fairly efficient and effective algorithms, only work well in unrestricted networks, where free movement is enabled.";* i.e. in case of private car, for instance. But for constrained public transport networks more complicated representations must be involved, which dramatically limit their efficiency. Basically, timetable-based route-search algorithms require enumeration, i.e. determination and subsequent examination of all possible routes, since the "shortest", i.e. most convenient one could be different in every particular moment.

Therefore, constrained enumeration procedures are used to generate exhaustive master sets. In this procedure, a connection tree between origin and destination of a trip based on a branch & bound rule is constructed. The extension of the tree with additional links (until the destination has been reached) is conditional on a number of built-in constraints that reflect cognitive, perceptual, and behavioural requirements. The types of adopted constraints imply that the choice set generation predominantly is preference driven. (Friedrich et al., 2001) developed an application to public transport networks, (Prato & Bekhor, 2006) to road, and (Hoogendoorn-Lanser, 2005) for multi-modal networks.

The disadvantage of the constrained enumeration is often too large choice set generated, and moreover, some attractive alternatives may be deleted. Normally, routes with fewer transfers provide better connection; another method, based on connectivity and transition matrices can be thus derived form the TPlanning algorithm described in (Liu, 2002). The algorithm could search the routes on very precisely coded network, starting from direct connections, subsequently adding routes with more transfers. Similarly, route trees can be also developed unrestrictedly – through a subsequent addition of all feasible destinations of routes serving a particular origin, or transfer point, respectively. These approaches provide exhaustive master set, however, require also a thorough filtering in post-processing phase to eliminate multiply found routes, (differing e.g. only with transfer point on collateral sections of routes) and routes with too long detours.

The final choice of (one or few several) routes is another weak point of more aggregated (non-time-table based) models. The route choice is the point of major simplifications in practical models, and thus source of imperfections, indeed. Because of those large choice sets, many route alternatives may differ only by fractional nominal time values. Every route alternative in public transport systems in fact provides an interval of possible total travel times, which usually overlap. That means that in every particular moment of the decision to start the trip, a completely different route may be the optimal or at least appropriate option.

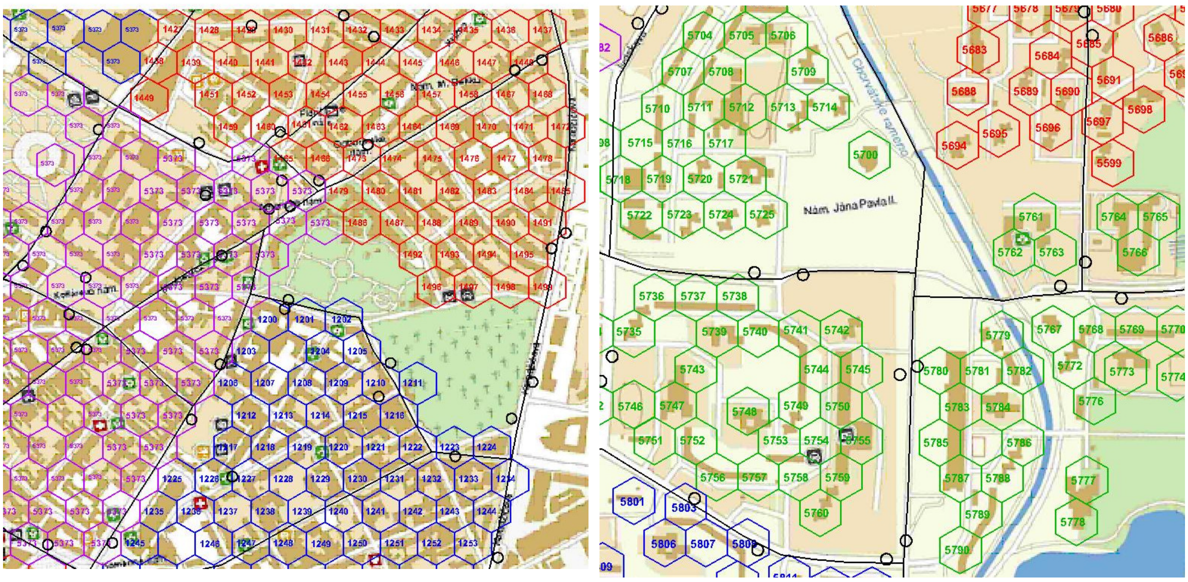
Actual choice behaviour among the routes from the master set is also ambiguous. Firstly, it is dependent on user preferences (as already mentioned, e.g. minimisation of the total travel, walk time and/or transfers). And secondly, it may consist of various forms of choice process: (i) sequential (Marzano & Papola, 2004), where decision is made in every subsequent decision point of the route; (ii) simultaneous, where routes are compared as a whole – from the origin to the destination at once; and (iii) strategic, where choice is made upon prevailing network conditions during the trip. Additional specific aspects are apparent in case of multi-modal networks (see Hoogendoorn-Lanser 2005, and Fiorenzo-Catalano, 2007).

