A DISCRETE TRANSPORTATION NETWORK DESIGN PROBLEM WITH COMBINED TRIP DISTRIBUTION AND ASSIGNMENT

D. E. BOYCE and B. N. JANSON

Department of Civil Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801, U.S.A.

Abstract—Three network design problems are formulated for optimizing discrete additions to link capacities. Trip distribution and assignment are functions of travel expenditure in each problem. A reformulation of Wilson's doubly-constrained trip distribution problem, in which entropy instead of total travel expenditure is constrained, provides zone-to-zone flows in each problem. Entropy is interpreted as a measure of zone-to-zone trip-making interaction. Computational results for a toy network are reported.

INTRODUCTION

The transportation network design problem has many alternative formulations. The appropriate objective function and constraints depend on the particular application. This paper addresses the following problem which often arises in transportation planning: minimize total travel expenditure subject to a budget constraint on the total construction cost of link capacity changes. Changes to link capacity can be modelled with discrete or continuous functions (see Abdulaal and LeBlanc, 1979; Dantzig et al., 1979). Discrete functions are more appropriate for transportation networks since continuous functions permit the solution to include fractions of highway lanes or fractions of railway lines. Thus, the discrete network design problem can be formulated as follows:

min
$$C(N) = \sum_{i=1}^{n} \sum_{j=1}^{n} T_{ij}(N) c_{ij}(N)$$
 s.t. $K(N) \le K_0$

where $T_{ij}(N)$ = number of trips from zone i to zone j on network N; $c_{ij}(N)$ = minimum path travel cost from zone i to zone j over network N; N = network comprised of existing link capacities plus a subset of the alternative link capacity changes; n = number of zones; each zone is implicitly regarded as having a centroid node (e.g. city, warehouse) from which trips originate and to which trips are destined; K(N) = construction cost of network N; K_0 = budget constraint on construction cost.

From each alternative network, the objective function C(N) will depend on the distribution and assignment of trips. The trip distribution can be fixed or allowed to vary as a function of zone-to-zone travel costs. Likewise, the travel costs for each link can be fixed, or an increasing function of link flows. Boyce et al. (1973) formulate and solve a network design problem where both the trip distribution and link travel costs are fixed. This paper extends their work to cases in which the trip distribution and link travel costs are variable. The next section discusses a formulation of the trip distribution problem which is compatible with the above network design problem. The third section considers three assignment problems which, together with the trips distribution problem, result in three formulations of the network design problem.

THE TRIP DISTRIBUTION PROBLEM WITH CONSTRAINED ENTROPY

Wilson's doubly-constrained trip distribution model is a solution to the following problem: maximize the entropy of the trip distribution subject to constraints on zone origins, zone destinations and total travel expenditure (see Wilson, 1970). However, a trip distribution problem which constrains total travel expenditure cannot be used with a network design problem which seeks to minimize this value (see Boyce and Soberanes, 1979). Erlander (1977) suggests a reformulation of the trip distribution problem which obviates this conflict by interchanging the maximum entropy objective functions with the constraint on total travel expenditure.

Erlander's formulation of the doubly-constrained trip distribution problem is:

$$\min C = \sum_{i} \sum_{j} T_{ij} c_{ij} \tag{1}$$

s.t.
$$\sum_{i} T_{ij} = 0_i$$
, for all i (2)

$$\sum_{i} T_{ij} = D_{j}, \quad \text{for all } j$$
 (3)

$$H = -\sum_{i} \sum_{i} p_{ij} \ln (p_{ij}) \ge H_0, \quad \text{where} \quad p_{ij} = T_{ij} / \sum_{ij} T_{ij}$$
 (4)

$$T_{ij} \ge 0$$
, for all (i, j) (5)

where 0_i = number of trips leaving zone i; D_j = number of trips arriving at zone j; c_{ij} = minimum travel cost path from zone i to zone j; n = number of zones.

The solution to this minimization problem is:

$$T_{ij} = A_i O_i B_j D_j \exp(-\beta c_{ij})$$
 (6)

where

$$A_i = \left(\sum_j B_j D_j \exp\left(-\beta c_{ij}\right)\right)^{-1} \tag{7}$$

$$B_{j} = \left(\sum_{i} A_{i} O_{i} \exp\left(-\beta c_{ij}\right)\right)^{-1}.$$
 (8)

Equations (6)-(8) are identical to those derived by Wilson (1970). However, the travel deterrence parameter β is now the Lagrange multiplier for the entropy constraint instead of the travel expenditure constraint.

