Issues in Travel Demand Forecasting

Huey-Kuo Chen

Abstract—Travel demand forecasting including four travel choices, i.e., trip generation, trip distribution, modal split and traffic assignment constructs the core of transportation planning. In its current application, travel demand forecasting has associated with three important issues, i.e., interface inconsistencies among four travel choices, inefficiency of commonly used solution algorithms, and undesirable multiple path solutions. In this paper, each of the three issues is extensively elaborated. An ideal unified framework for the combined model consisting of the four travel choices and variable demand functions is also suggested. Then, a few remarks are provided in the end of the paper.

Keywords—Travel choices, B algorithm, entropy maximization, dynamic traffic assignment.

I. INTRODUCTION

TRAVEL demand forecasting, the core of transportation planning, has been extensively studied for more than 60 years. A common practice adopts the so-called sequential demand forecasting procedure in which trip generation, trip distribution, modal split and traffic assignment are treated in a top-down sequential manner and sometimes with feedbacks [1]. Although this approach has been widely used, indeed there exist some drawbacks need to be improved. In the following sections, issues of interface inconsistencies among the sequential travel choices, inefficiency of commonly used solution algorithms, and undesired multiple path solutions, will be elaborately discussed in order.

II. INTERFACE INCONSISTENCIES AMONG SEQUENTIAL TRAVEL CHOICES

Four travel decisions, i.e., trip generation, trip distribution, modal split and traffic assignment must be dealt with in travel demand forecasting which in turn constructs the essential component in transportation planning. When a sequential travel demand forecasting procedure is adopted, the output from one travel decision would naturally become the input of its lower level travel choice submodule. For example, in the first submodule of trip generation, the inputs are travel times and the output is the number of generated trips, which in turn become the inputs to the second submodule of trip distribution. Similar input-output relationship applies to the last two submodules of modal split, and traffic assignment as well. When the last submodule, traffic assignment, has been successfully performed, "equilibrated" travel times would be yielded as the output, which unfortunately are hardly to be consistent with the previous inputs to the first submodule of

H. K. Chen is with the National Central University, Jhong-Li, Taoyuan 80305 Taiwan (phone: +866-3-4227151 ext 34115; fax: +866-3-4252960; e-mail: ncutone@ncu.edu.tw).

trip generation, resulting in a phenomenon of internal inconsistencies in the sequential travel demand forecasting procedure. This inconsistent problem affects the precision of the solution to a great extent, and sometimes may even diverge.

To cope with the internal inconsistency problems, two strategies have been proposed in the past. The first strategy incorporates feedback steps into the sequential demand forecasting procedure with a hope that the final result will gradually converge from iteration to iteration. Bar-Gera and Boyce [2] have done extensive experiments on the sequential traffic demand forecasting with feedbacks. Their results showed that a pair of constant weights around (0.25, 0.75) between two consecutive iterations performs better than any other pairs of weight combinations and should be used for the future applications.

The second strategy adopts unified framework to treat two or more travel choices simultaneously, resulting in the combined models. In detail, centered on the traffic assignment, one or more other travel choices can be incorporated to form a combined model. The advantage of this approach is that the internal inconsistencies occurred between different submodules vanishes. However, the disadvantage is that the more submodels combined together, the higher difficulty would be encountered in both model formulation and problem solving. The earliest combined model appeared in the literature is the trip distribution and assignment problem [3]. Two algorithms, i.e., linearization method and partial linearization methods, are proposed for the optimal solution. According to the experimental results, the partial linearization method (aka Evans algorithm) performs better in terms of computation time. Unfortunately, even with the partial linearization method, its converging speed and degree of precision are not satisfactory for large scale problems. Other two dimensional [4]-[5] or three dimensional [6] combined models can be formulated by analogy. Recently, the most complicated combined model consisting of the four travel choices as well as variable demand functions has been formulated and can be regarded, by way of a suitable supernetwork representation, as an extended traffic assignment problem which can be easily solved by a nested solution algorithm embedding the FW algorithm [7]. Notice that FW algorithm is easy to understand but is relatively inefficient due to undesired zigzagging converging behavior. Therefore more efficient solution algorithms must be developed for practical applications. In a recent research, the above nested solution algorithm embedding B algorithm, instead of FW algorithm, has been applied to several larger networks and got satisfactory results [8].