3. Methodology of the analysis

As introduced in the title and first chapter, the objective of the analysis in this paper is to investigate the influence of spatial preciseness and stop choice behaviour on values of travel times in real public transport systems. A part of the real public transport system in the city of Bratislava was selected as a study area. It comprises the largest residential area Petržalka (approximately 150 thousands inhabitants) on the right bank of the river Danube and the central city district on the other bank. For the purposes of the study, the area has been divided into 10 zones of origin (in Petržalka) and 5 destination zones (in the city centre).

The spatial preciseness is changing on several levels; hence the zones are further divided in smaller spatial units, called here “cells”. The smallest is the one-minute walk time cells (OMC – with the radius of cca 75 meters / and extent of ~0,45 ha). Hexagon-shape[1] cells were used to cover entire area. Larger “uninhabited” areas (which are not potential origins or destinations of work- or shop- related trips) were excluded; example is shown in the Figure 1. These cells are further progressively coupled into larger spatial units: two-minute (2-3 OMC / ~1,25 ha) and three-minute (5-7 OMC / ~2,5 ha) cells, thereafter already amorphous “quarter-zones” (comprising roughly 10-12 OMC / ~ 5 ha), “semi-zones” (cca 25 OMC / ~12,5 ha) and “full-zones” (70-100 OMC / > 30 ha), finally.

Figure 1 Example of cell definition – „empty“ areas are omitted



[1] Using the hexagonal shape, walk time difference between two neighbouring cells is the same in every direction.

For each cell, access and egress times are determined to all stops in the area within a reasonable walking distance. Total door-to-door and detail component travel times (access, egress, wait, transfer and in-vehicle times) are then subsequently simulated minute-by minute over characteristic hour (i.e. 60 times) in two day-time periods: (i) morning peak (7-8 a.m.), and (ii) mid-day off-peak (11-12 a.m.). Timetables valid in March 2008 were used.

Six types of user preferences are considered to investigate the influence of different stop choice behaviour on travel time calculations:

- *Optimal choice*: describes the maximal utility that is possible to exploit form the system; (in fact unrealistic) perfect choice behaviour is assumed – the stop of origin as well as route is chosen in order to minimise the total travel time;
- *Minimal walk*: describes a specific behaviour when, both, stops of origin and destination are selected in order to minimise the access and egress times, nonetheless considering the total travel times as well[2]*;*
- *Best Knot*: describes the prevailing behaviour, when the knot of origin is selected, which provides the highest probability of a good connection; the route is chosen afterwards (naturally from a limited choice set already), again to minimise the total travel time;
- *Least transfers*: describes a users' preference for routes requiring less and easier transfers[3], even though total travel time is then slightly higher;
- *First arrival*: describes the behaviour derived from the "*best knot"* when the first arriving connection going to the destination zone is taken, thus minimising first wait times;
- *Best Route:* describes the behaviour of the travellers unfamiliar with the system, who are just aware about one particular route (the route may comprise several service lines that have exactly the same origin, destination and transfer knots, however).

Accomplishment of such a detailed study would be quite problematic in commercial software packages (VISUM, OMNITRANS, etc...). Timetable-based assignment requires a lot of data, which is often input quite inconveniently and laboriously. Practically a separate model would be required to be built for each level of spatial detail. Further, application of several different

[2] Several different knots may be at the similar distance, and the closest one could only be served by not very attractive line(s). Therefore, a subset of knots, access/egress times of which are not more than two minutes higher than the absolute minimum ones, is taken in consideration. Those "minimal walk" knots are then selected out of the subset, which provide the best total travel time averages (over the whole analysed period).

[3] From two alternatives with the same (minimal) number of transfers, the one is selected that fulfils second condition on minimal transfer walk distance, or in the third order only, the minimal average total travel times.

choice behaviour models is complicated, if possible at all. And chiefly, the modelling results are not provided in a convenient form, suitable for further processing. Therefore, to accommodate the issues concerning travel time computation according to presumptions outlined in the previous chapter of this paper, a new software tool has to be developed.

3.1 The ACCEPT Simulator

The tool is named ACCEPT Simulator. It allows a simultaneous simulation of several hundreds of routes and thousands of O-D pairs. Travel times are computed with one minute preciseness according to factual timetables. The network is defined purely through a precise network coding[4]. Each network object attains a unique number to be distinguished from the others. Lines are encoded with five-digit number in the form *L=mllld*; where *m* denotes mode (metro, tram, bus etc...), *l* denotes line number (most often lines in real systems are designated with two or three digits) and *d* discern direction, or eventually sub-line (when route of line is changing).

Three levels of nodal objects are distinguished: stop, knot and the node. Stops are understood as interruption of drive of the vehicle to allow alighting and embarking of passengers. They are labelled with unique code of line and two other digits to attain final form *S=mllldss*; where s is order of stop from start of line. Groups of stops of different lines that share the same spatial coordinates (i.e. physical stops) and served by service lines in the same direction are called knots. Groups of knots located in the same area, allowing transfers in different directions are called nodes, usually having a unique name of street or other important object adjacent (e.g. “Station”). Nodes are coded by three-letter abbreviations of their names, and another digit is attached to denote the respective knots of the node.

