Network Assignment and Equilibrium for Disaggregate Models

John Gibb
Senior Transportation Engineer
DKS Associates
8950 Cal Center Dr, Suite 340, Sacramento, CA, USA, 95826
jag@dkssacramento.com

Submitted May 29, 2009 to the proceedings of the 12th National Transportation Planning Applications Conference of the Transportation Research Board, Houston, Texas, May 2009

Abstract

Practical implementations of activity-based travel demand models normally apply a disaggregate process to individual households and persons, generating their activity demands and other decisions leading ultimately to a list of individual trips. This enables heterogeneous choice behavior, finer spatial detail than zones (TAZs), deeply nested choice sets, and preserves linkages between persons and their itineraries and trips. However, assignment of those trips to the network typically relies on software and methods built for customary four-step aggregate models, which assign centroid-to-centroid in few discrete classes of path choice criteria (such as HOV lane use eligibility). Finer spatial detail and heterogeneity in path choice are deterred by the excessively large matrices and/or many shortest-path tree computations required. This study examines the practicality of individual trip assignment at overcoming these limitations, in terms of computational performance and empirical convergence toward equilibrium. The tested method uses the A-star path-choice algorithm and a simple iterative loading procedure. One path per individual trip is chosen, and a full relational linkage is retained between each trip, path, link, itinerary, and person in the model. Using a set of individual trips from an activity-based demand model, a disaggregate network assignment was performed, and shown to converge toward equilibrium with comparable performance to a conventional trip-based assignment.

Introduction

For regional travel demand forecasting, application of behaviorally rich activity-based models has been made practical, to a significant extent, by the method of disaggregate microsimulation (Bradley, Bowman, Lawton, 1999). Here this means that choices are modeled for each person in each household in the modeled region, individually (or in small bundles at nearly individual scale), and one outcome is drawn for each choice, randomly with the model's predicted probability. (Microsimulation here does not necessarily refer to a finely-detailed sequential process in time, as the term is sometimes used.)

Disaggregate microsimulated activity-based models produce a synthetic itinerary of trips, trip-groups such as tours and day-patterns, all linked to each another and the individual person making them. Application at individual or small units of scale enables modeling of heterogeneous behavior with numerous explanatory variables. Single-outcome application, usually with Monte Carlo sampling, enables deeply-nested choices and submodels, by "paring the tree" down from exploding numbers of choice alternatives. Each demand-unit's choices can be saved, retaining a full relational linkage between all households, persons, itineraries, and trips. Application upon large numbers of small demand units causes total demand to each location, mode, etc. to approximate what would result from aggregate allocation to every possible choice.

By contrast, conventional four-step models are applied to quantities of households and employees in each zone, each in limited numbers of classifications. These models split demand to all alternatives in proportion to their probabilities. To conserve storage space, demand is re-aggregated at various stages of application, as distinctions that determine choices are no longer necessary for later stages. (For example, in conventional trip generation, several different land use types each generate trips in several trip purposes; trips are output by trip purpose, but no longer distinguished by their source land use. Later, origin-to-destination trip matrices distinguish trips by time-of-day, but seldom retain trip purpose distinctions.) Linkages between trips and trip-makers are lost. It is difficult to apply any but the simplest activity-based models in this aggregate manner.

As far as the author has determined, no practical regional travel demand model extends disaggregate microsimulation into the trip assignment stage. In common practice, trips produced by the activity-based model are aggregated into matrices, classified into a few time periods, occupancies (for eligibility for HOV lanes, etc.), and possibly toll-paying and toll-refusing classes, for use by conventional static equilibrium assignments such as in Cube Voyager (Citilabs), EMME (INRO Consultants), or TransCAD (Caliper Corp.). Lengthy runtimes, efficiency from bulk processing, and memory requirements tend to dictate the number of different matrices assigned to be minimized.

