



CIRRELT

Centre interuniversitaire de recherche
sur les réseaux d'entreprise, la logistique et le transport

Interuniversity Research Centre
on Enterprise Networks, Logistics and Transportation

Logit Choices in Strategy Transit Assignment

Michael Florian
Isabelle Constantin

September 2011

CIRRELT-2011-53

Bureaux de Montréal :

Université de Montréal
C.P. 6128, succ. Centre-ville
Montréal (Québec)
Canada H3C 3J7
Téléphone : 514 343-7575
Télécopie : 514 343-7121

Bureaux de Québec :

Université Laval
2325, de la Terrasse, bureau 2642
Québec (Québec)
Canada G1V 0A6
Téléphone : 418 656-2073
Télécopie : 418 656-2624

www.cirrelt.ca

Logit Choices in Strategy Transit Assignment

Michael Florian^{1,2,*}, Isabelle Constantin²

¹ Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation (CIRRELT)

² INRO, 376 Victoria Avenue, Suite 200, Montréal, Canada H3Z 1C3

Abstract. Since it was first developed (see Spiess and Florian (1)) the strategy-based transit assignment has been extensively used and its properties are well understood. The computation of an optimal strategy for one destination is relatively fast and is comparable, more or less, to the computation of a shortest path tree for one destination. However, since it is the solution of a linear program it produces extreme solutions. The set of strategies that are considered in a transit assignment may be enlarged by introducing elements of discrete choice modeling into the route choices that are considered. One can include a larger number of strategies in the assignment by allowing multiple strategies to be used, one from each connector access to the transit network. In addition some walk choices may also be added by considering explicitly choices made between different transit stops that are served by different lines and require one or more walk links for access. This development is based in part on the work of Nguyen et al (2) in which an algorithm for a logit choice of strategies was developed. The difference between the model considered in the Nguyen et al paper and this development is that walk was not considered as an option in the construction of the strategies. The walk mode was considered only on connector links from an origin node to the transit network. The basic transit assignment based on the computation of optimal strategies is generalized first to compute simultaneously the best paths that do not require walking and the best paths that may access a transit stop with one or more walk links. Then the set of walk links considered is augmented by adding “efficient” walk links (following a basic idea due to Dial (3)). The modified strategy transit assignment algorithm was shown to produce more realistic results in networks where the zones are large and there are several connector links; in dense transit networks where relatively short walks are required for access to attractive alternative transit paths; and in better modeling of alternative fares.

Keywords. Transit assignment, optimal strategies, discrete choice models.

Results and views expressed in this publication are the sole responsibility of the authors and do not necessarily reflect those of CIRRELT.

Les résultats et opinions contenus dans cette publication ne reflètent pas nécessairement la position du CIRRELT et n'engagent pas sa responsabilité.

* Corresponding author: Michael.Florian@cirrelt.ca

Dépôt légal – Bibliothèque et Archives nationales du Québec
Bibliothèque et Archives Canada, 2011

© Copyright Florian, Constantin and CIRRELT, 2011

1. INTRODUCTION

Since it was first developed (see (1)) the strategy-based transit assignment has been extensively used and its properties are well understood. The computation of optimal strategies is relatively fast and is comparable, more or less, to the computation of a shortest path tree for one destination (the optimal strategy algorithm is given in the Appendix). However, since it is the solution of a linear program it produces extremal solutions. See the simple example in Figure 1.

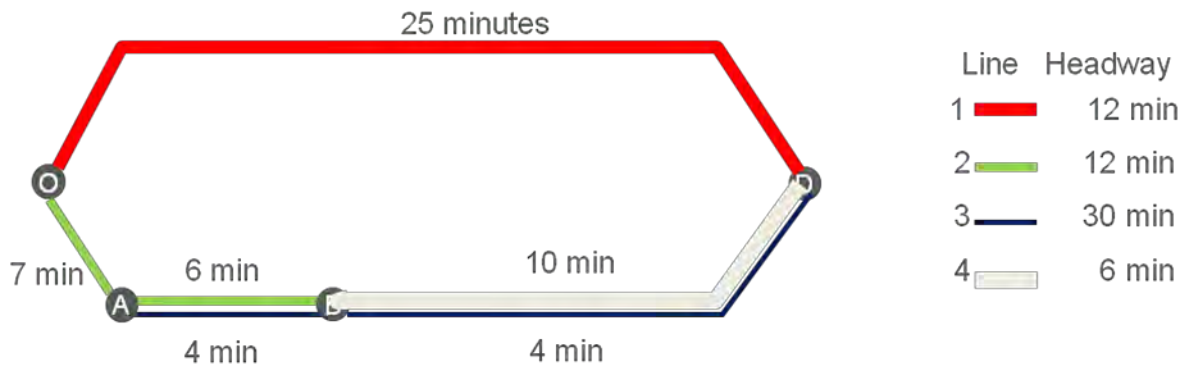


Figure 1. A 4 line example.

There are 4 transit lines that provide service from node O to node D. The line headways are given in the legend and the travel times are indicated on the transit line segments. The number of travelers that board at node O is 100 and they are all destined for D. The volumes corresponding to the optimal strategy are shown in Figure 2.

