Alternative formulations of a combined trip generation, trip distribution, modal split, and trip assignment model

Zhong Zhou a, Anthony Chen b,*, S.C. Wong c

a Citilabs Inc., 1040 Marina Village Parkway, Alameda, CA 94501, USA
b Department of Civil and Environmental Engineering, Utah State University Logan, UT 84322-4110, USA
c Department of Civil Engineering, The University of Hong Kong, Pokfulam Road, Hong Kong, China

A R T I C L E   I N F O

Article history:
Received 30 September 2007
Accepted 28 July 2008
Available online 8 August 2008

Keywords:
Transportation
Combined travel demand model
Random utility
Mathematical programming
Variational inequality

A B S T R A C T

The traditional four-step model has been widely used in travel demand forecasting by considering trip generation, trip distribution, modal split and traffic assignment sequentially in a fixed order. However, this sequential approach suffers from the inconsistency among the level-of-service and flow values in each step of the procedure. In the last two decades, this problem has been addressed by many researchers who have sought to develop combined (or integrated) models that can consider travelers’ choice on different stages simultaneously and give consistent results. In this paper, alternative formulations, including mathematical programming (MP) formulation and variational inequality (VI) formulations, are provided for a combined travel demand model that integrates trip generation, trip distribution, modal split, and traffic assignment using the random utility theory framework. Thus, the proposed alternative formulations not only allow a systematic and consistent treatment of travel choice over different dimensions but also have behavioral richness. Qualitative properties of the formulations are also given to ensure the existence and uniqueness of the solution. Particularly, the model is analyzed for a special but useful case where the probabilistic travel choices are assumed to be a hierarchical logit model. Furthermore, a self-adaptive Goldstein–Levitin–Polyak (GLP) projection algorithm is adopted for solving this special case.

Published by Elsevier B.V.

1. Introduction

The traditional four-step model has been widely used in travel demand forecasting. It considers trip generation (travel choice), trip distribution (destination choice), modal split (mode choice) and traffic assignment (route choice) sequentially in a top-down sequential process (Ortuzar and Willumsen, 2001). The outputs of one step serve as the inputs of the next step. In practice, it has been criticized by its inherent weakness, such as lack of a single unifying rationale that would explain or legitimize all aspects of demand jointly. It also suffers from inconsistent consideration of travel times and congestion effects in various steps of the procedure (see Boyce (2002, 2007), Oppenheim (1995), Garret and Wachs (1996), McNally (2000)). A common approach to remedy this inconsistency is to introduce a 'feedback' mechanism into the computational procedures. However, the convergence is not guaranteed. Thus, many simultaneous models have been proposed using behavioral assumptions (typically from random utility theory) as mathematical conditions to seek solutions that satisfy these conditions. The word ‘simultaneous’ is to solve the various travel choices concurrently to resolve the inconsistency problem in the sequential four-step model. Therefore, these models are often referred to as ‘combined’ or ‘integrated’ models.