Some application capabilities gained or made more practical using disaggregate assignment include:

1) Spatial resolution of origins and destinations, to finer detail than typical TAZ centroids,
2) Heterogeneous path choice criteria, such as differing sensitivity to price, reliability, or other variables,
3) Stochastic path choice, with either random "noise" throughout the network, or structured random variation of certain links,
4) Linkage between paths and the particular vehicles and persons on them, and their personal and household data, available to post-run analysis.
5) A means to warm-start an assignment after the activity-based model has revised the itineraries of some or all of the model's persons.
6) A closer integration between the assignment and the activity-based model itself.

A further consideration is to eventually integrate dynamic assignment with activity-based demand modeling. Lin, Eluru, Waller, and Bhat (2009) argue that both types of models, under separate development for several years, have much to gain by integration. In particular, both activity-based demand models and dynamic traffic assignment represent people's movement in continuous time. With an activity-based demand model, static assignment "undoes much of the advantages of predicting travel patterns in continuous-time." And providing demand for a dynamic assignment in a few aggregate time periods "does not exploit the very purpose for which DTA models have been developed." Their paper studied convergence properties of a unified model system, but did not question the use of aggregate trip matrices as the conveyance of demand.

While this paper presents results limited to static assignment, it can be stated that direct dynamic assignment of individual trips from the activity-based model is a logical and beneficial form of integration of the two models. Individual trips, each with a departure time and a path, are "simulation-ready." It is noteworthy that some commercial dynamic assigners synthesize individual vehicles from demand matrices for sake of their simulations (e.g. Yang 2009). But if individual trips are already available from an activity-based model, it is a loss of information and a waste of storage to aggregate them to matrices only to re-synthesize a different set of individual vehicles within assignment.

TRANSIMS has a disaggregate route choice module, and an interrelated dynamic simulation module. While TRANSIMS is gaining gradually larger-scale deployments, applications known to the author remain highly specialized in scope and coverage, focused on high spatial, temporal, and traffic-control detail, and expensive in data preparation and computing. This study explores a more basic and generic disaggregate assignment intended for practical regional model applications for planning, not attached to any particular network performance model, yet which can be built up to greater detail.

Path choice heterogeneity in one continuous dimension has been developed for aggregate assignment in T2 (Dial 1995). T2 overcomes a problem of multiclass assignment methods: if discrete values-of-time are applied to costs, then some paths may be overlooked and the rest suffer "lumpy" loadings. While a significant achievement for aggregate modeling, such multicriteria assignment is simple in a disaggregate setting, where it readily extends to more than one criterion dimension and/or dynamic assignment.

This study focuses on empirical results for large-scale practical application. Individual-trip assignment has rarely been studied theoretically, unlike the well-studied problems of equilibrium flow in a continuous domain. Work by Bernstein (1990) was brought to the author's attention shortly before this paper was due for submittal, and will be reviewed for further guidance. It is anticipated that some amount of random "noise" in assignment, as well as the activity-based demand model, limits the possible or meaningful convergence of individual trips toward equilibrium. To pursue the limits of convergence is deferred to further study.

The remainder of this paper identifies a well-established path-choice algorithm suitable for individual trips, proposes a procedure to load trips onto the network using this algorithm, and presents results for their computational performance and convergence toward equilibrium applied to regional model data.

Reducing the Computational Cost of Disaggregate Assignment

It might seem that the computational cost of assigning individual trips would be prohibitive. Conventional assignment relies on the bulk efficiency of assigning a whole matrix row at a time. Specifically, the classic path-building algorithm of Dijkstra (1959) builds the shortest-path tree from an origin to all destinations at a time. Then the respective matrix row is assigned to the links by an efficient one-pass inward accumulation from leaf to root. Individual trip assignment requires building a tree for each trip, which in many models numbers in the millions, instead of hundreds or thousands of TAZ-based trees in conventional assignment.

The two methods of this study to reduce the computational burden of individual-trip assignment are:

1) The A-star algorithm, and
2) Assignment of only a small fraction of trips per iteration.