A network design problem, using this reformulation, can be stated as follows: find the network that minimizes total travel expenditure subject to constraints on origins, destinations, entropy and the total construction cost of link capacity changes. A later section of this paper examines what it means to hold entropy constant for different networks. This formulation is not meant to suggest that the entropy of a trip distribution in the "real world" necessarily remains constant following network changes. Likewise, this formulation does not suggest that tripmakers seek to minimize total travel expenditure subject to some observed level of entropy. This is simply a mathematical manipulation which permits the use of Wilson's model with the network design problem.

THE ASSIGNMENT PROBLEM—THREE ALTERNATIVES

The previous section discussed the trip distribution problem for fixed zone-to-zone travel costs with c_{ij} representing the minimum travel cost path from i to j. If travel costs are fixed, then one typically uses all-or-nothing assignment on the assumption that all trip-makers choose their minimum cost path. This assumption may be satisfactory for interurban railway or freeway networks. For an urban highway network where congestion does affect travel costs, the problems of distribution and assignment must be solved jointly.

Consider the assignment problem with variable travel costs and a fixed distribution of trips. Wardrop's principle of equal travel costs states that, at equilibrium, the average travel cost must be equal for all used paths from zone i to zone j and no unused path has a lower average cost (Wardrop, 1952). If travel costs are unaffected by congestion, then this principle is satisfied by assigning all trips to the minimum cost paths. If travel costs are affected by congestion, then the link flows and travel costs must satisfy the above condition. This is usually referred to as the user-equilibrium assignment since no trip-maker can decrease his or her cost by switching to a different path.

The equilibrium assignment problem (see LeBlanc et al., 1975; Nguyen, 1974) can be formulated as follows:

$$\min \sum_{a} \int_{0}^{v_a} S_a(x) \, \mathrm{d}x \tag{9}$$

s.t.
$$v_a = \sum_i \sum_k \sum_k \delta^{ak}_{ij} x^k_{ij}$$
 for all a (10)

$$\sum_{k} x_{ij}^{k} = T_{ij} \qquad \text{for all } (i, j)$$

$$v_{a} \ge 0 \qquad \text{for all } a \qquad (12)$$

$$v_a \ge 0$$
 for all a (12)

$$x_{ii}^k \ge 0$$
 for all (i, j) (13)

where x_{ij}^k = number of trips from zone i to j on path k; $\delta_{ij}^{ak} = 1$, if link a belongs to path k from zone i to j; v_a = total flow of trips on link a; $S_a(v_a)$ = average cost per trip on link a for flow v_a .

Equation (9) represents the sum of the areas under the link cost-flow functions. Equation (10) relates each link flow to the specified zone-to-zone trips. Equation (11) assures that the trips on each path from zone i to j sum to total trips from i to j.

A problem very similar to user-equilibrium assignment is system-optimal assignment. The system-optimal assignment problem is to minimize total travel cost (eqn 14) subject to the same constraints (eqns 10-13) which apply to the user-equilibrium problem.

$$\min C = \sum_{a} S_a(v_a) v_a. \tag{14}$$

In place of Wardrop's principle, the solution to the system-optimal assignment problem satisfies the condition that marginal travel costs are equal for all used paths from zone i to zone j and no unused path has a lower marginal cost (see LeBlanc and Morlok, 1976). If this principle is satisfied, no trip-maker can decrease his or her marginal cost by switching to a different path. There is no basis for assuming that individual trip-makers select their paths so as to minimize their own marginal cost. Given the same network and trip distribution, the total travel cost for a user-equilibrium assignment must be greater than or equal to the total travel cost for a system-optimal assignment. Thus, a system-optimal assignment provides a lower bound on total travel cost for the user-equilibrium assignment. This property is utilized below in the context of network design. It should be noted that in the case of fixed link travel costs, the all-or-nothing assignment to minimum cost paths satisfies both the user equilibrium and the system-optimal assignment conditions since no trip-maker can decrease his or her travel cost and all marginal costs equal zero.