The first constrained convex optimization formulation for the user equilibrium assignment problem with elastic demand was proposed by Beckmann et al. (1956) that considered travelers between every origin–destination (OD) pair to be a function of the travel service for that OD pair. Florian et al. (1975) and Evans (1976) extended the convex optimization formulation to consider destination choice by providing a combined distribution and assignment (CDA) model, where the trip distribution follows a gravity model with a negative exponential deterrence function and the traffic assignment follows a user equilibrium model. This approach was further extended by Florian (1977) and Florian and Nguyen (1978) to include modal split, where the two modes (auto and transit) may be either independent or interdependent. Later, the location choice and travel choice were incorporated together into a combined model proposed by Boyce et al. (1983, 1988), Safwat and Magnanti (1988) proposed a combined model by integrating all four steps in sequential demand forecasting, where trip generation was based on random utility maximization of travelers’ behavior through an accessibility measure, trip distribution and modal split were given by a logit model, and traffic assignment was based on Wardrop’s user equilibrium principle. Oppenheim (1995) also made similar extension to simultaneously consider...
the travel-destination-mode-route choice, which is based on the multinomial logit model in a hierarchical structure by assuming each traveler is a customer of urban trips, whose choice is reflected by random utility theory and budget constraints. The solution of the model was shown to correspond to individuals as well as aggregate utility maximization. Recently, several researchers further extended the combined model to account for multiple user classes, such as Lam and Huang (1992), Boyce and Bar-Gera (2001, 2004) and Wong et al. (2004). Based on the assumption that travel cost structures are either separable or symmetric, the above models were formulated as convex optimization programs, which have the advantage of a unique solution and readily available convergent solution algorithms. However, the separable or symmetric assumption of the cost structure may not be realistic in certain situations (e.g., non-separable link costs for modeling intersection delay with asymmetric interactions (Smith, 1982; Heydecker, 1983; Meneguzzer, 1995), asymmetric interactions between cars and truck (Mahmassani and Mouskos, 1988; Wu et al., 2006), non-additive route cost structures (Gabriel and Bernstein, 1997; Lo and Chen, 2000; Chen et al., 2001)). In order to model the asymmetric interactions, some general combined travel demand models were formulated as a variational inequality (VI) problem (e.g., Florian et al., 2002; De Cea and Fernandez, 2001; De Cea et al., 2003; Garcia and Marlin, 2005; Hasan and Dashiti, 2007) and a fixed point problem (Bar-Gera and Boyce, 2003, 2006). Compared to the convex optimization approaches, these general approaches allow general demand and cost functions and promise solution properties, such as existence and uniqueness, under mild continuity and monotonicity assumptions. Boyce and Xiong (2007) discussed the challenges and opportunities for applying these advanced combined travel demand models in China. More comprehensive reviews on the combined travel demand model can be found in Boyce and Bar-Gera (2004).

In this paper, we propose alternative formulations for a combined travel demand model that integrates trip generation, trip distribution, modal split, and traffic assignment, which is based on the well-established random utility theory in microeconomics, and provides an explicit, rigorous framework where each individual traveler is regarded as a consumer of urban trips such that their travel behavior can be interpreted as outcome of a rational decision-making process. First, an unconstrained optimization formulation is provided, which extends Daganzo’s (1982) and Sheffi and Powell’s (1982) unconstrained minimization optimization formulation for the stochastic user equilibrium (SUE) traffic assignment problem to represent the multi-dimensional hierarchical choice structure and maintains the behavioral richness through random utility theory. Similar to Daganzo’s (1982) and Sheffi and Powell’s (1982) formulation, the proposed formulation has the flexibility to handle general probabilistic distributions. It implies that, by choosing an appropriate probabilistic function, the proposed formulation has the ability to avoid the difficulties associated with the multinomial logit model (Daganzo and Sheffi, 1977). Second, to explicitly treat the general link cost structures (i.e., non-separable link cost functions), a VI formulation is proposed. It handles the situation where link travel time of each mode may be dependent on the flows of other modes on that link and this interaction may be asymmetric. Under certain probabilistic assumptions, our formulations include Oppenheim’s (1995) convex mathematical program as a special case. Some qualitative properties of the proposed formulations are also given to ensure the existence and uniqueness of the solution. Overall, the proposed formulations allow a more general and consistent treatment of travel choice over different dimensions as well as behavioral richness.

The remainder of the paper is organized as follows. In the next section, we introduce a combined travel demand model as well as some basic definitions of random utility theory that will be used in the development of two alternative formulations. Then, an unconstrained optimization formulation and a VI formulation are proposed in Section 2. Existence and uniqueness properties if the two formulations are also rigorously proved. A self-adaptive Goldstein–Levitin–Polyak (GLP) projection algorithm is introduced in Section 3. Section 4 provides some conclusions and future research. In addition, two appendices are provide for readers interested in the qualitative properties of the hierarchical multinomial logit model.

2. Model and mathematical formulations

This section presents the combined travel demand model and the development of two alternative formulations. Notation is provided first for convenience, followed by the combined travel demand model and the mathematical development of an unconstrained optimization formulation and a variational inequality formulation.