Route search algorithm for the ACCEPT Simulator[5] is a composite of the unrestricted enumeration method and slightly modified and adapted TPlanning algorithm described by (Liu, 2002). At first, the algorithm searches all non-transfer routes between selected origin and destination knots (the coding of stops allows for a proper differentiation of directions of the service). However, sometimes, routes that require more transfers can be in fact more convenient. Therefore, the master set generation continues with searching for routes with higher number of transfers. The maximum number of transfers was limited to two (three leg routes) for the purposes of this study.

[4] Disregarding of the graphical network representation allows more convenient input of model data, and also to increase the speed of simulation.

[5] The algorithm is considered as the most suitable one, but still not inherent component of the methodology.

In the end of the route set generation, the journey times of all routes found are preliminary evaluated according to minimal (no wait time) and maximal journey times (wait time equal to headway), in order to exclude those, minimal journey time of which is higher than minimum of maximal journey times set. The results of route-search – routes' itineraries with origin, destination and transfer knots and respective lines – are stored in external file “RouteCube”.

Thus, an exhaustive master set of all routes that may provide minimal travel times (under particular conditions) is created. In the next stage, route travel properties are set according to itineraries. Since practically all lines in Bratislava are regular, each route in each period of analysis is sufficiently described through: (i) travel time, (ii) headway and (iii) first minute of departure for each leg of a route; and (iv) transfer walk times at each transfer node.

Appropriate routes, their itineraries and properties for all periods, are sorted out according to knots in origin and destination zones and are stored in the “RouteInputs” files. Access/egress walk times[6] from each cell (as minimum unit of spatial preciseness) to all knots of the respective zones are stored in another type of input files, called “StopWalkInputs”. Data from these files are then loaded just before the start of the actual simulation, according to selected zone of origin and destination and day-period, analysis of which is intended to be carried out.

In the initial phase of the actual simulation, net journey times (yet excluding access and egress) are computed for all considered routes from their properties, and stored in the so-called “Stop-to-Stop” matrix. Transfer necessity and/or perception coefficients can be considered already. In the next phase, each pair of cells from origin and destination zones is examined separately. Access and egress walk times to respective stops (stored again in external file for each zone) are loaded automatically. The final total travel times are computed for each route and every minute of departure (decision to start the travel) during the analysed period. Consecutively, they are summed and averaged to obtain an indication of the average travel times over the period for each route.

Finally, respective total travel times (identified in the previous step) are averaged again for each stop and route choice type, with all respective component times (access/egress, (first) wait, transfer, and in-vehicle times) as well. Results, absolute and relative (to optimal) values for each choice type are stored in an external “Exports” file, where they are accessible for further analysis, after the simulation is finished for all cell O-D pairs of selected pair of origin and destination zones.

[6] The access and egress times to/form knots can be determined manually or computed automatically in any available transportation analysis or GIS software (that enables pedestrian network analysis).

4. Results of the analysis

The outcomes of simulations (Export files) of every O-D zone pair are post-processed to obtain comparable values for each level of the spatial preciseness. It is assumed that each cell within the zone is the source (or termination) of the same number of trips (total number of trip between the respective zone of origin and destination, divided by the total number of cells within the zone). Thus, the relative weights are assumed to be equal[7] and the overall average (mean value) of a zone (TT_Z) is calculated as a simple quotient of the sum of travel times (total or any component) of all cells (TT_C) divided by total number of cells in the zone (N_C):

$$TT_Z = \Sigma (TT_C) / N_C$$

The resulting figures of TT_Z were compared among each others. The basic statistics of the analysis of the influence of the spatial preciseness on Total Travel Time averages of a Zone (TTT_Z) for all O-D zone pairs and choice behaviour types are summarised in the Table 1.

Table 1 Influence of the spatial preciseness – relative TTT_Z figures (all behaviour types)