THREE NETWORK DESIGN PROBLEMS

So far, the problems of network design, trip distribution and assignment have been discussed as separate entities. This section considers three network design problems which synthesize these separate formulations. In each problem, trip distribution is a function of zone-to-zone travel costs, a given value of entropy and fixed numbers of origins and destinations at each zone. Link travel costs are fixed in Problem 1, but are variable functions of link flows in Problems 2 and 3. Problem 2 incorporates user-equilibrium assignment while Problem 3 incorporates systemoptimal assignment.

Problem 1—Variable trip distribution with fixed link travel costs

Problem 1 consists of solving the discrete network design problem for the trip distribution given by eqns (1)-(5) for each alternative network. All trips from i to j are implicitly assigned to the minimum cost path.

Problem 2—Combined distribution and user-equilibrium assignment

Problem 2 consists of solving the discrete network design problem for the trip distribution and assignment given by eqns (9)–(13) and (2)–(5) for each alternative network. Augmenting the equilibrium assignment problem with constraints (2)–(5) allows the trip distribution (T_{ij}) to respond to network flow conditions, subject to origin, destination and entropy constraints. An efficient algorithm to solve the problem of combined distribution and assignment was designed by Evans (1976); see also Florian *et al.* (1975).

Since minimizing total travel cost is *not* the objective function of the combined distribution and user-equilibrium assignment problem, total travel cost could increase in response to an increase in the capacity of one or more links. Such a result is known as Braess' paradox (see Murchland, 1970; LeBlanc, 1975). As a consequence, Problem 2 is non-convex. The implications of non-convexity with respect to the network design algorithm will be discussed after Problem 3.

Problem 3—Combined distribution and system-optimal assignment

Problems 3 consists of solving the discrete network design problem for the trip distribution and assignment given by eqns (14), (10)-(13) and (2)-(5) for each alternative network. Evans' algorithm for combined distribution and user-equilibrium assignment can be easily converted to system-optimal assignment.

NON-CONVEXITY AND THE NETWORK DESIGN ALGORITHM

Boyce et al. (1973, 1974) formulate and test a tree-search algorithm which finds the optimal subset of link additions for a fixed trip distribution and fixed link travel costs. This paper extends the use of this algorithm to cases where the trip distribution and link travel costs are variable. The search procedure is based on the algorithm of Beale et al. (1967), and is suited to examine link additions, link deletions or any discrete adjustments to link capacity. The examples in this paper, like those of Boyce et al. (1973), consider only the case of link additions.

Since the structure of the algorithm is fully described by the above references, it is sufficient here to examine the way in which non-convexity affects the algorithm. If the network design problem is convex, then the algorithm can use certain bounding rules which obviate the need to evaluate all possible subsets of network changes. The algorithm makes repeated use of the following property defined for networks N_1 and N_2 :

if
$$N_1 \supset N_2$$
, then $C(N_1) \leq C(N_2)$.

At the outset of the search, the algorithm computes an unconditional threshold $U(\lambda)$ for each link capacity increase. $U(\lambda)$ equals the value of the network design objective function, i.e. total travel cost, for the network which includes all possible link capacity increases except λ . The unconditional threshold is used at each iteration of the search to test whether any link capacity increase must be included in the optimal solution. If at any stage in the search, $U(\lambda) > C^*$, where C^* is the lowest value of total travel cost found thus far, then all networks which do not include λ can be eliminated from the search.

It should be noted that this "threshold" rule is valid only if the network design objective function responds monotonically to any change in the capacity of one or more links. For a network design problem which seeks to minimize total travel cost, this property of "monotonicity" can be stated as follows: for a given network, any increase in the capacity of one or more links will not increase total travel cost and any decrease in the capacity of one or more links will not decrease total travel cost. If capacity increases on some links are performed simultaneously with capacity decreases on other links, then the objective function may vary in either direction regardless of whether the sum of the capacities for all links increases or decreases. The network design problem is "non-convex" if this property of monotonicity is not satisfied. As stated earlier, this property is not satisfied by Problem 2 since total travel cost may increase if the capacity of one or more links is increased.