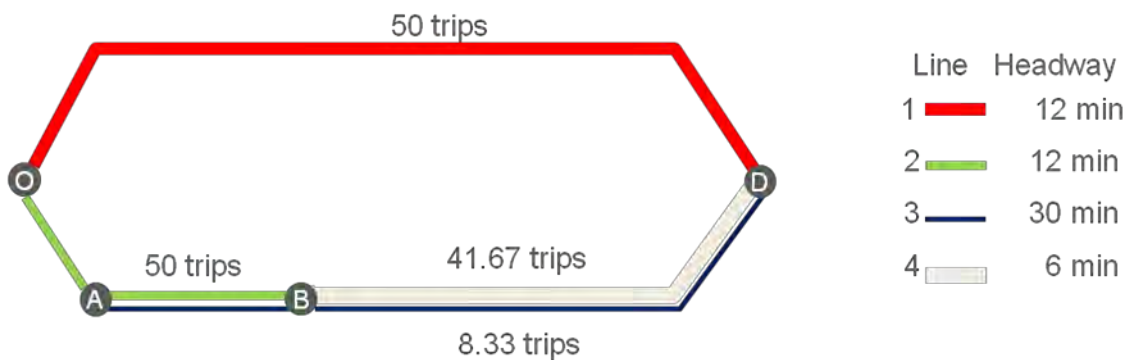


Figure 2. A 4 line example: the optimal strategy.

At node O there are 50 trips that board line 1 and alight at node D; the other 50 trips board line 2, alight at node B and then use lines 3 and 4 as shown in Figure 2. The expected travel time from O to D is 27.75 minutes (note that the headway fraction was assumed to be 0.5).

Consider now the addition of another line that requires a walk from node O to node E in order to access line 5 which was added to the network. This is shown in Figure 3.

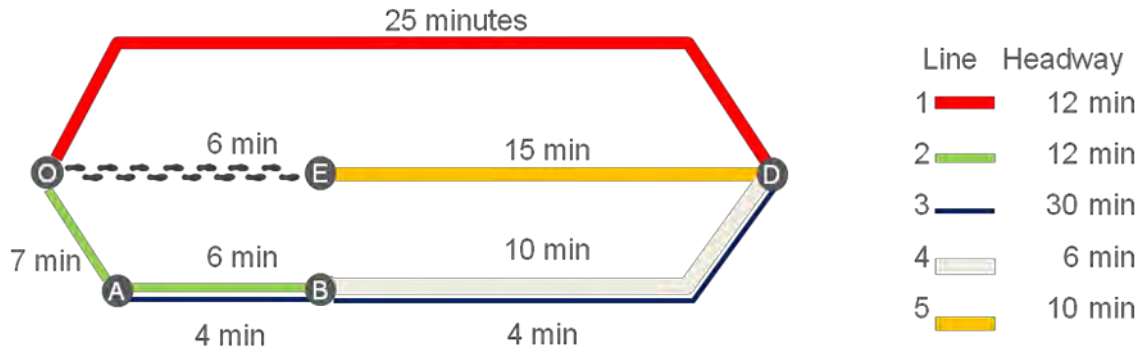


Figure 3. A 5 line example: adding a walk to line option.

The new line provides a total waiting and travel time of 26 minutes which is less than the 27.75 minutes of the optimal strategy before line 5 was added. Hence the new optimal strategy for all 100 trips would be to walk to node E, board line 5 and alight at node D. None of the lines used before line 5 was added are used now as indicated in Figure 3 below. All 100 trips will walk from node O to node E and then board line 5 to the destination D.

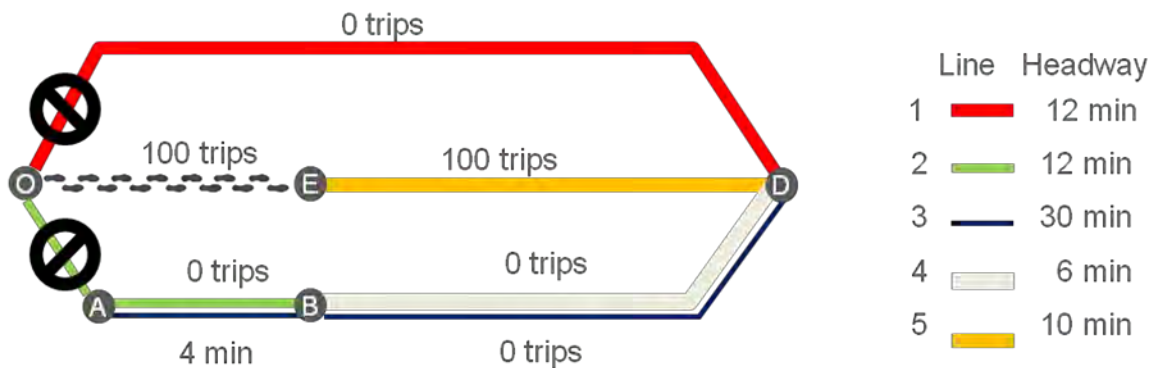


Figure 4. A 5 line example: walk to line strategy is optimal.

Now suppose that a logit function is applied to the expected waiting and travel times were to be used to make the choice between these two alternative strategies (one includes the walk link and line 5; the other includes all other lines). The probability (also referred to as the proportion) of the first strategy, which does not include line 5, when the scale parameter $\theta = .1$ would be:

$$\exp(-\theta * 27.75) / (\exp(-\theta * 26) + \exp(-\theta * 27.75)) = .458.$$

The resulting line segment flows are shown in Figure 5. The logit function choice is applied to the two strategies and for each strategy the flow subdivision on lines is done according to the optimal strategy. The new impedance is 26.80 since the resulting assignment is not optimal in the sense of least cost strategies. However, it may represent better the choices made by transit travelers.