The A-star (A*) algorithm (Hart, Nilsson, Raphael, 1968) modifies the Dijkstra algorithm by choosing the next node of outward search in order of estimated total impedance of the trip to the destination. This estimate is the actual individual-dependent impedance from the origin to the node, plus an optimistic estimate of impedance from that node to the destination, known unfortunately as the “heuristic” impedance.

A-star’s use of “heuristic” impedance does not mean a non-shortest path is chosen. The path chosen is shortest if the heuristic impedance at a given node is less than or equal to that to any other connected node, plus the link impedance from that node (explained in *e.g.* Beeker 2004). The choice of heuristic (meeting this condition) determines the efficiency of the algorithm in finding a true best path, but not the actual path found (except possibly in case of a tie). In fact, if the heuristic impedance is set to zero for all movements, the resulting algorithm is equivalent to Dijkstra’s. (A more liberal condition for the heuristic is valid with a slight modification to the algorithm.)

The chosen heuristic impedance in this study is the shortest-path time using the free-flow travel time on all links except zone-access links, whose possible minimum is zero. Before any trip loading begins, the test program fills this heuristic Nodes × Zones matrix using the Dijkstra algorithm. While fairly large, this matrix stays in an array for the duration of the trip assignment process.

The other method to minimize the computational burden of individual trip assignment is to assign only a fraction of the trips during each iteration. Each iteration of conventional assignment chooses new paths for all trips, but reroutes only a fraction of the trips to those new paths. That fraction is the “step size.” A disaggregate assignment with indivisible trips doesn’t permit fractions of trips to be rerouted. Instead, an alternative meaning is applied here: an iteration draws a sample of trips to reroute; the step size is the relative size of that sample out of the population of all trips. All trips are processed in “passes,” so that every trip is sampled once per pass through the population of trips.

Step sizes are chosen in advance; no "line search" for an optimal step size is applicable, because the trips are indivisible. Link times are simply adjusted to the new flows after each iteration. Computational effort is mainly proportional to the number of passes through the trips, not the number of delay update iterations taken. Each pass takes numerous iterations to complete.

Before the first pass is completed, the population loaded is incomplete. Consequently a temporary scaling factor proportionally expands accumulated link flows, to estimate the full-population equivalent.

During the second and subsequent passes, while each trip is reassigned a new path, its previous path is reviewed to:

1) Deduct the trip from flows on the old path, and
2) Report that particular trip's travel impedance of its old path and the new path (both reckoned at the current link impedances).

Afterwards, the old path is not used or otherwise present in the model; only the final path remains. Link flows represent the latest paths only.

An Experimental Test Assignment

The experimental test assignment reported here loaded about 4,500,000 trips from an "unofficial" test version of the SACSIM activity-based demand model using DaySim (Bowman and Bradley, 2006) onto a network having nearly 1500 zones, 11,400 nodes (including zones, 14,900 counting unused node numbers), and 25,200 directional links. Links representing HOV lanes are present in the network, and the individual trips are marked for their eligibility to use them. For comparability with SACSIM's trip-based assignment, the day was divided into four time periods (AM and PM 3-hour peak periods, mid-day, and evening-overnight) and each trip categorized into one of these periods. The trip-based assignment's volume-delay functions were programmed into the disaggregate assignment program.

All trips in SACSIM begin and end at "parcels," of which there are over 600,000 in the model system. To improve the spatial resolution of assignment of individual trips, a technique was devised to adjust travel times on centroid connectors and the immediately connected street links based on the coordinates of the parcels of origin and destination. The two methods of this technique are:

1) For a given trip, all centroid connectors of the origin and destination TAZ are adjusted to the actual length between the parcel and the other node (as if the centroid were moved to the parcel);
2) For movements through a parcel-adjusted centroid connector, and a next-connected link (besides another zone connector) that make an acute-angle turn, the point of the turn is moved to the closest point within the adjacent link. (Additional virtual links and nodes are generated, similar to those needed to correctly model turn penalties.)