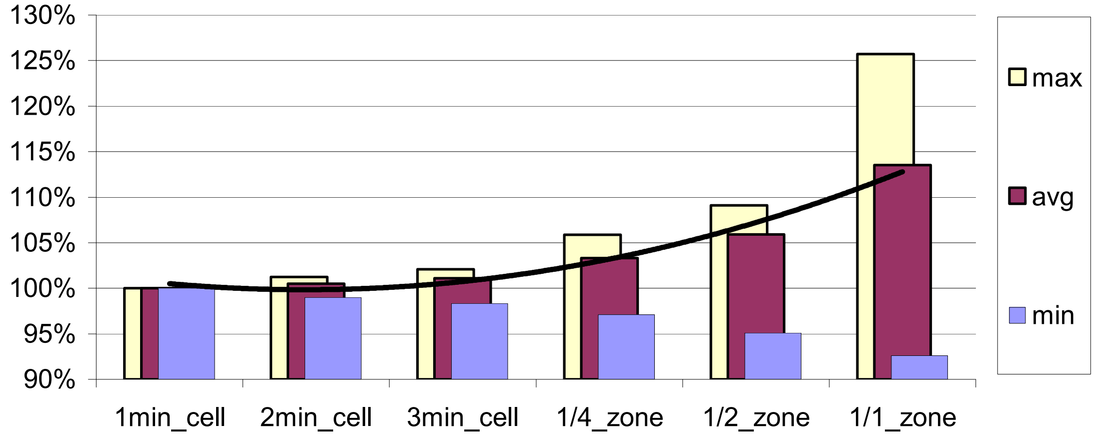
	Morning Peak				Midday Off-peak			
Spatial unit	Min	Mean	Max	St Dev	Min	Mean	Max	St Dev
1-min cell	-	100,0%	-	0,0%	-	100,0%	-	0,0%
2-min cell	99,0%	100,5%	101,2%	0,7%	99,1%	100,7%	101,3%	0,7%
3-min cell	98,3%	101,1%	102,1%	1,1%	98,6%	101,0%	102,7%	0,9%
quarter-zone	96,6%	103,3%	105,9%	2,3%	96,5%	102,4%	106,0%	2,4%
semi-zone	95,1%	105,7%	109,1%	5,1%	94,7%	105,5%	109,4%	5,3%
full-zone	92,6%	113,5%	125,7%	9,9%	91,3%	112,0%	124,9%	10,1%

If it is assumed that travel times calculated for division at one-minute cells are the most “veritable” ones, then it is clearly seen, how is the error increasing with the cell size – up to over 10% on average and maximal deviation even 25% at full-zone level. Interestingly, decreasing level of the spatial detail leads mostly to overestimation of travel times. Although off-peak travel time averages are on average by 13% higher compared to morning peak, impacts of the spatial preciseness are concordant.

[7] Indeed, the exact values are in fact dependent on factual proportion on total number of inhabitants, jobs or retail opportunities. Nevertheless, the zones are defined based on prevailing type of urbanisation and land use; hence, so the assumption is appropriate. But of course, it could be abandoned if the more detail data are available.

From the Figure 2 it is obvious that the error is increasing with the decreasing spatial preciseness exponentially. Yet at 3-min cell size level is the error practically negligible, then both, maximum and minimum deviations starts to grow rapidly.

Figure 2 Influence of the spatial preciseness on total travel times (morning peak)



This finding is also supported with figures in the Table 2. At the 3-min cell level all component travel times are almost the same as at 1-min cell level. The variance of total travel times at lower levels of the spatial detail is primarily caused by the variance in walk and wait times, which have the highest standard deviation (except of transfer, which is not applicable to every route, however). Note that walk and wait times are exactly those component times, which are just estimated in practical modelling techniques, thus are possible sources of errors.

Table 2 Influence of the spatial preciseness on individual component travel time values

Component time	3-min cell spatial preciseness				full-zone spatial preciseness			
	Min	Mean	Max	St Dev	Min	Mean	Max	St Dev
Total	98,8%	100,3%	101,2%	0,7%	99,4%	109,1%	114,4%	4,8%
Walk	97,6%	101,0%	103,4%	1,7%	96,1%	128,8%	155,3%	18,5%
Wait	97,7%	100,9%	105,6%	2,3%	48,1%	77,4%	111,3%	20,7%
Transfer	93,6%	102,7%	111,5%	5,0%	0,0%	67,2%	134,2%	58,2%
In-vehicle	97,5%	99,1%	102,5%	1,14%	75,3%	95,8%	116,6%	11,7%
Total	98,8%	100,3%	101,2%	0,7%	99,4%	109,1%	114,4%	4,8%

Off-peak period, Optimal behaviour type only

– figures are relative to values for 1-min cell preciseness (equal to 100%)

The relative error variance is not dependent on absolute mean travel time values; i.e. one zone may exhibit almost no variance of travel time averages with the changing spatial preciseness, but another one, with the same absolute distance may be very sensitive. The variance is also (a bit surprisingly) not dependent on type of coverage of the zone with stops and service lines (if they are cross-traversing or tangential with respect of the zone area). Most probably, the relation is too complex for derivation of simple rules, neither based on distance between the origin and destination, nor coverage of the area with the network.

Surely, mean values of travel times calculated for lower levels of spatial detail are strongly dependent on selection of cells, which are to represent the centroids of larger spatial units. Out of thousands of O-D cell relations calculated for each pair of origin and destination zones, about one third have total travel times close to overall zone average (within 5% margin). Nonetheless, the cells, total travel times of which are close to e.g. zone average are not always to be found in the expected positions. This is depicted in the Figure 3, where an example of absolute values of relative deviations |Δ| for one selected O-D zone relation (from Zone 51 to Zone 10) is presented; cells closest to average (darkest ones) are at the left edge of the zone and not in the middle, as could be expected normally.

Figure 3 Absolute values of relative deviations of total travel time from the zone average

$|\Delta| = |(TTT_C - TTT_Z) / TTT_Z|$; averages from all cells of origin Zone 51; destination Zone 10
1-min cell spatial preciseness; Optimal behaviour type; Morning peak;

Even if the total travel time is similar, the portions of component times are very different. Only couple of percent of O-D cell relations fulfil (quite generous) 10% margin condition for at least three out of four component times (total, walk, wait and in-vehicle). All four conditions are fulfilled (if only) for less than two promille (!) of all O-D cell relations on average. Moreover, the individual pairs from the same origin to different destinations (or from different origins to one destination) are often completely diverse.