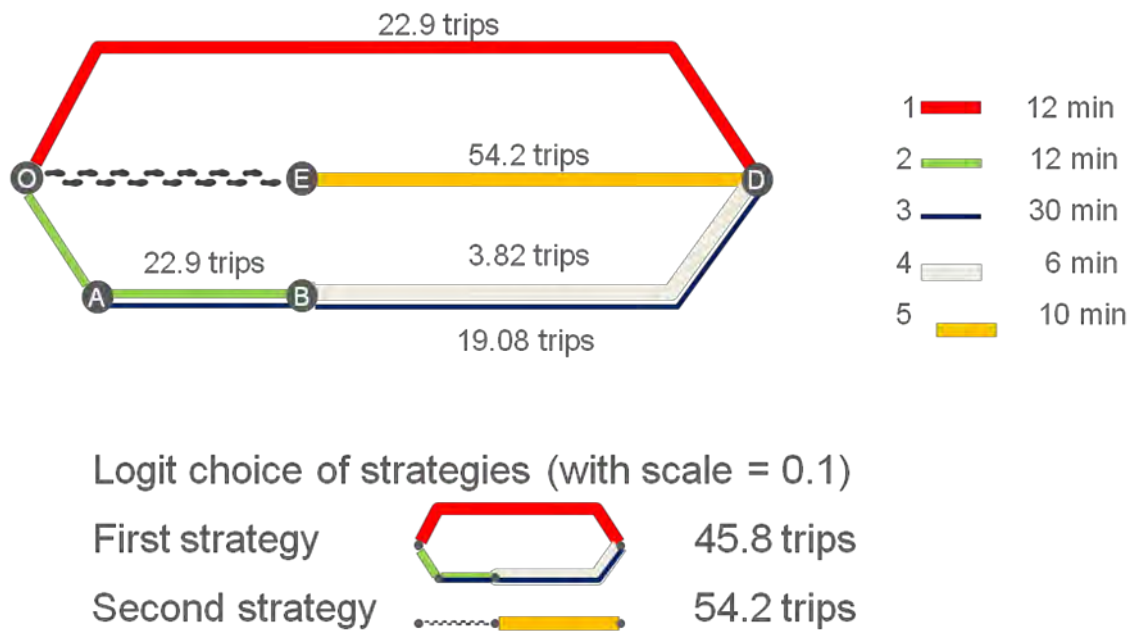


Figure 5. A 5 line example: logit choice of strategies.

It is worthwhile to note that, in iterative transit equilibrium assignments (see (1) and Cepeda et al (4)), where congestion in transit services is modeled, this extreme property of strategies would be attenuated: if the first strategy were to become congested then the walk option would be part of the optimal strategy; in an equilibrium solution both the attractive lines at O and the walk to transit options are likely to be used.

The theoretical work on which this document draws on is a paper by Nguyen et al (2) in which they develop an algorithm for a logit choice of strategies (hyperpaths) by adapting a basic idea due to Dial (3). The term *hyperpath* refers to an acyclic directed graph which is synonymous with a strategy. The difference between the model considered in (2) and this development is that *walk was not considered as an option* in the construction of the strategies. The walk mode was only considered on connector links from an origin node but not at nodes of the network where there is a choice between boarding a vehicle at a stop and walking to another stop to board a different line.

2. EFFICIENT ARCS

Clearly, one must first identify, in a very general way, the walk links and transit line segments that would be considered in an enlarged choice set of strategies. This leads to a modification of the algorithm that computes optimal strategies in order to compute simultaneously two optimal strategies:

- one that does not include any walk links from a node that is served by a transit line and
- one that includes walk links from a node that is served by transit line to another node which is served by a different transit line.

Suppose that these two optimal strategies have been computed and each node of the network has two labels, u_i^t, u_i^w which correspond to the optimal strategies without walk and with walk at a node i . These

labels contain the smallest expected waiting and travel time from the node to the destination without and with a walk to another node. The set of walk arcs and transit line segments considered may be further augmented by using a basic idea due to Dial (3) for implementing a logit route choice for road networks. One defines the “efficient” walk arcs (i, j) and “efficient” transit segments s obtained, after the computation of the optimal strategies as mentioned above, as follows:

- a walk arc is efficient if, by taking it, one gets nearer to the destination;
- a transit segment, not belonging to the optimal strategy, is efficient if, by boarding it, the alighting stop in the non-optimal strategy, is nearer to the destination.

The efficient walk arcs and transit segments are easily identified, after the computation of an optimal strategy, by using the computed node labels.

3. ATTRACTIVE LINES IN AN OPTIMAL STRATEGY

Before the development of the logit strategy model and algorithm, a review of some notation that refers to the transit network and to the computed quantities in a strategy is necessary.

The network considered $G = (N, A)$ consists of nodes N and arcs A . The transit line segments are represented as links; boarding and alighting arcs are used to represent activities at stops. Each arc $a \in A$ has two attributes (t_a, f_a) which are the travel time and the frequency of the arc. Figure 6 describes the activities at a stop:

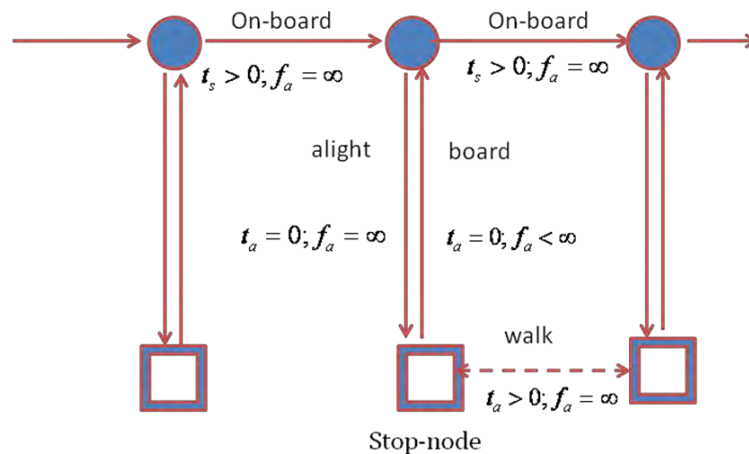


Figure 6. The representation of lines and stops.