III. Inefficiency Of Commonly Used Solution Algorithm

The traffic assignment problem can be formulated as a quadratic optimization problem in which the objective function is nonlinear and the associated constraints are linear [9]. In the past, the traffic assignment problem is commonly solved using the Frank-and-Wolfe algorithm (FW) [10], also known as convex combinations method. Unfortunately, the FW method is slow and indeed hard to attain the required precision due to undesired zigzagging converging behavior. The results obtained are often imprecise especially for large networks and sometimes may lead to incorrect decisions.

To alleviate the zigzagging converging behavior of the FW method, a revision called the parallel tangent method (PARTAN) was developed. In addition, many path-based solution algorithms like gradient projection (GP) [11], disaggregate simplicial decomposition (DSD) [12], and aggregate simplicial decomposition (ASD) [13], have also been proposed. Though sufficient improvement over the FW method can be observed, however, these new methods are still inadequate in attaining sufficient precise solutions especially for large networks. In addition, for path-based solution algorithms the memory requirement for large networks can be large, which is of course prohibitive in reality. Therefore a new type of quick and precise solution algorithms, which take advantage of tree-like network structure, must be developed.

In the past two decades, several quick and precise solution algorithms such as traffic assignment by paired alternative segments (TAPAS) [14], B [15], linear user cost equilibrium (LUCE) [16], origin-based algorithm (OBA) [17] and projected gradient (PG) [18] algorithms have been developed and indeed are satisfactory in solving large scale traffic assignment problems in terms of computation time and memory requirement. One immediate question is raised: which of these quick and precise solution algorithms is most useful and should be chosen for (extended) traffic assignment problems? This question leads us to the discussion of another issue: undesired multiple path solutions.

IV. Undesired Multiple Path Solutions

As aforementioned, the traffic assignment problem, or more generally the extended traffic assignment problem, can be formulated as a quadratic problem to which under certain regularity conditions a unique link-flow solution can be obtained. Unfortunately a unique link-flow pattern does not necessarily imply a unique path-flow pattern. According to the definitional relationship between link flows and path flows, a unique link flow pattern may associate with multiple path-flow solutions.

A unique path-flow solution, rather than multiple path-flow solutions, is indeed desired in transportation applications. Two possible applications pertaining to the unique path-flow solution are briefly discussed in the following. First, for travelers path information is more useful than link information in vehicle route guidance. In view of today's advancement in the intelligent transportation systems (ITS), especially in the

advanced transportation information systems (ATIS), on-line personalized and customized traffic information such as departure time and route choice is definitely needed. Without having unique path-flow solutions on hand, vehicle route guidance is hard to be realized. The second application has something to do with the network design problem. The most well-known network design problems in transportation are signal optimization and traffic assignment problem and origin-destination matrix estimation problem, both of which can be formulated as bi-level model. In the past, these two types of network design problems were generally solved by employing the implicit function theorem. This theorem, by way of variable perturbation [19], requires a path-flow solution in exploring the descent direction. As aforementioned, a unique link flow solution may associate with multiple path-flow solutions and different path-flow solution may in fact result in different converging behavior and, hence, different solutions. Therefore even if we can obtain a path flow solution using a path-based solution algorithm such as GP, how do we know the one obtained is the most reasonable path-flow pattern? This question has been resolved by a novel solution algorithm called TAPAS [14]. By introducing a flow proportionality principle, which is essentially a reminiscence of entropy maximization [2], [20], a unique path-flow solution can be obtained. The implication of flow proportionality is "no route will be left behind." Of course, this property does not mean that every route should be used, since under the user equilibrium (UE) assumption [21] only least cost routes are used. So the assumption of proportionality requires that "no route should remain unused, unless there is a good reason." A formal definition of proportionality is discussed in: [14], [20], [22]. Surprisingly, for any network configuration the path-flow solution obtained by TAPAS is unique.

Note that TAPAS is in the category of quick and precise solution algorithms and its computational efficiency has been proved and acknowledged in the literature [23]. Similar evidence can also be observed in a recent study on solving the doubly constrained distribution/assignment problem by TAPAS [8]. The superiority of TAPAS justifies its potential in solving large scale network problems.

V. An Ideal Framework for the Four-Step Combined Model with Variable Demands and Associated Solution Algorithm

So far we have elaborately discussed three important issues, i.e., interface inconsistencies, inefficiency of commonly used solution algorithms, and undesired multiple path solutions, in travel demand forecasting. Considering all these three issues together, one may be curious about what would be the ideal framework for the four-step combined model with variable demands and what would be the suitable solution algorithm? In the author's opinion, the ideal framework and suitable solution algorithm should conform to the following two requirements:

1) The combined model approach is certainly superior to the corresponding sequential procedure for treating different travel choices, as internal inconsistency problem can be

avoided. Moreover, the combined model, via the so-called supernetwork representations [24], can be regarded as an extended traffic assignment problem, which is conceptually as easy as the traditional traffic assignment problem. As a consequence, any traffic assignment algorithm can be adopted for the combined models.