Although the unconditional threshold does not provide a strong bound on the objective function, computational experience for simpler versions of the network design problem, and for certain statistical problems, has shown that this bound can be highly effective. One measure of the computational difficulty of a network design problem is the number of links whose unconditional thresholds exceed the objective function for the optimal network. As this number approaches the number of links in the solution, the size of the search is sharply reduced. Moreover, a related bound, called the *conditional threshold*, can be defined for specific subsets of links which further limits the search. Details may be found in Boyce *et al.* (1973, 1974) and Beale *et al.* (1967).

Since the network design problem with user-equilibrium assignment (Problem 2) is a non-covex problem, a branch-and-bound procedure similar to that of LeBlanc (1975) can be used to determine the optimal network. For a given network and level of entropy H_0 , total travel cost obtained for a system-optimal assignment is a lower bound on total travel cost for a user-equilibrium assignment. In LeBlanc's procedure, the system-optimal lower bound is used to eliminate certain sub-optimal networks in Problem 2. In general, the solution to Problem 2 may include a very different set of network changes than the solution to Problem 3. The computational examples presented next provide a case where the two network solutions are quite different.

SOLUTIONS TO PROBLEMS 1, 2 AND 3 FOR A TOY NETWORK

To explore the properties of Problems 1, 2 and 3, solutions were computed for the ten node Network 1 used by Boyce *et al.* (1973). To simplify calculations and reduce computational effort, eleven links were designated as the "existing network". These links consisted of the minimum spanning tree plus two additional links. Origin and destination totals were specified for each node i such that $O_i = D_i$.

From the remaining 34 links, nine links were selected as candidates for addition to the existing network. The construction cost of each link was defined as its length. The total length of the nine links was approximately 4000. Two budget constraints were specified equal to 30 and 60 per cent of total cost, i.e. 1200 and 2400.

The results for the three problems are shown in Table 1. An initial value of β (0.090) was selected, and Problem 1 was solved on the existing network to determine the corresponding value of entropy (3.91782). Then, throughout the network algorithm and for both budget constraints, the trip distribution problem was solved for each network with $H_0 = 3.91782$. The iterative solution procedure which calibrates the trip distribution model was halted when the value of entropy for (T_{ij}) was within 0.3 per cent of H_0 .

For Problem 1, the network design algorithm chose three links for the first budget constraint and six links for the second. The second solution includes the three links from the first; two of

Problem	Budget constraint	Construction cost	User cost (obj. fn.)	Interaction (entropy)	β Value	Links added†	Number of iterations‡
Fixed costs	existing 1200	-	10,764	3.91782	0.0900		-
	2400	1166 2278	10,248 10,217	3.90734 3.91680	0.0978 0,1005	1, 4; 1, 7; 1, 10. 1, 3; 1, 4; 1, 5; 1, 7; 1, 10; 8, 9.	1 8
User equilibrium variable costs	existing 1200	1007	17,172 12,474	3.71326 3.71292	0.0900 0.1350	1, 4; 1, 10; 8, 9.	
	2400	2157	11,857	3.71104	0.1595	$\overline{1,3}; \overline{1,4}; 1,5; \\ \underline{1,10}; 6,9; 8,9.$	8
System optimal variable costs	existing 1200	1152	15,939 12,036	3.71575 3.71219	0.0494 0.1206	1, 4; 1, 5; 1,10	
	2400	2324	11,337	3.71276	0.1422		5

Table 1. Optimal link additions to a ten node, eleven link network for three design problems

[†]Underlined links passed the unconditional threshold.

[‡]An iteration of the design algorithm is defined in Boyce et al. (1974).

these links pass the unconditional threshold which reduces the computational effort very substantially.

To obtain comparable solutions to Problems 2 and 3, Problem 2 was solved with a value of β (0.090) on the existing network and H_0 (3.71326) was calculated. This value of H_0 was then used for both Problems 2 and 3. For Problem 2, three links were selected for the first budget constraint and six for the second. Problem 3 resulted in the choice of three and five links respectively for the two budget constraints. Links (1,4) and (1,10) were selected in all six solutions, and in each case, passed the unconditional threshold. Otherwise, the links chosen were rather different.