It is convenient to refer to the "on-board" arcs as transit line segments $s, s \in S \subset A$, and to the walk arcs as $a^w \in A^w \in A$. Since the alighting and boarding arcs are handled implicitly in the solution algorithm, and do not require specific notation, a walk link a^w will be referred to simply as walk link a .

In an optimal strategy, at each stop that has multiple transit lines that can be boarded, an “attractive set” of transit lines is computed. This set of attractive lines is used by transit paths that lead from each node to the destination. Entire paths leading from the origin nodes to the destination can be recovered by analyzing the optimal strategy.

The frequency of a line l is denoted by f_l ; $t_{s(l)}$ is the travel time on the line segment and t_a is the walk time on link a . The optimal strategy is denoted \bar{A} and the set of attractive lines (or the stops of the attractive lines) at node i is denoted \bar{A}_i^+ . The set of walk links that are outgoing from node i are referred to as a , $a \in \bar{A}_i^+$. The combined frequency of the attractive lines is $f_i = \sum_{l(i) \in \bar{A}_i^+} f_{l(i)}$, where $l(i)$ denotes an outgoing line at node i ; or equivalently, $f_i = \sum_{l \in \bar{A}_i^+} f_l$ where \bar{A}_i^+ is the set of outgoing lines from node i in the optimal strategy \bar{A} from node i . The proportion of trips that leave from stop (boarding) node i on line $l \in \bar{A}_i^+$ is

$$p_l = f_l / \sum_{l \in \bar{A}_i^+} f_l, \quad l \in \bar{A}_i^+. \quad (1)$$

The average waiting time at stop (boarding) node i is $w_i = 1 / \sum_{l \in \bar{A}_i^+} f_l$. For any arc and walk arc (as well as access/egress arcs) a there is a travel time t_a and the flow proportion equals 1; for any continuation on the itinerary of a transit line there is no waiting time and the incurred segment travel time is t_s .

4. THE ALGORITHMIC APPROACH: LOG LIKELIHOODS AND NODE PROPORTIONS

A general description of the method is presented next. The notation E is used to denote the set of "efficient" walk links and transit line segments which are added to the optimal strategy during the computations. This larger set of transit line segments and walk links is denoted \bar{A} , $\bar{A} = \bar{A} \cup E$. First an optimal strategy is computed by using a modification of the algorithm given in the Appendix. It is useful to recall that, in an optimal strategy, a transit trip that boards a particular line at a stop alights at the node which is determined by the optimal strategy algorithm to be the best sub-path from the boarding.

The modification consists of computing two labels for each node: u_i^t is the least expected waiting and travel time to the destination by boarding a transit line at node i and u_i^w is the least expected waiting and travel time to the destination by using a walk link from the node, if possible. Then the walk links and transit line segments of the set E are identified by using the node labels and the sub-paths of the optimal strategy.

A walk link $a = (i, j)$ is "efficient" if

$$\min(u_i^t, u_i^w) > u_j^t, \quad a = (i, j). \quad (2)$$

A transit segment, $s = (i, j)$, which is *not part* of the optimal strategy, is "efficient" if

$$\min(u_i^t, u_i^w) > u_k^t, \quad s = (i, j) \quad (3)$$

for the strategy sub-path leading from node i to alighting node k for the transit segment.

The arcs used in the optimal strategy form an acyclic graph. The definition of efficient walk arcs and efficient transit line segments, which are added to the optimal strategy, maintains the topological order of node labels. As a consequence, the network defined by \vec{A} is acyclic as well.

Next the logit choice between transit and walk is considered. It is useful to recall that when using a logit path choice for highway networks, where the waiting times for vehicle services is not modeled, the flow assigned to each path k is:

$$pl_k = \exp(-\theta T_k) / \sum_{k' \in K_{od}} \exp(-\theta T_{k'}) \text{ for } od \in OD, \quad (4)$$

where pl_k is the logit proportion (probability), θ is a scale parameter, K_{od} is the set of paths for origin-destination pair od and T_k is the cost of path k . In all efficient algorithms that compute the logit path choice (see for instance Trahan (5)) an equivalent sequential form is used based on the conditional proportion:

$$pl_{(ai)} = \exp(-\theta(t_a + T_{K(j)})) / \sum_{(a') \in A_i^+} \exp(-\theta(t_{a'} + T_{K(j)})), \text{ all } a=(i, j) \in A_i^+$$

where $T_{K(j)}$ is defined as the log sum $T_{K(j)} = 1/\theta \ln W_j$, $W_j = \sum_{(j,k) \in A_j^+} \exp(-\theta c_{i,j}) W_k$, $t_{(i,j)} (= t_a)$ is

the cost of arc a and A_i^+ is the set of arcs outgoing from node i that satisfy $u_i > u_j$. In the case of road networks the two methods produce the same proportions. This is not necessarily the case for transit networks due to the waiting costs which depend on the destination. The sequential method is adapted for the logit choice of strategies since it does not require an explicit enumeration of all strategies and it can be integrated with the computation of the optimal strategies from destination to all origins in an efficient way. The difference lies in the way that common lines at a node are handled.