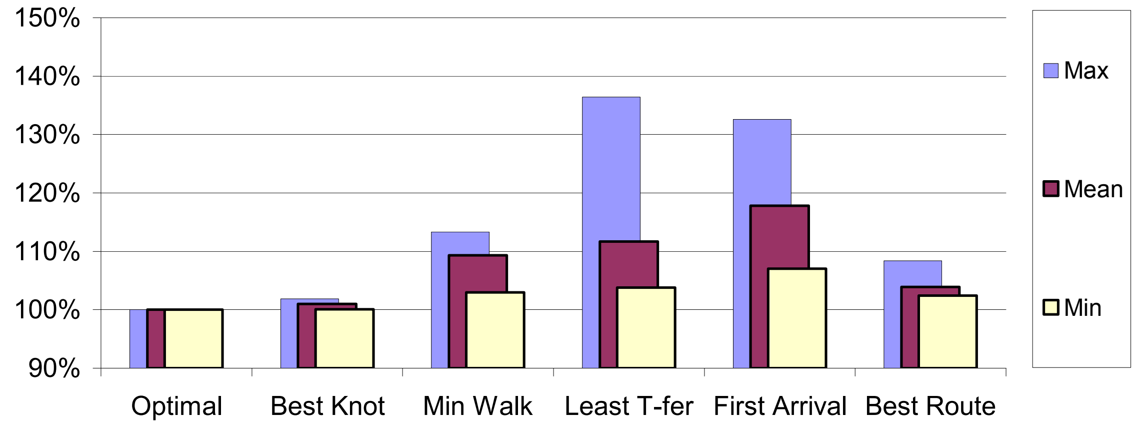
A similar analysis is done for all types of the stop and route choice behaviour (see Chapter 3 to remind). The basic statistical figures of this analysis are summarised in the Table 3.

Table 3 Influence of the choice behaviour – relative TTT_Z figures (1-min cell preciseness)

Choice behaviour	Morning Peak				Midday Off-peak			
	Min	Mean	Max	St Dev	Min	Mean	Max	St Dev
Optimal	-	100,0%	-	0,0%	-	100,0%	-	0,0%
Best Knot	100,1%	101,0%	101,9%	0,4%	100,4%	102,3%	104,6%	1,3%
Min Walk	103,0%	109,3%	113,3%	2,9%	107,3%	112,7%	115,6%	2,8%
Least T-fer	103,8%	111,7%	136,4%	10,6%	107,0%	113,7%	139,2%	10,6%
First Arrival	107,0%	117,8%	132,6%	7,6%	105,2%	114,3%	126,2%	6,9%
Best Route	102,4%	103,9%	108,4%	2,0%	103,5%	106,4%	108,7%	1,7%

The most surprising outcome of this point of analysis is that "Best Knot" choice behaviour type leads to almost exactly the same total travel time values as "Optimal" type. That means that in reality impossible perfect choice behaviour would not lead to significant time savings. Also "Best Route" behaviour type brings fairly comparable times. More specific stop and route choice behaviour types, "Minimal Walk", "Least Transfer" and "First Arrival" lead already to more distinct deviations, over 10%. Off-peak values (relative to off-peak optimal choice values) are only slightly higher compared to morning peak.

Figure 4 Influence of the stop and route choice behaviour on total travel times (morning peak)



Interesting figures come out also from the comparison of component times for each type of stop and route choice behaviour, which are summarised in the Table 4. It is obvious that (from the objective of total travel time minimisation) "Optimal" stop and route choice behaviour leads to slightly higher access times, since sometimes, farther stops are to be chosen. In-vehicle (i.e. actual ride) times are fairly similar for all behaviour types. The desired time savings are achieved through notable reductions of other component times (curiously, including egress!).

Table 4 Influence of the stop choice behaviour on component time values

Choice behaviour	Optimal	Best Knot	Min Walk	Least T-fer	First Arrival	Best Route
Access	100,0%	97,9%	84,8%	97,9%	97,9%	99,6%
Egress	100,0%	100,9%	80,1%	130,3%	153,1%	105,7%
(Sum) Walk	100,0%	99,6%	81,0%	116,3%	129,0%	103,0%
(First) Wait	100,0%	127,1%	162,3%	171,0%	77,3%	170,0%
Transfer	100,0%	112,3%	390,8%	36,7%	201,7%	97,5%
Off-vehicle	100,0%	102,9%	110,7%	118,3%	125,5%	109,8%
In-vehicle	100,0%	98,0%	107,5%	102,6%	108,3%	95,0%
Total (Trip)	100,0%	101,0%	109,3%	111,7%	117,8%	103,9%

Morning-peak; average figures for all observed O-D pairs

– figures are relative to Optimal choice behaviour (equal to 100%)

"Best Knot" behaviour leads to very similar component travel times like "Optimal" choice, with an exception of (first) wait time, which is higher almost by 30% on average. "Minimal Walk" walk times are about 15-20% lower than walk times of all others, but the saving is diminished with significantly higher waiting and transfer times. Opting for "Least Transfer" type may save two thirds of transfer time, which is however quite low in absolute values. By "First Arrival" choice, 25% saving of waiting time is overbalanced especially with 50% longer egress times and doubled transfer times. "Best Route" choice type leads to weighty 70% increase of wait times (because of reduction of alternative lines). Relative (to "Optimal") values in the off-peak period are practically the same as presented morning peak ones[8].