These node and link distance adjustments do not appear in the resulting loaded network; instead, they are processed like an individualized perception of the network.

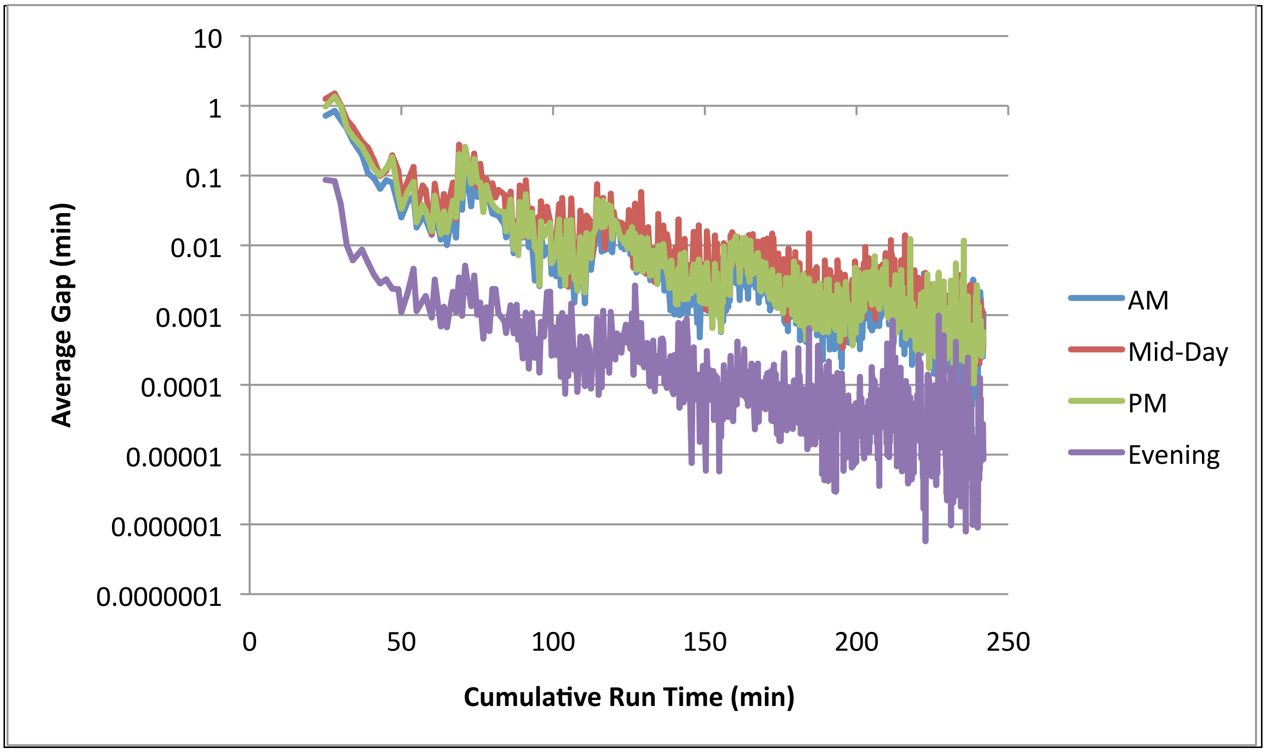
An experimental loading step-size schedule was programmed into the disaggregate assigner having 6 complete passes through all 4,500,000 (approx.) vehicle trips, 900 total iterations (link-delay updates), gradually-declining step sizes containing about 300,000 trips in early iterations, gradually declining to nearly 7,100 trips in the last iteration. All four modeled time periods are assigned concurrently.