2.1. Notation

\begin{align*}
& i \quad \text{origin index} \\
& j \quad \text{destination index} \\
& m \quad \text{mode index} \\
& r \quad \text{route index} \\
& t \quad \text{represents travel choice} \\
& a_m \quad \text{link index of mode } m \\
& \beta \quad \text{parameters in the logit-based combined travel demand model: } \beta_1, \beta_m, \beta_x \text{ and } \beta_t \text{ are positive parameters associated with the route, mode, destination, and travel choice, respectively} \\
& \beta' \quad \text{Rescaled parameters, where } \frac{1}{\beta_t} = \frac{1}{\beta_x} - \frac{1}{\beta_r} = \frac{1}{\beta_m} \frac{1}{\beta_x} \frac{1}{\beta_m} = \frac{1}{\beta_t} \\
& \tau \quad \text{a scalar attached to travel time in the utility function (value of time)} \\
& h \quad \text{constant term in the utility specification, which can be specified as a linear function of socio-economic characteristics} \\
& h_i \quad \text{is the travel propensity of origin } i \\
& h_j \quad \text{is the attractiveness of destination } j \text{ from origin } i \\
& h_{im} \quad \text{is the attractiveness of mode } m \text{ between } i \text{ and } j \\
& d_{ir}^m \quad \text{link-route incidence indicator, 1 if link } a \text{ on the route } r \text{ from origin } i \text{ to destination } j \text{ on mode } m, 0 \text{ otherwise} \\
& d_{ir}^m \quad \text{link-route incidence matrix for trips from origin } i \text{ to destination } j \text{ on mode } m \\
& c_{mr} \quad \text{fixed cost of travel on link } a \text{ of mode } m \\
& c_{mij} \quad \text{fixed cost of taking route } r \text{ on mode } m \text{ from origin } i \text{ to destination } j; \quad c_{mij} = \sum_{a} c_{mr} a_{mr} \quad (\text{unconditional) joint probability of } x \text{ and } y \text{ (e.g., } P_{ijmn} \text{ is the probability that a traveler in origin } i \text{ travels to destination } j \text{ on mode } m \text{ through route } r \text{)} \\
& P_x \quad \text{a vector of } P_{ij}, \text{ i.e., } P_x = (\ldots P_{ij}, \ldots)^T \\
& P_{xy} \quad \text{Conditional probability of choosing } y \text{ given } x \text{ (e.g., } P_{ijmn} \text{ is the probability of choosing route } r \text{ given that a traveler in origin } i \text{ has chosen to travel to destination } j \text{ on mode } m) \\
& P_{xij} \quad \text{A vector of } P_{xij}, \text{ i.e., } P_{xij} = (\ldots P_{ij}, \ldots)^T \\
& V_{xy} \quad \text{total received utility of choosing an alternative } x \text{ (e.g., a route in the route choice level, a mode in the mode split level)} \\
& U_o \quad \text{observed (or measured, systematic) utility of choosing an alternative } x \\
& V_{ijmn} \quad \text{total received utility of choosing alternatives } x \text{ and } y \text{ (e.g., } V_{ijmn} \text{ is the total utility received from a single trip from origin } i \text{ to destination } j \text{ on mode } m \text{ and route } r \text{)} \\
& U_{xy} \quad \text{observed utility of choosing alternatives } x \text{ and } y \text{ (e.g., } U_{ijmn} \text{ is the observed utilities for a traveler making a trip from location } i \text{, choosing destination } j \text{, choosing mode } m \text{ and choosing route } r \text{)} \\
& V_{ij} \quad \text{conditional utility of choosing alternative } y \text{ given } x \text{ (e.g., } V_{ijmn} \text{ represents the conditional utility that a traveler received from a trip on route } r \text{ given that origin } i \text{, destination } j \text{, and mode } m \text{ have been chosen)} \\
\end{align*}
2.2. Combined travel demand model

The combined travel demand model studied here is shown in Fig. 1, where the individual traveler’s decision process is represented by a hierarchical structure of the traveler choice process.

Note that, the decision tree above is structured sequentially just for analytical purpose, they are made simultaneously in reality, i.e., these choices are interrelated with each other. The order above is somewhat arbitrary and has to be decided based on real situations.