Once the set \vec{A} has been identified the computation of the logit proportions are carried out by computing node weights for all “efficient” walk links and transit line segments in order to compute the logit strategy choice. This set can be scanned in topological order due to the definition of efficient walk arcs and transit line segments that are added to the optimal strategy. The “efficiency” criterion maintains this topological order just as in Dial’s (1971) method. The set A_i^+ contains all the efficient outgoing transit line segments and walk links that are outgoing of node i .

The node weights W_i , or likelihoods, are used to compute recursively as indicated below:

- a) $W_i = 1$ if i is the destination;
 - b) $W_i = \exp(-\theta w_i) \prod_{k \in A_i^+} W_k^{(f_i/f_k)}$, if i is a boarding node;
- (5)

$$c) W_i^w = \sum_{j \in \vec{A}_i^+} \exp(-\theta t_a) W_j, \text{ if } u_i^w < \infty; \text{ otherwise } W_i^w = 0; a = (i, j);$$

$$d) W_i^t = \sum_{j \in \vec{A}_i^+} \exp(-\theta t_s) W_j, \text{ if } s = (i, j) \text{ is a transit line segment of the same line,}$$

where $(i, j) \mid j \in \vec{A}_i^+$ are the outgoing optimal strategy and “efficient” links from node i .

In general, a node may have two or more values for the travel times to the destination by “efficient” links; this would occur if a set of attractive lines and one or more walk links are possible options. While the presentation here is for one transit option and one walk option it may well occur that there will be more than one walk option at a node. In this case the logit splitting proportions may be determined at each node by the formulae:

$$a) pl_i^t = W_i^t / (W_i^t + W_i^w); \tag{6}$$

$$b) pl_i^w = W_i^w / (W_i^t + W_i^w).$$

The node proportions and the node weights can be computed simultaneously in a single scan of the network from destination to origin. The entire algorithm is stated next.

TRANSIT ASSIGNMENT: LOGIT CHOICE OF OPTIMAL STRATEGIES

Step 1: Compute the optimal strategies with binary choice (transit, walk) for one destination:

Apply the optimal strategy algorithm from destination to origins and find the best transit strategy from node i (if it exists) and the best sub strategy that uses a walk link from node i (if it exists) to obtain labels (u_i^t, u_i^w) . Note that some of these label values may be infinite.

Step 2: Identify and add efficient walk links:

While the network is scanned from destination to origins, after computing node labels u_i^t, u_i^w :

determine the efficient walk links and transit line segments by applying criteria (2) and (3);

all efficient walk arcs and transit line segments are added to the optimal strategy \vec{A} to obtain $\vec{\bar{A}}$.

Step 3: Compute node weights and node logit proportions:

For all nodes (stops) and links included in $\vec{\bar{A}}$, in topological order, compute the node weights W_i^t, W_i^w ;

if a node is reachable by transit and a walk link then a node is assigned two node likelihoods, one for transit and one for the walk link(s);

the node local logit splitting proportions are determined according to (6).

Note that the node weights and the node proportions can be computed simultaneously.

Step 4: Load the network:

Scan the network in reverse topological order of assigning labels to nodes, that is from origins to the destination.

The logit node splitting proportions are used to subdivide demand between transit services at a node and walk links. Then the trips assigned to transit lines are subdivided according to the optimal strategy.

The algorithm that computes optimal strategies is polynomial (see Spiess and Florian (1)). The definition of efficient walk arcs and efficient transit line segments, which are added to the optimal strategy, maintains the topological order of node labels. As a consequence, the entire algorithm has a polynomial complexity. The only additional computations are those that identify efficient arcs and transit line segments. The network considered is acyclic and the scan from origins to the destination, in topological order, is polynomial as well.

5. A NUMERICAL EXAMPLE

The numerical example presented in the first section of the paper is used to demonstrate the algorithm. It is the simple example presented at the beginning of the article. It has been constructed to show a simple example of the identification of the “efficient” segments and/or walk links and the application of the logit model to expand the choices at nodes where walk to a transit line is an attractive option. The network data was shown in Figure 3 but is repeated in Figure 7 below.