The test computer has dual Intel processors at 2.4 GHz and 2 GB of RAM. The test program is in compiled Visual Basic 2008 with no programmed multithreading. All files in and out were text. Complete link sequences of paths were written, each pass's being about 1.3 GB in size. Reading the network and computing the heuristic time matrix took around 12 seconds. The first assignment pass through all the trips took about 20 minutes, and all passes thereafter took around 40 minutes. (From that 2:1 ratio, it is speculated that input and output of paths, rather than path-building, dominate these timings.)

Individual trip lengths on old and new paths (for all but the first pass) were valuable in validating the program's functionality. The new path's impedance must be less than or equal to the total current impedance along the old path. The difference is the individual trip gap. Gaps of the wrong sign (beyond numerical precision error) would clearly indicate path-choice or measurement error. All gaps should approach zero if and when the system converges toward equilibrium.

Figure 1 presents the gaps of individual trips averaged by assignment period of day and iteration, beginning with the second pass through the trips. Gaps generally decline during each pass, although trips earlier in the sequence of each pass tend to have larger gaps than those later, probably due to fluctuations in flows during the first pass. After nearly four hours of computation, gaps average roughly 0.001 minutes for all periods except the much less congested evening-overnight period. The apparently growing ranges of fluctuations are largely due to the smaller numbers of trips being averaged in late iterations versus early.

Figure 1
Average Gap of Assigned Trips in Each Iteration, Disaggregate Trip Assignment



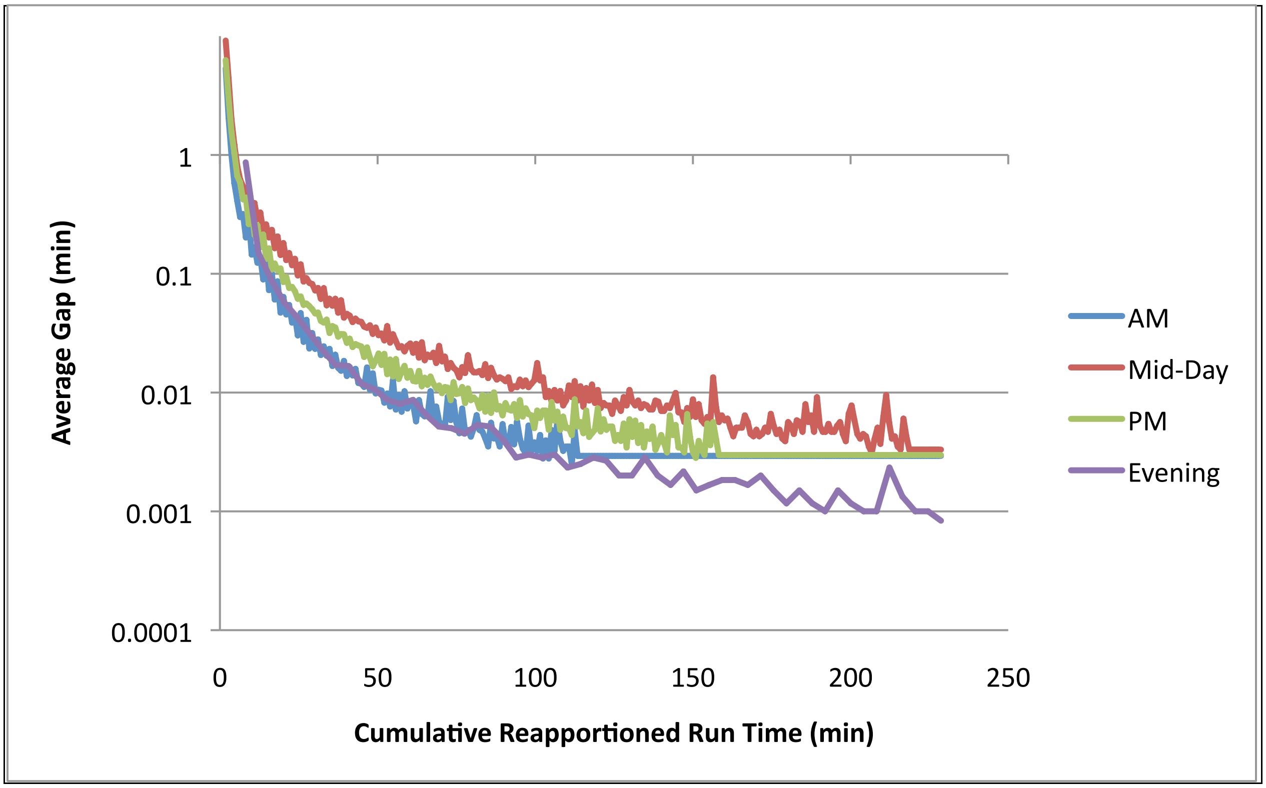
The average gap of all trips is similar to the gaps reported as a convergence measure by common modeling software programs, when expressed in time units. Specifically, for conventional aggregate assignment:

□□□□□□□□□□ □□□□□□□□ □□□= ∀□□□□□ □(□) −∀□□ □□□□□□□ ∀□□ □□□,

where $f(v)$ is each link's volume-delay function of flow v, and for movement ij, $d_{ij}$is the vehicle-trip demand and t_{ij} is the minimum path impedance.

Figure 2 presents the average time gaps from a conventional Frank-Wolfe assignment of the same trips, on the same network and delay functions, using popular commercially-available modeling software, on the same computer. All periods except the evening-overnight failed to continue to converge when the average gap neared 0.003 minutes, apparently due to a limit of numeric precision. (The vendor was one at this conference presenting much more precise and fast convergence using new non Frank-Wolfe methods in their newest version.) The assignments for each of the four time periods were run sequentially with a total runtime of nearly 230 minutes; in Figure 2, the runtimes are reapportioned into this same total as if run concurrently.

Figure 2
Average Gap of Each Iteration, Conventional Assignment



Had the convergence of three of the periods not stalled, it might be surmised that those periods' gaps would have continued to not be significantly less than the evening period's, which finished near 0.001 minutes. This is close to the average gap of most trips on their last path-updates in the disaggregate test.

While both the disaggregate and conventional assignment gaps both measure differences in trip times, there are some significant differences in definition. The disaggregate gaps in an iteration are from only those trips whose paths are being updated. To measure the gaps of all trips at once would require computing the shortest paths of all trips – an exercise without much economy if not presently updating the paths of any but a few. The trips being updated at any point in the process are the trips with the "oldest" paths, which haven't been updated since last encountered in the preceding pass. Consequently, the reported disaggregate gaps are likely "worst case"; gaps of more recently-updated trips are likely lesser, since the network flows haven't had as much opportunity to change since they were updated.

Since gaps are determined individually for each trip, these may be examined and summarized in other ways than averages. In the final pass, several individual trips had gaps exceeding a little over one minute. Most of these trips are much longer than average. Individual trips with large gaps may be examined in detail – a priority for further study.

Conclusions

Two different assignment methods were applied to the same trips and network – a disaggregate assignment of individual trips from an activity-based demand model, and a conventional Frank-Wolfe assignment of those same trips aggregated to matrices. Run for about the same total runtimes, they achieved roughly similar average trip-time gaps. In the author's opinion, this indicates disaggregate assignment is practical, suitable for special purpose or even routine usage with a regional activity-based model system as a substitute for conventional aggregate static assignment. Warm-start ability adds to its practicality for system equilibrium with "feedback."

Spurred by developments by Bar-Gera (1999), and Dial (2006), new versions of popular modeling software converge much faster than the conventional Frank-Wolfe assigner used for comparison in this report (Slavin et al. 2006, Marimo and Zhou 2009, Florian et al. 2009). This disaggregate assigner does not compete with those for speed upon the standard matrix-assignment problem. Instead, it pursues different capabilities. Nonetheless, the disaggregate assigner used for this report stands to be sped up too. Several significant impediments against its computational and convergence performance have been identified in this report that can be overcome with straightforward programming improvements. Experimenting with different iteration step-size schedules may also yield improvements.