The computations performed by the algorithm were examined to determine if the monotonicity property was ever viaolated in Problem 2. For the budget constraint of 1200, 21 networks were solved in which one link was added to the existing network or to the existing network plus one or two links. The addition of a link (6, 9) to a network of 12 links resulted in an *increase* in the objective function; all other increases in link capacity substantially decreased the objective function. For the second budget constraint of 2400, 30 networks were solved, but no increase in the objective function occurred. The violation of monotonicity in the case of link (6, 9) means that the optimal solutions shown could be local optima.

INTERPRETATION OF THE ENTROPY CONSTRAINT

As noted near the outset of this paper, the entropy constrained formulation of the trip distribution and assignment problems was adopted primarily for reasons of mathematical compatibility with network design objective function. What remains is to interpret the use of entropy as a constraint in this class of problems. One interpretation is to view entropy as a measure of zone-to-zone "interaction" based on Erlander (1977). A second interpretation is to view entropy as the likelihood of the predicted trip distribution based on Wilson (1970). These interpretations will be discussed next.

Erlander (1977) describes entropy "as a measure of the spread of the distribution of journeys over the cells of the trip matrix." In that article, Erlander referred to entropy as a measure of accessibility, but this implies that the formula for entropy includes travel costs. In their later article, Erlander and Stewart (1978) refer to entropy as a measure of zone-to zone "interactivity". The term "interaction" or "level of interaction" would seem to be a more appropriate expression to describe this aggregate measure of trip-making activity. Wilson has also referred to his broad class of models involving zone-to-zone flows as "interaction" models. Thus, two trip distributions with the same entropy can be viewed as having the same level of zone-to-zone interaction.

For a given set of row and column totals, the graph of entropy versus total travel cost has a slope equal to β and resembles a parabola which is concave from below (see Erlander and Stewart, 1978). This is expected since, in Wilson's formulation, β is the Lagrange multiplier for total travel cost. Since this curve has its maximum at $\beta = 0$, the distribution with the largest value of entropy is the one whose entries are proportional only to these row and column totals. This distribution is written as:

$$T_{ij} = \frac{O_i D_j}{\sum_{ij} \sum_{i} T_{ij}}$$
 for all (i, j) pairs.

This distribution corresponds to the case in which travel costs have no bearing on the behavior of trip-makers.

The entropy of a distribution which is affected by travel costs, i.e. $\beta \neq 0$, is lower than the entropy given by the above distribution. The difference in these values of entropy can be viewed as the degree to which travel behavior responds to zone-to-zone travel costs for a given network. For different networks, trip distributions with the same entropy have the same level of trip-making interaction but in general have different total travel expenditures. Thus, the network which affords the lowest travel expenditure for a given level of interaction and network investment can in one sense be considered to be the best choice. Then, for each level

of interaction, the trade-off between total travel expenditure and network investment can be examined by varying the budget constraint and solving for the best network. Finally, one could vary interaction to determine whether the same network tends to be optimal for a given investment budget and different levels of interaction as the findings of Boyce and Soberanes (1979) suggest.

Neither the interaction-constrained model employed here, nor Wilson's travel cost-constrained model, confront the basic question of whether interaction or total travel cost observed in an actual transportation system are independent of network structure. If these constraints are independent of network structure, then these models should provide useful results for policy making. However, if interaction or total travel cost are functions of the network under consideration, then one must be very cautious when interpreting the results.

It is also tempting to interpret the entropy constraint as meaning that the predicted trip distributions for different networks are equally likely. However, two distributions are equally likely only if their prior probability distributions are uniform. Hence, the network design problems presented here are not comparing trip distributions which are equally likely for their respective networks. Viewing the entropy constraint in a likelihood context can easily lead to this misinterpretation.

CONCLUSION

Three network design problems incorporating variable trip distributions with both fixed and variable link costs have been presented. Other formulations which include variable demand but do not require entropy to be constrained can also be considered. For instance, the cost deterrence parameter β could be held constant for each alternative network. Unfortunately, there is no reason to expect that β remains constant as the network varies, just as there is no reason to expects that entropy or total travel expenditure remains constant. Hence, a problem with β exogenously defined offers no improvement to an entropy-constrained problem.

If trip-makers are hypothesized to expend a given amount on travel regardless of network changes, then other network design objective functions which do not attempt to minimize this value might be considered. One objective might be to maximize consumer surplus, a measure which Williams (1976) and others have suggested to evaluate the benefits of transport systems. Another objective might be to maximize the level of interaction subject to a travel expenditure constraint. This approach would incorporate Wilson's original formulation to estimate the trip distribution for each network.