2) Suitable solution algorithms must be able to produce unique path-flow solution, in addition to quick, precise, mild memory requirement. This is because unique path-flow solution is critical in providing useful traffic information in vehicle route guidance and good search direction for solving the network design problem.

Taking these two essential requirements together, the reasonable choice among all the available solution algorithms for solving extended traffic assignment problems (resulted from the corresponding combined models) would be a nested solution algorithm embedding TAPAS. To illustrate, we take as an example the combined four-step choice model with variable demands (hereafter combined TO/TD/MC/TA/VD problem which is acronym for trip origin, trip distribution, mode choice, traffic assignment and variable demand). According to [7], [25], the corresponding combined model can be formulated as a quadratic optimization problem (cf. Appendix I) in which the objective function and constraints must take all four travel choice dimensions into consideration. This combined model is definitely more difficult than the traditional traffic assignment. Fortunately, it is not as difficult as it looks like. In fact, it can be regarded as an extended traffic assignment problem via the supernetwork representation, as follows.

r, r': origin; s, s': destination; m: mode; M: attractiveness; θ, γ, η: dispersion parameter
1, 2, 1', 2': origin; 3, 4, 5, 3', 4', 5': destination; O: superorigin; O': superdestination
TA: traffic assignment; MC: mode choice; TD: trip distribution; TG: trip generation;
VD: variable demand
$c^{Or} = 0$; $c_{Transit}^{rs} = u_{Transit}^{rs}$; $c^{sm} = 0$; $c^{ms'} = \frac{1}{\theta}\ln q_m^s - M_m$; $c^{s'r'} = \frac{1}{\gamma}\ln q^{rs} - M^s$; $c^{r'O'} = \frac{1}{\eta}\ln q^r$
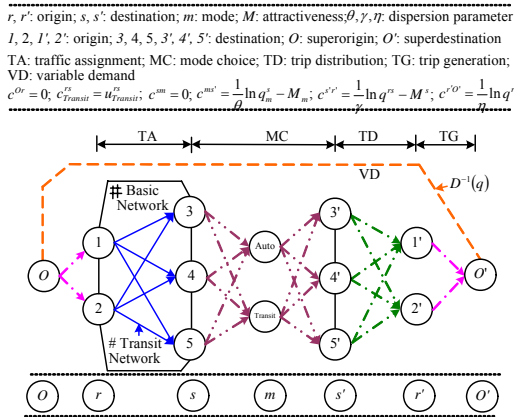

Fig. 1 Supernetwork for the combined TO/TD/MC/TA/VD model

To solve the combined TO/TD/MC/TA/VD model, a nested solution algorithm embedding any traffic assignment algorithm was developed (cf. Appendix II). So far, this new algorithm embedding the F-W algorithm or B algorithm has been successfully applied to some instances. In the future, TAPAS should be adopted instead of the F-W algorithm or B algorithm, as a unique path-flow solution is required.

VI. CONCLUDING REMARKS

In this paper, we have discussed three important issues in travel demand forecasting. First, to avoid interface inconsistencies among four travel choices, the combined model approach is required. Second, to improve the computational efficiency, quick and precise solution algorithms such as TAPAS, LUCE, B, OBA, or PG should be adopted because combined models must be solved to a required precision level within an acceptable period of time. Third, in view of practical applications in transportation, a unique path-flow solution is needed. For the reader's reference, the combined four-step choice model with variable demands and a nested solution algorithm are also provided. Note that though not mentioned in this paper, temporal dimension should be incorporated into the combined model which is certainly a subject worth exploring in the near future.

APPENDIX I

As an optimization model, the unified framework for the four-step combined model with variable demands must consist of an objective function and a feasible region that includes three types of constraints (i.e., flow conservation, non-negativity and definitional constraints).

$$\min_{(\mathbf{x},\mathbf{q}_m^{rs},\mathbf{q}^{rs},\mathbf{q}^r)\in\Omega} z =$$

$$\sum_m \sum_l \int_0^{x_{ml}} c_{ml}(\omega)d\omega + \sum_s \sum_m \int_0^{q_m^s} c_m^s(\omega)d\omega +$$

$$\sum_m \sum_{s'} \int_0^{q_m^{s'}} c_m^{s'}(\omega)d\omega + \sum_{s'} \sum_{r'} \int_0^{q^{s'r'}} c^{s'r'}(\omega)d\omega \qquad (1)$$

$$+ \sum_{r'} \int_0^{q^{r'}} c^{r'O'}(\omega)d\omega - \int_0^q D^{-1}(\omega)d\omega$$

where the feasible region Ω is defined by the following constraints.