Acknowledgements

The test data were created during activity-based model development, testing, and upgrading for the Sacramento Area Council of Governments, in collaboration with John Bowman and Mark Bradley, but do not represent this model as it has been applied for policy analysis.

The author thanks Howard Slavin for bringing to attention the discrete trip assignment equilibrium work of David Bernstein.

References

Bradley, Mark A. , John L. Bowman and T. Keith Lawton (1999) *A Comparison of Sample Enumeration and Stochastic Microsimulation for Application of Tour-Based and Activity-Based Travel Demand Models*, European Transport Conference, September 1999, Cambridge, UK.

Bowman, J. L., Bradley, M. A. (2006) Activity-Based Travel Forecasting Model for SACOG: Technical Memos Number 1 through 11 (available at `http://jbowman.net`)

Lin, Eluru, Waller, Bhat (2009) *Integration of Activity-Based Modeling and Dynamic Traffic Assignment*, The University of Texas at Austin, Department of Civil, Architectural & Environmental Engineering,
`http://www.ce.utexas.edu/prof/bhat/ABSTRACTS/CEMDAP_VISTA_pub.doc`

Dial, Robert (1995) *Multicriterion Equilibrium Traffic Assignment: Basic Theory and Elementary Algorithms, Part 1, T2: The Bicriterion Model,* Volpe National Transportation System Center, Cambridge MA

Yang, Qi (2009) *A Dynamic Traffic Simulation Model on Planning Networks*, presented at 12th National Transportation Planning Applications Conference of the Transportation Research Board, Houston, Texas, May 2009

TRANSIMS: `http://tmip.fhwa.dot.gov/community/user_groups/transims` or `http://tmip.fhwa.dot.gov/resources/clearinghouse/docs/transims_fundamentals/` and linked documents.

Bernstein, D.H., 1990, *Programmability of continuous and discrete network equilibria,* Ph.D. dissertation, University of Pennsylvania

Dijkstra, E. W. (1959), "A note on two problems in connexion with graphs", *Numerishe Matematik* 1, 269-271.

Hart, P. E.; Nilsson, N. J.; Raphael, B. (1968). *A Formal Basis for the Heuristic Determination of Minimum Cost Paths*. IEEE Transactions on Systems Science and Cybernetics SSC4 4 (2)

Beeker, E., (2004), Potential Error in the Reuse of Nilsson's A Algorithm for Path-Finding in Military Simulations, *Journal of Defense Modeling and Simulation, Applications, Methodology, Technology*, Vol. 1, Issue 2, April 2004, The Society for Modeling and Simulation International, San Diego, `http://www.scs.org/pubs/jdms/vol1num2/beeker.pdf`

Bar-Gera, H. (1999) *Origin-Based Algorithms for Transportation Network Modeling*, National Institute of Statistical Sciences, Technical Report Number 103.

Dial, Robert, (2006) A path-based user-equilibrium traffic assignment algorithm that obviates path storage and enumeration, *Transportation Research Part B, Methodological*, Vol. 40 Issue 10, Elsevier.

Slavin, Brandon, Rabinowicz (2006) An Empirical Comparison of Alternative User Equilibrium Traffic Assignment Methods, European Transport Conference 2006.

Martimo, M., Zhong Zhou (2009) *New Assignment Methods in Voyager – Theory and Tests*, Citilabs, `http://www.citilabs.com/usergroup/files/futurapresentations/MMartimo-ZZhou_NewAssignmentMethodsinVoyager.pdf`, accessed May 4, 2009

Florian, M., Constantin, I., Florian, D. (2009) *A New Look at the Projected Gradient Method for Equilibrium Assignment*, Transportation Research Board 88th Annual Meeting