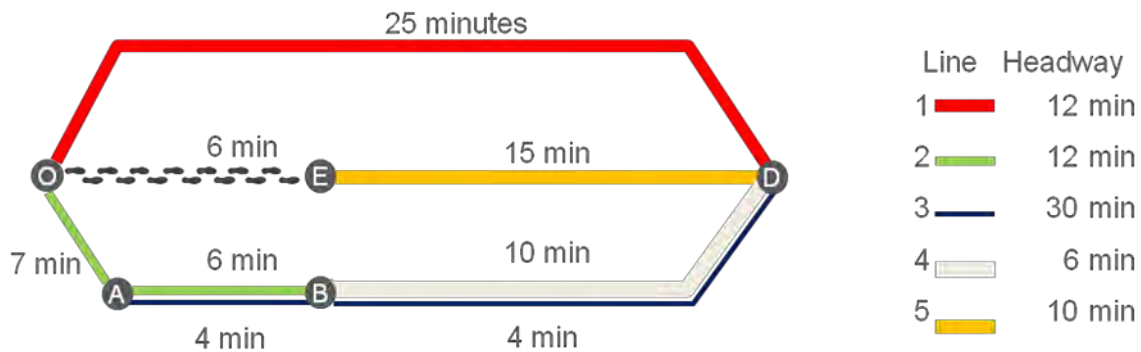


Figure 7. The 5 lines and walk link numerical example

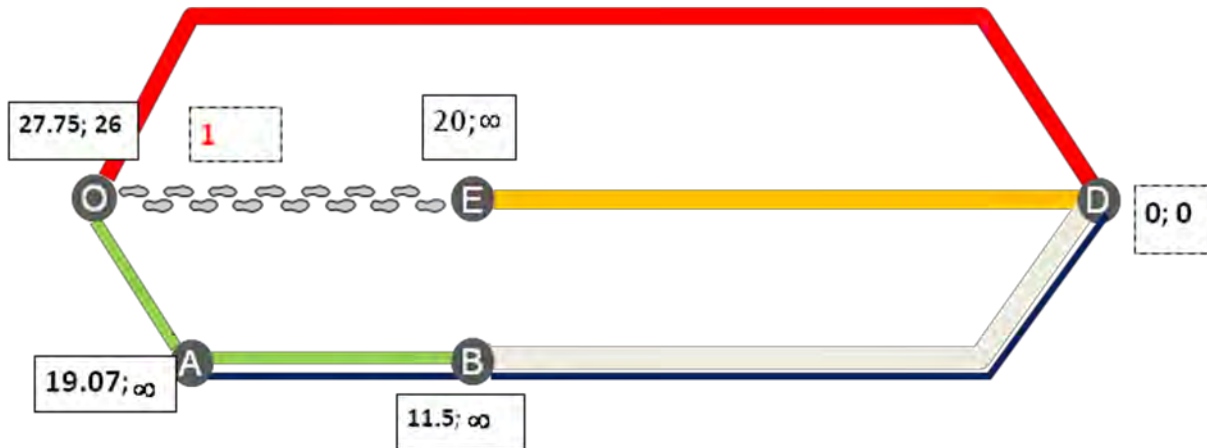


Figure 8. Computation of the optimal strategy

The labels u_i^t, u_i^w are indicated in the solid line boxes at nodes. The proportion of flow assigned to transit lines in the optimal strategy is indicated in the dashed boxes on each link. The only node where there is a choice between transit and walk is node O; in the optimal strategy, O->E->D, the walk option is selected. In this example there are no additional “efficient arcs” that are added. However, all the other lines are “efficient” transit segments.

Next, the computation of the logit assignment is done by using the node labels found when computing the optimal strategy. In this example all the links retained in the optimal strategy are “efficient”. Then, the node weights are computed in another scan of the network from destination to origins as follows:

$$W_D = 1;$$

$$W_B^t = (\exp(-.1 * 2.5) * (\exp(-.1 * 4))^{.167} * (\exp(-.1 * 10))^{.833}) = 779 * .935 * .435 = .318;$$

$$W_A^t = \exp(-.1 * 4.286) * (\exp(-.1 * 17.5))^{.714} * (\exp(-.1 * 8))^{.286} = .651 * .287 * .795 = .149;$$

$$W_E^t = \exp(-.1 * 5) * \exp(-.1 * 15) = .135;$$

$$W_O^w = \exp(-.1 * 6) * W_E = .074 * .135 = .074$$

$$W_O^t = \exp(-.1 * 3) * (\exp(-.1 * 25))^{.5} * \exp(-.1 * 24.5)^{.5} = .741 * .287 * .294 = .062$$

Transit lines segments (O,D) and (O,A) are “efficient” since

$$\min(27.75, 26.00) = 26 > 11.5 \text{ (alighting at B), and}$$

$$\min(27.75, 26.00) = 26 > 0 \text{ (alighting at D) respectively.}$$

As a consequence the strategy that they belong to is included in the flow computation.

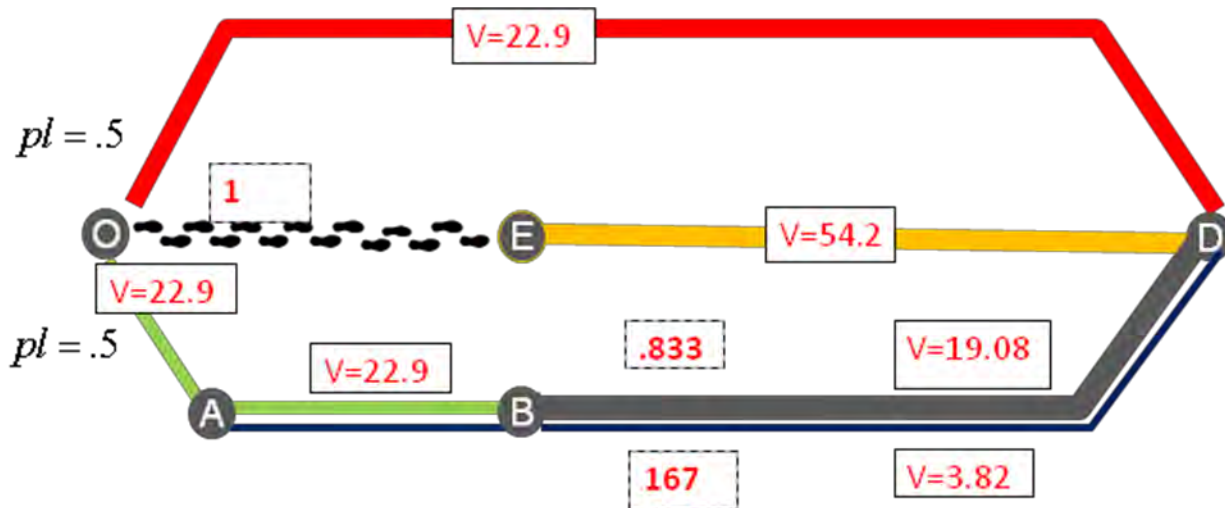


Figure 9. Logit choice of strategies; the flows are indicated in the heavy-edged boxes

It is relevant to note that the likelihood of node A was computed without the use of W_A since there is no alighting at node A.