Stop and route choice has also some effects on deviations related with decreasing level of the spatial detail. While maximum error at the full-zone level for the "Optimal" behaviour type is about 20%, for "Min Walk" or "Least Transfer" it is over 25%. Unexpectedly, the least variance is observed for the "First Arrival" choice type (see Table 5 for more details).

Table 5 Cross-influence of the choice behaviour and spatial preciseness on TTT_Z values

Choice behaviour	3-min cell spatial preciseness				full-zone spatial preciseness			
	Min	Mean	Max	St Dev	Min	Mean	Max	St Dev
Optimal	98,4%	100,3%	101,0%	0,8%	98,2%	109,6%	120,8%	6,6%
Best Knot	98,4%	100,3%	101,1%	0,9%	97,4%	109,6%	120,4%	6,6%
Min Walk	98,3%	101,1%	99,9%	1,1%	95,5%	113,5%	125,0%	9,9%
Least T-fer	99,3%	100,4%	102,1%	0,5%	95,4%	109,2%	126,7%	9,5%
First Arrival	98,9%	100,1%	101,4%	0,9%	92,6%	104,2%	113,9%	7,9%
Best Route	98,4%	100,3%	101,0%	0,8%	98,2%	109,6%	120,8%	6,6%

Off-peak period, Optimal behaviour type only

– figures are relative to values for 1-min cell preciseness (equal to 100%)

4.1 Other provisional results

There are hundreds of feasible routes for each of the O-D zone pairs (i.e. routes, total travel time of which could be minimal at particular circumstances). The ACCEPT simulator allows

[8] Note a special feature of the public transport system in Petržalka that may affect this feature; Missing capacity system (tram) is substituted with an extensive bus supply. Main lines are running with the frequency of 3 or 4 minutes during the morning peak! Also in off- peak period run these lines less in headways less than 8 minutes.

also for identification of individual routes that are selected for each of the behaviour types. The provisional analysis insinuates that vast majority of this routes are forsooth "chosen" during the simulation; i.e. they fulfil the conditions of any stop and route choice type at least for one of the cell O-D pairs, and are selected in at least one simulated minute.

However, only several routes are really considerable, that means that they are "chosen" regularly (several times during the simulation hour) at least for several cell O-D pairs, and only few routes are most convenient option in more than only a few per cents of all cases.[9] This is further underlain with the figures for the "Best Route" choice type, as they do not differ notably from "Optimal" or "Best Knot" choice types.

[9] The simulation of one relation between one origin and one destination zone in the ACCEPT Simulator comprises computation of several thousands of O-D cell pairs, repeated 60 times (for every minute of a hour); that makes roughly 100 thousands of possible cases.

5. Conclusions

Influence of the spatial preciseness and stop choice behaviour has been an object of the analysis presented in this paper. The existence of the modifiable area unit problem has been confirmed also for the public transport networks on the example of the part of the network in the city of Bratislava. That means, that the calculation of average travel times for a study area at different levels of spatial detail (i.e. the study area is divided into different number and extent of zones) leads to different outcomes.

Travel times averaged for zones divided into cells with diameter equal to 3-minutes of walk (R = 200-250m and cell area < 3 ha) are yet practically the same as those computed out of 1-minute walk time cells (area ~ 0,45 ha). But if the study area is divided into zones with the size over 30ha (denoted in this case study as full-zones), the error may reach more than 25% in some cases. The impacts of decreasing level of spatial detail, measured in average relative deviations in total travel time calculations, are rising exponentially (see Figure 2). This suggests existence of a threshold, below which increased level of the spatial detail brings no significant improvement of outcomes precision. This could be important for reducing of computing time and memory requirements.

The mean travel times computed at lower levels of spatial preciseness (larger cell units) are mostly overestimated compared to ones computed at one minute preciseness. This is partially caused by the fact that larger cell units may comprise already few "empty" areas, which were eliminated by definition of the one-minute cells. The estimation of the average walk and wait times is the main source of calculation inaccuracy in practical models. Surely, these component times depend merely on spatial coverage of the zone with the network (stops) and levels of service provided on each of its sections. Due to complexity of these relations, no general rules could be derived for a proper estimation of the two times, however.