Flow conservation constraints:

$$\sum_p f_{mp}^{rs} = q_m^{rs} \qquad \forall r,s,m \tag{2}$$

$$\sum_m q_m^{rs} = q^{rs} \qquad \forall r,s \tag{3}$$

$$\sum_r q_m^{rs} = q_m^s \qquad \forall r,s \tag{4}$$

$$\sum_s q^{rs} = q^r \qquad \forall r \tag{5}$$

$$\sum_r q^r = q \tag{6}$$

$$q + e = \overline{q}_{\max} \tag{7}$$

Non-negativity constraints:

$$f_{mp}^{rs} \geq 0 \qquad \forall r,s,m,p \tag{8}$$

Definitional constraints:

$$x_{ml} = \sum_r \sum_s \sum_p f_{mp}^{rs} \delta_{mlp}^{rs} \qquad \forall m, l \qquad (9)$$

$$c_{mp}^{rs} = \sum_l c_{ml} \delta_{mlp}^{rs} \qquad \forall r, s, m, p \qquad (10)$$

where

c_{ml} : the travel cost function associated with link l and mode m

c_m^s : travel cost function that characterize mode choice for destination s

$c_m^{s'}$: travel cost function that characterize destination s for mode m

$c^{r'O'}$: travel cost function that characterizes trip generation for origin r

$c^{s'r'}$: travel cost for link $s'r'$ that connect origin r and destination s

$D^{-1}(q)$: the inverse demand function; equivalent to the excess demand function $E(e)$

e : the excess flow associated with the entire area, $e = \bar{q}_{max} - q$

q : the total OD trip rate from the entire area

q^r : the total OD trip rate from origin r; flows on link $r'O'$

$q^{r's'}$: trip rate between origin r' and destination s'; flows on link $s'r'$

q_m^{rs} : the total flow by mode m between origin r and destination s

$q_m^{s'}$: the total flow by mode m between origin r and destination s; flows on link ms'

q_m^s : the total flow by mode m between origin r and destination s; flows on link sm

$\bar{q}_{max}$: the upper limit of total OD trip rate from the entire area

δ_{mlp}^{rs} : 1, if link l associated with mode m is on path p between O-D pair rs.

Equation (1) defines the objective function by summing the integrals of link travel costs for all links, including real links for traffic assignment and virtual links for modal choice, trip distribution, trip origin as well as variable demand. (2) conserves flows for each O-D pair by mode m. (3) conserves flows for each O-D pair (4) conserves flows for destination s by mode m. (5) conserves flows for each origin r. (6) conserves flows for the entire area (7) sets the upper limit of total traffic demand for the entire area. Equation (8) requires path flow associated with each mode and route be negative. Equations (9) and (10) are definitional constraints.

APPENDIX II

With the above description, we now proposed a nested solution algorithm that solves a series of combined travel choice models as follows:

Step 0. Input traffic data including upper limit of trips for the entire traffic area, traffic demand functions respectively for the entire traffic area, for each origin-destination pair, for each modes as well as link relevant data such as free flow link travel costs, link capacities and link cost functions.

Step 1. Solve the area-wide trip assignment problem with variable demand (ATO/VD) resulting in the trip rates for the entire area.

Step 2. Solve the trip origin and assignment (TO/TA) problem, resulting in trip rates for each origin.

Step 3. Solve the trip distribution and assignment (TD/TA) problem, resulting in O-D trip rates.

Step 4. Solve the modal split and assignment (MC/TA) problem, resulting in the O-D trip rates by mode.

Step 5. Solve the mode-specific traffic assignment problem.

Step 6. Stop.

REFERENCES

[1] D.E. Boyce, M. Lupa, and Y. Zhang, "Introducing 'feedback' into four-step travel forecasting procedure vs. equilibrium solution of combined model," *Transportation Research Record*, vol. 1443, 1994, pp. 65–74.

[2] H. Bar-Gera, and D.E. Boyce, "Solving the sequential procedure with feedback," paper presented at *the Sixth International Conference of Chinese Transportation Professionals*, Dalian, China, 2006.