The link proportions are computed by using the product of the node “logit proportions” and the “node “strategy proportions” computed for each strategy in reverse topological order from origins to destination:

At node O: Proportion “board line” = $W_o^l / (W_o^l + W_o^w) = .062 / (.074 + .062) = .458$; apply now the strategy proportions: the assigned trips to (O,D) line 1 are $45.8 * .5 = 22.9$ and the trips assigned to (O,A) line 2 are 22.9 as well; and proportion “walk” = .542; hence the assigned trips on (O, E) are 54.2;

At node E: The proportion of trips on (A, B) is 1.00 in the strategy; the number of trips is 54.2;

At node A; there are no transit boarding at this node.

At node B: The proportion of trips on (B,D) on line 3 is .167; apply now the strategy proportions; the number of trips on (B,D) of line 3 is 3.82 and the number of trips on (B,D) of line 4 is 19.08.

At a node that has both labels with finite values, the choice will be made with a logit function. This choice may use node specific values for the scale parameters of the logit function.

While the numerical example shows only a binary choice between “line” and “walk to line”, any number of walk links may be considered

6. PRACTICAL APPLICATIONS

The transit assignment algorithm described above has been implemented and used in several applications. One main observations made is, that in addition to the consideration of walk trips, there is a better distribution of the transit demand out of a centroid to the available connectors. The optimal strategy assignment would use only one connector which yielded the optimal strategy. As a consequence the distribution of transit demand on parallel service corridors is more even. As an illustration of such results,

Figures 10 and 11 show the differences in centroid connector and transit flows between the optimal strategy and the logit on strategies assignments.

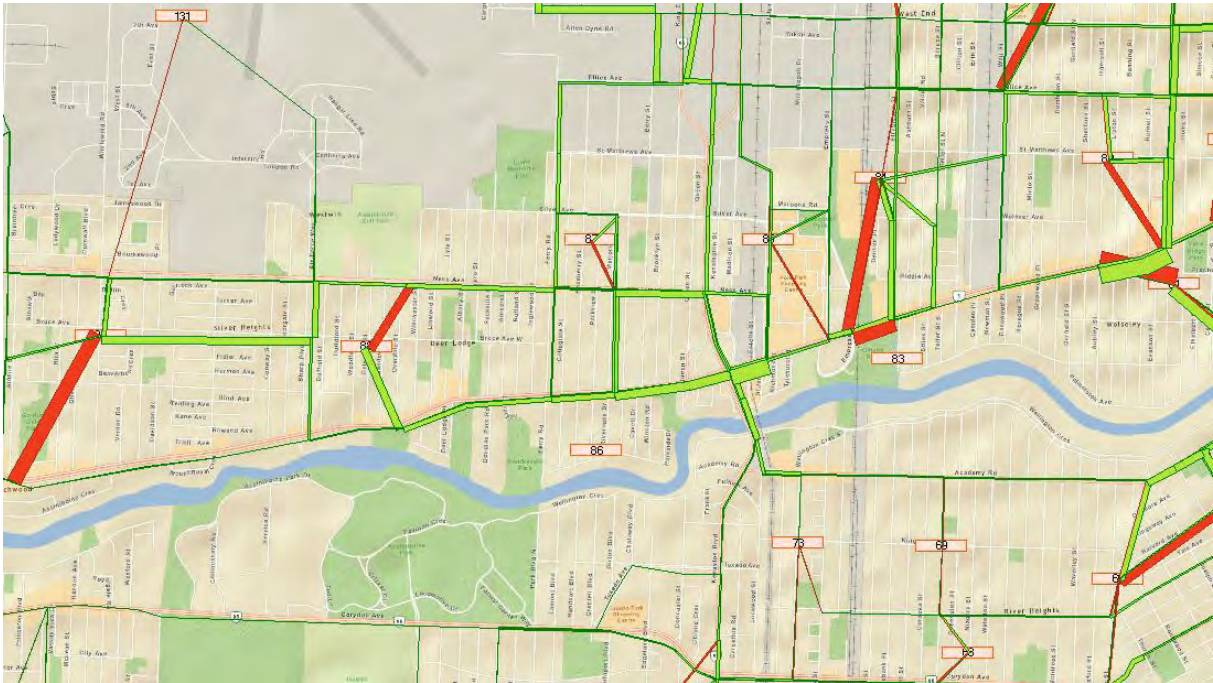


Figure 10. Differences in flows on centroid connectors

The logit on strategies assignment uses all the connectors from the centroids whereas the optimal strategy assignment uses only one connector. The connector links indicated in green are those that do not have any flow in the optimal strategy assignment.