The actual values of travel time are affected by the user's stop choice behaviour as well. "Optimal" behaviour with variable stop choice for every moment of departure from the origin would not bring a significant reduction of total travel times compared to (usual) choice of the "Best Knot". More specific types of behaviour, "Min Walk "Least Transfer" and "First Arrival" may lead to prolongation of the total travel times by more than 10%. Even more weighty increments are likely for individual component times. The variance of total travel times among individual behaviour types is principally caused by different walk and wait times. This may have consequential impacts on soundness of the models, in which different weights of component times are applied in total travel time or utility calculations.

5.1 Related issues

Factually, a representative cell with "average" values of the total travel time and its most important components cannot be found for larger zones. Hence, it is not possible to calculate the travel time values of a zone from a single point (centroid) as it is de facto (though virtually) presupposed by the common classical modelling techniques.

Absolute maximal variance of 2 or 3 minutes observed for different levels of the spatial preciseness or stop and route choice behaviour may not seem critical at the first sight. This is certainly true for many tasks of general transportation modelling. However, it can seriously bias the modelling results when larger zones/cells are used; particularly, when popular logit function is applied for the mode choice, and if total travel times by public transport and private car modes (as primary variables of utility) are similar for the particular O-D relation.

On the other hand, several thousands of 3-minute cells are needed just to cover the built-up area of a medium-sized city like Bratislava. This is indeed far beyond applicable number of zones, for which some other necessary modelling data that can be obtained (for instance to built an O-D matrix or to validate the model). In other words, spatial disaggregation is limited due to feasibility constraints in practice.

This problem can be overcome if two kinds of spatial units are distinguished: (i) zones, which are the elementary modelling units (i.e. as they are understand now); and (ii) cells (as a unit of higher spatial preciseness) for which travel times (and perhaps also some other aggregated model variables) could be computed. The zones could be large enough; adjusted for instance to normal administrative division or available amount of data. Each zone would be further divided into several cells (smaller than 3 ha).

The average travel times of a zone required for further modelling could be then determined as the mean values from all cells. If the zones are not homogeneous enough with respect to land use functions, weighted averages need to be calculated instead. Furthermore, if a model does not discern different user groups (e.g. general accessibility models), travel times could be averaged according to (surveyed) particular proportions of stop and route choice behaviour.

Another issue is related with reliability. In reality, every public transport system runs with imperfections, such as delays or cut-off die to accidents. The figures presented in this paper do not involve considerations on system failures. In general, the travel times by means of public transport are variable around these values with certain probability distribution. Therefore, they should be further averaged accordingly. For a case study on impacts of reliability on travel times see (Ľos et al., 2008). Accommodation of reliability issues requires further modifications of the current computation procedures in the ACCEPT Simulator. A detailed discussion is beyond the subject of this paper.

5.2 Further applications

Simplifications and errors in travel time calculations are definitely not the only ones in transportation modelling. Surely, in many cases effects of the resulting errors brought into models are negligible compared to other impacts. However, true sensitivity analysis is rarely done. Thus, although it may be not necessarily required, elimination of this issue could be welcomed to enhance the validity and credibility of practical (public transport) models.

For instance, it is not quite possible, veritably to evaluate the changes of accessibility brought about with modifications of the public transport system, such as alternation of routing of the lines or their timetables. Effects of alternative options on travel times may be often of minor variance that could be hardly detected with existing modelling techniques. A tool for precise travel time computations like ACCEPT Simulator could be therefore appreciated by optimisation of the public transport networks.

There are several other applications planned with using the ACCEPT Simulator software. The analysis of impacts of spatial preciseness and stop choice is actually first part of a broader intended study. Existing accessibility indicators, measures and models adopt more or less point of view of authorities, planners and/or operators (see Ľos 2007a). The general objective is therefore to create an assessment tool of accessibility in real public transport systems from the passengers' perspective).

In 2005, a fundamental change in organisation of public transport in Bratislava-Petržalka has been executed. The network of many direct links of lower frequency has been replaced by a system of more frequent lines, but more routes that require transfers. This offers a unique opportunity to compare this two approaches in real conditions, from the point of view of the direct accessibility benefits of users that should come out from theoretically intended and expected reductions of the travel times (if there are any) and factual savings of the operation costs.

A similar application that is planned is analysis of the synergy effects of different public transport lines and/or sub-systems (metro, tram, bus). The principle of the objective is to find out what is the enhancement of accessibility compared to increased costs of operation, or its decrement when one of the subsystems is out of operation.

The last idea is to examine relation between the accessibility and mobility indicators, particularly modal split. Based on precise travel time calculations in the framework of the full ACCEPT methodology (see Ľos, 2007b) the could arbitrated, if accessibility and provided public transport supply quality is really of so intrinsic as it is expected, or other aspect, such as cultural and historical conventions are of more importance.

6. Acknowledgement

This work is supported by the European Community within the TRANSPORTNET's Marie Curie Fellowship for Early Stage Researchers Training framework.

7. References

Alvanides, S., Openshaw, S., Duke-Williams, O. (2000) Designing zoning systems for flow data in GIS and Geocomputation, In: Innovations in GIS 7 Eds P. Atkinson, D. Martin (Taylor and Francis, London), pp. 115 -134.