Finally, it might be the objective of network design to minimize total travel expenditure so long as the existing trip distribution is not altered beyond a given bound. The trip distribution problem could be formulated to minimize total travel cost subject to a constraint on the information divergence between the observed and predicted trip distribution. Kadas and Klafszky (1976) show that the maximum entropy trip distribution for the observed total travel cost also minimizes the information divergence of the observed distribution from the predicted distribution. Thus, it might be useful to constrain the observed-predicted information divergence and compare the minimum total travel cost for each network.

Unfortunately, these alternative formulations do not eliminate the problem of non-convexity which arises with user-equilibrium assignment. The characteristics of the actual network and link capacity changes must be examined in order to estimate the degree to which violations of convexity may produce sub-optimal networks. In other words, in actual computational situations, how often does the total travel cost increase when the capacity of a link or links is increased. It is not known to what extent, if ever, this problem occurs in large networks and in real-world transportation systems. These questions need to be examined further.

Acknowledgements—We wish to thank the University of Illinois at Urbana-Champaign and the School of Engineering and Applied Sciences of the University of Pennsylvania for the use of their computational facilities. We are grateful for comments by Sven Erlander, Michael Florian, Marc Los and the detailed comments of a referee. Remaining errors and omissions are, of course, our responsibility.

REFERENCES

Abdulaal M. and LeBlanc L. J. (1979) Continuous equilibrium network design models. *Transpn Res.* 13B, 19-32. Beale E. M. L., Kendall M. G. and Mann D. W. (1967) The discarding of variables in multivariate analysis. *Biometrika* 54, 357-366.

Boyce D. E., Farhi A. and Weischedel R. (1973) Optimal network problem: a branch-and-bound algorithm. *Environm. and Plan.* 5, 519-533.

Boyce D. E., Farhi A. and Weischedel R. (1974) Optimal Subset Selection: Multiple Regression, Interdependence and Optimal Network Algorithms. Lecture Notes in Economics and Mathematical Systems, No. 103. Springer-Verlag, Heidelberg.

Boyce D. E. and Soberanes J. (1979) Solutions to the optimal network design problem with shipments related to transportation cost. *Transpn Res.* 13B, 65-80.

Dantzig G. B., Harvey R. P., Lansdowne Z. F., Maier S. F. and Robinson D. W. (1979) Formulating and solving the network design problem by decomposition. *Transpn Res.* 13B, 5-17.

Erlander S. (1977) Accessibility, entropy and the distribution and assignment of traffic. Transpn Res. 12, 291-293.

Evans S. P. (1976) Derivation and analysis of some models for combining trip distribution and assignment. *Transpn Res.* 10, 37-57.

Florian M., Nguyen S. and Ferland J. (1975) On the combined distribution-assignment. Transpn Sci. 9, 43-53.

Kadas, S. A. and Klafszky E. (1976) Estimation of the parameters in the gravity model for trip distribution: a new method and solution algorithm. Regional Sci. and Urban Economics 6, 439-457.

LeBlanc L. J. (1975) An algorithm for the discrete network design problem. Transpn Sci. 9, 183-199.

LeBlanc L. J. and Morlok E. K. (1976) An analysis and comparison of behavioral assumptions in traffic assignment. In *Traffic Equilibrium Methods* (Edited by M. Florian). Springer-Verlag, Heidelberg.

LeBlanc L. J., Morlok E. K. Pierskalla W. P. (1975) An efficient approach to solving the road network equilibrium traffic assignment problem. *Transpn Res.* 9, 309-318.

Murchland J. D. (1970) Braess' paradox of traffic flow. Transpn Res. 4, 391-394.

Nguyen S. (1974) An algorithm for the traffic assignment problem. Transpn Sci. 8, 209-216.

Wardrop J. G. (1952) Some theoretical aspects of road traffic research. Proc. Institution of Civil Engineering, Part II Vol. 1, 325-378.

Williams H. C. W. L. (1976) Travel demand models, duality relations and user benefit analysis. J. Regional Sci. 16, 147-166. Wilson A. G. (1970) Entropy in Urban and Regional Modelling. Pion, London.