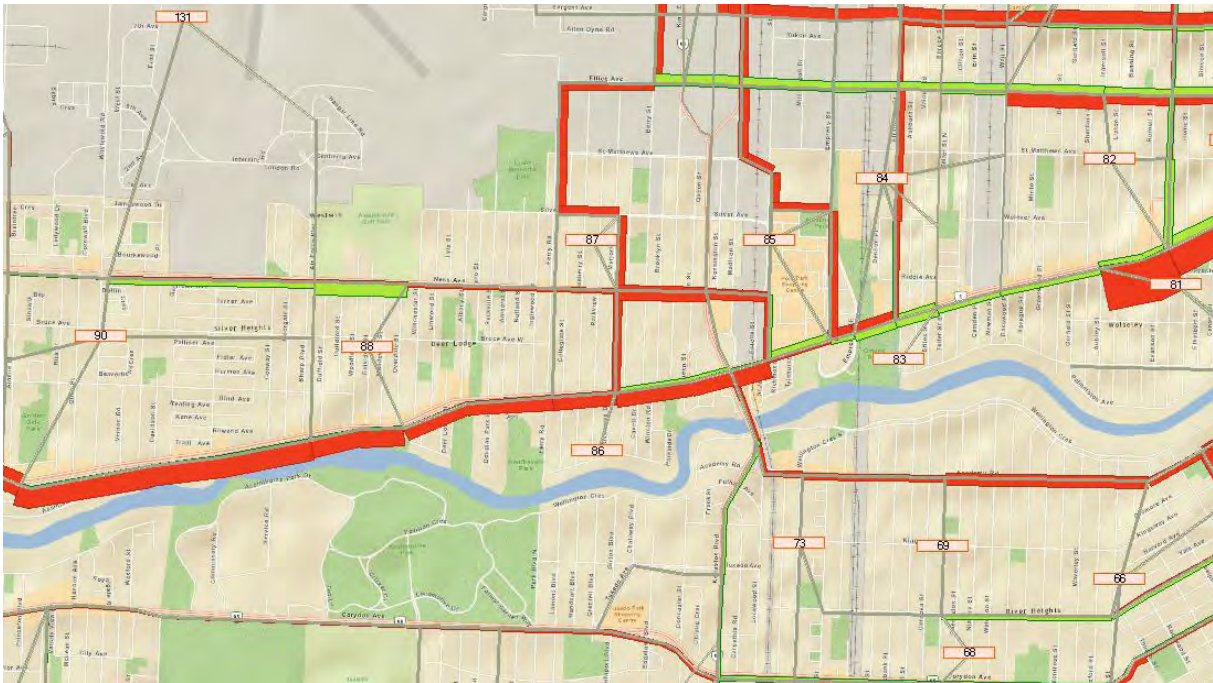


Figure 11. Differences in link flows on links

The flows indicated in green show the use of parallel transit corridors by the logit on strategies assignment.

7. CONCLUSIONS

This development opens up new possibilities for better representing the choices that travelers make in transit route choices. The issue of calibrating the scale parameter of the logit function is an issue which has not been addressed in this paper, but is relevant. Another aspect that must be considered is the possible consequence of the IIA property of the logit function. In the applications carried out so far no aberrant flows were observed. This is probably due to the fact that the logit function was used to choose among strategies. The selective application of the logit choice at certain nodes of the network, and not all, is advisable in practice.

8. APPENDIX: Algorithm for computing optimal strategies (from Spiess and Florian, (1))

Part 1: Find optimal strategy

- 1.1 (Initialization) $u_i = \infty, i \in N - \{d\}; u_d = 0;$
 $f_i = 0, i \in N;$
 $S = A; \bar{A} = \phi.$
- 1.2 (Get next link) If $S = \phi$ then STOP,
 Otherwise find $a = (i, j) \in S$ which satisfies
 $u_j + t_a \leq u_{j'} + t_{a'}, a' = (i', j') \in S;$
 $S = S - \{a\}.$
- 1.3 (Update node label) If $u_i \geq u_j + t_a$ then
 $u_i = f_i u_i + f_a (u_j + t_a) / (f_i + f_a)$
 $f_i = f_i + f_a, \bar{A} = \bar{A} + \{a\};$
 go to step 1.2

Part 2: Assign demand according to optimal strategy

- 2.1 (Initialization) $V_i = g_i, i \in N;$
- 2.2 (Loading) Do for every link $a \in A$, in decreasing order of $(u_j + t_a);$
 if $a \in \bar{A}$ then $v_a = V_i (f_a / f_i),$
 $V_j = V_j + v_a,$
 otherwise $v_a = 0.$

The variable V_i denotes the total flow that is leaving node i and g_i is the travel demand from node i to the destination. The destination node is d . The set S includes all the links at the beginning of the computations and the algorithm terminates when all the links were scanned.

9. REFERENCES

1. Spiess, H. and Florian M. (1989) Optimal Strategies: A New Assignment Model for Transit Networks. *Transportation Research B* 23(2), 83-102
2. Nguyen, S., Pallottino, S. and Gendreau M. (1998) Implicit Enumeration of Hyperpaths in a Logit Model for Transit Networks. *Transportation Science* 32(1), 54-64
3. Dial, R.B. (1971) A Probabilistic Multipath Traffic Assignment Model which Obviates Path Enumeration. *Transportation Research* 5, 88-111.
4. Cepeda, M., Cominetti, R. and Florian, M. (2006) A Frequency-based Assignment Model for Congested Transit Networks with Strict Capacity Constraints: Characterization and Computation of Equilibria. *Transportation Research B* 40, 437-459.
5. Trahan, M, (1974) Probabilistic Assignment: An Algorithm, *Transportation Science* 8(4), 311-320.

10. ACKNOWLEDGMENT

We would like to thank our colleague Calin Morosan for a careful reading of the paper.