Ben-Akiva, M.E. et al. (1984) Modelling inter-urban route choice behavior. In: Proceedings 9th International Symposium on Transportation and Traffic Theory. (Editors: Volmuller, J. and R. Hamerslag), VNU Science Press, pp. 299-330.

Bliemer, M.C.J. & Taale, H. (2006) Route generation and dynamic traffic assignment for large networks. Paper DTA Conference Leeds, June 2006.

Bovy, P. H. L. (2007) Modeling Route Choice Sets in Transportation Networks: A Preliminary Synthesis, TRISTAN V Conference, Phuket, Thailand, June 2007.

Cascetta E. and A. Papola (2001) A model of route perception in urban road networks, *Transportation Research B* 36, pp. 577-592.

Chang, K. T., Khatib, Z., Ou Y. M. (2002) Effects of zoning structure and network detail on traffic demand modeling, In: Environment and Planning B 29, pp. 37-52.

Chou, Y-H. (1991) Map resolution and spatial autocorrelation, Geographical Analysis 23, pp 228-246.

Ding, C. (1998) The GIS-based human-interactive TAZ design algorithm: examining the impacts of data aggregation on transportation-planning analysis, Environment and Planning B: Planning and Design 25, pp. 601- 616.

Duckham, M., Mason, K., Stell, J., Worboys, M. (2001) A formal approach to imperfection in geographic information, Computers, Environment and Urban Systems 25, pp. 89-103.

Eagleson, S., Escobar, F., Williamson, I. (2002) Hierarchical spatial reasoning theory and GIS technology applied to the automated delineation of administrative boundaries, Computers, Environment and Urban Systems 26, pp. 185-200.

Fiorenzo-Catalano, S. (2007) Choice set generation in multi-modal transport networks. PhD Thesis, TRAIL, Delft University of Technology, The Netherlands.

Friedrich, M., Hofsass, I. & Wekeck, S. (2001) Timetable-based transit assignment using branch and bound. Transportation Research Record 1752, 100-107.

Hoogendoorn-Lanser, S. (2005). Modeling travel behavior for multi-modal transport networks, TRAIL Thesis Series T2005/4, TRAIL, The Netherlands.

Liu, C.-L., Pai T.-W., Chang C.-T. and Hsieh C.-M. (2001) Path-Planning Algorithms for Public Transportation Systems. In: 2001 IEEE Intelligent Transportation Systems Conference Proceedings, Oakland (CA) USA, August 2001, pp. 1061-1066.

Liu, C.-L. (2002) Path-Planning Algorithms for Public Transportation Systems. In: IEEE 5th International conference on Intelligent Transportation Systems Conference Proceedings, Singapore, September 2002, pp. 834-839.

Ľos, P. (2007a) Mobility Indicators and Accessibility of Transport, in: Slovak Journal of Civil Engineering, Vol. XV, 2007/1, Bratislava, Slovakia, pp. 24-32.

Ľos, P. (2007b) Measuring Public Transport Accessibility and Supply Quality from the Users' Perspective: ACCEPT Methodology, TRANSPORTNET section at Swiss Transport Research Conference 2007, Monte Verita, Switzerland, September 2007.

Ľos, P., Coelho, P., Spina, G. F., and Karyda, M. (2008) Assessment of Urban Public Transport Systems from Passengers' Perspective, paper submitted for NETNOMICS journal in May 2008.

Marzano, V. & Papola, A. (2004) A link-based path-multilevel logit model for route choice which allows implicit path enumeration. Proceedings ETC Strasbourg, France.

Openshaw, S., Taylor, P. (1979) A million or so correlation coefficients: three experiments on the modifiable areal unit problem, in: Statistical Methods in the Spatial Sciences Ed. N Wrigley (Routledge and Kegan Paul, London), pp. 127-144.

Openshaw, S., & Rao, L. (1995) Algorithms for Reengineering, In: Census geography, Environment and Planning A 27, pp. 425-446.

Ortuzar, J. & Willumsen, L. (2001) Modeling Transport, 3rd edition (JohnWiley, New York)

Prato, C. G. & Bekhor, S. (2006) Applying branch & bound technique to route choice set generation. Paper TRB Annual Meeting, CD-Rom.

PTV, (2007) VISUM 10.0 User manual, PTV AG, Germany.

Ramming, M.S. (2002). Network knowledge and route choice, PhD thesis, MIT, USA.

Viegas, J. M., Martinez, L. M., Silva, E. A. (2008) Effects of the modifiable area unit problem on the delineation of traffic analysis zones, Environment and Planning B: Planning and Design, advance online publication, doi:10.1068/b34033.

Wu, Q. & Hartley, J. (2004) Accommodating User Preferences in the Optimization of Public Transport Travel. In: 4th European Congress and Exhibition on ITS, 2004

Xie, Y. (1995) The overlaid network algorithms for areal interpolation problem, *Computers, Environment and Urban Systems* 19, pp. 287-306.

Zhang, M. & Kukadia, N. (2005) Metrics of urban form and the modifiable areal unit problem *Transportation Research Record* # 1902, pp. 71-79.