[3] S.P. Evans "Derivation and analysis of some models for combining trip distribution and assignment," *Transportation Research*, vol. 10, 1976, pp. 37–57.

[4] M. Florian, S. Nguyen, and J. Ferland, "On the combined distribution-assignment of traffic," *Transportation Science*, vol. 9, 1975, pp. 43–53.

[5] M. Abdulaal, and L. J. LeBlanc, "Methods for combining modal split and equilibrium assignment models," *Transportation Science*, vol. 13, 1979, pp. 292–314.

[6] M. Florian, and S. Nguyen, "A combined trip distribution modal split and assignment model," *Transportation Research*, vol. 12, 1978, pp. 241–246.

[7] H.K. Chen, 2013.07.27-28, "A combined model with the four travel choices and variable demand," *International Academic Conference on Social Sciences* (IACSS 2013), Istanbul, Turkey.

[8] H.K. Chen, and K.C. Yen, "A algorithm for the combined model with the four travel choices and variable demand," 2014. (In Chinese)

[9] M. Beckmann, C. B. McGuire, and C. B. Winsten, *Studies in the Economics of Transportation*, Yale University Press: New Haven, 1956.

[10] M. Frank, and P. Wolfe, "An algorithm for quadratic programming", *Naval Research Logistics Quarterly*, vol. 3, 1956, pp. 95–110.

[11] R. Jayakrishnan, W.T. Tsai, J.N. Prashker, and S. Rajadhyaksha, "A Faster Path-Based Algorithm for Traffic Assignment," University of California Transportation Center, Working Papers qt2hf4541x, University of California Transportation Center, 1994.

[12] T. Larsson, M. Patriksson, Simplicial Decomposition with Disaggregated Representation for the Traffic Assignment Problem, *Transportation Science*, vol. 26, no. 1, 1992, pp. 4-17.

[13] M. Patriksson, *The Traffic Assignment Problem: Models and Methods*, VSP: Utrecht, 1994.

[14] H. Bar-Gera, "Traffic assignment by paired alternative segments (TAPAS)," *Transportation Research part B*, vol. 44, 2010, pp. 1022–1046

[15] R.B. Dial, "A path-based user-equilibrium traffic assignment algorithm that obviates path storage and enumeration," *Transportation Research part B*, vol. 40, 2006, pp. 917–936.

[16] G. Gentile, "Linear user cost equilibrium: a new algorithm for traffic

assignment," Submitted to *Transportation Research Part B*, 2009.

[17] H. Bar-Gera, "Origin-based algorithm for the traffic assignment problem," *Transportation Science*, vol. 36, no. 4, 2002, pp. 398–417.

[18] M. Florian, I. Constantin, and D. Florian, "A new look at the projected gradient method," In *CD-ROM of TRB 2009 Annual Meeting*, 2009.

[19] R.L. Tobin, and T. Friesz, "Sensitivity analysis for equilibrium network flow," *Transportation Science*, vol. 22, no. 4, 1988, pp. 242–250.

[20] H. Bar-Gera, and D. Boyce, "Route flow entropy maximization in origin-based traffic assignment," In: Ceder, A. (Ed.), *Proceedings of the 14th International Symposium on Transportation and Traffic Theory*, Jerusalem, Israel, Elsevier Science, Oxford, UK, 1999, pp. 397-415.

[21] J.G. Wardrop, "Some theoretical aspects of road traffic research," In *Proceedings of Institution of Civil Engineers, Part II*, no. 1, 1952, pp. 325–378.

[22] H. Bar-Gera, "Primal method for determining the most likely route flows in large road networks," *Transportation Science*, Vol. 40, No. 3, pp. 269-286, 2006.

[23] Y. Nie, "A class of bush-based algorithms for the traffic assignment problem," *Transportation Research part B*, vol. 44, no. 1, 2010, pp. 73–89.

[24] Y. Sheffi, *Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods*, Prentice-Hall Inc: New Jersey, USA, 1985.

[25] H.K. Chen, "Supernetwork representations for combined travel choice models," *Open Transportation Journal*, Vol. 5, 2011, pp. 92-104.

Huey-Kuo Chen received his B.S. (1977) from National Cheng-Kung University, Taiwan, M.S. (1979) from National Taiwan University, Taiwan, and Ph. D. (1989) from University of Illinois at Urbana-Champaign, USA. He is currently a professor in the Department of Civil Engineering at National Central University, Taiwan. His research interest covers transportation and network modeling, supply chain management and logistics, as well as hierarchical linear model.