According to random utility theory of users’ behavior, we say that the received utility of choosing alternative k (e.g., a route in the route choice level, a mode in the mode split level) is

\[ V_k = U_k + \epsilon_k, \]

where \( U_k \) is the observed (or measured, systematic) utility, which depends on both the attributes of the alternative and the characteristics of the traveler, and \( \epsilon_k \) is the random (or error) term, which represents the uncertainty factors of the utility.

For the hierarchical structure above, all levels of choice have to be considered simultaneously. That is, the demand patterns, \( T_i, T_j, T_{ij}, \) and \( T_{ijm} \), have to be estimated jointly. Thus, the total utility received from a single trip from origin i to destination j on mode m and route r can be specified as

\[ V_{ijmr} = U_{ijr} + \epsilon_{ijr} + U_{ijm} + \epsilon_{ijm} + e_{ijm}. \]

where \( U_{ijr}, U_{ijm}, U_{ijmr} \) are the observed utilities for a traveler making a trip from location i, choosing destination j, choosing mode m and choosing route r, respectively, and \( \epsilon_{ijr}, \epsilon_{ijm}, \epsilon_{ijmr} \) are the corresponding error terms (i.e., unobserved attributes) associated with each travel choice level. The hierarchical structure in Fig. 1 implicitly imposes certain restrictions on the error terms. That is, the nested structure does not allow choice alternatives to share common unobserved attributes along all the travel choice dimensions and the unobserved attributes are additive components. This restriction is adopted in all the current combined travel demand models reviewed in the introduction, including the one presented in this paper. However, no specific distribution assumption on the error terms is needed here. All we assume is that the error terms are continuous random variables with continuously differentiable density functions. Without loss of generality, let

\[ U_{it} = h_i, U_{ij} = h_{ij}, U_{ijm} = h_{ijm} \quad \text{and} \quad U_{ijmr} = -g_{ijmr}, \]

where \( h_i \) is the traveling propensity of origin i, \( h_{ij} \) is the attractiveness of destination j from origin i, \( h_{ijm} \) is the attractiveness of mode m between i and j, and \( g_{ijmr} \) is the generalized travel cost of taking route r on mode m from origin i to destination j. Note that the travel propensity and attractiveness in the observed utilities can also be specified as a linear function of socio-economic characteristics. Furthermore, based on random utility theory, the probability that a traveler chooses an alternative k is equal to the probability that the utility of alternative k is greater than (or equal to) that of any other alternatives, i.e.,

\[ P_k = \text{Pr}(V_k \geq V_r \, \forall r \in D), \]

where D is the set of alternatives. Following the same line of thought from several earlier studies (e.g., Daganzo, 1979; Williams, 1977; Ben-Akiva and Lerman, 1978; Sheffi and Daganzo, 1978; Safwat and Magnanti, 1988), the satisfaction function \( W(U) \) can be defined as

\[ W(U) = E \left[ \max_{V \in D} \left( U + e \right) \right] \]

to capture the expected received utility that a traveler received from the set of alternatives.

Thus, we can easily derive the satisfaction of each stage (without loss of generality, we set \( U_0 = 0, V_0 = U_0 + \epsilon_0 \)), i.e.

\[ W_m = \max_{V} \left[ \max_{V_{ijm}} \left( W_j + \epsilon_j \right) \right] \]

\[ 0 = \max_{V} \left[ \max_{V_{ijm}} \left( W_r + \epsilon_r \right) \right] \]

\[ 0 = \max_{V} \left[ \max_{V_{ijm}} \left( W_r + \epsilon_r \right) \right] \]

where \( V_{ijm} = U_{ijm} + \epsilon_{ijm} \), \( V_{ij} = U_{ij} + W_{ijm} + \epsilon_{ijm} \), \( V_{ij} = U_{ij} + W_{ijm} + \epsilon_{ijm} \), and \( V_{ij} = U_{ij} + W_{ijm} + \epsilon_{ijm} \) represent the conditional utility that a traveler received from a trip conditioned by the travel choice
of the upper level, respectively. For example, $V_{ijrm}$ represents the conditional utility that a traveler received from a trip on route $r$ given that origin $i$, destination $j$, and mode $m$ have been chosen.

It is well known that the satisfaction function $W(U)$ has three important properties (Sheffi, 1985):

(i) it is convex with respect to (w.r.t.) $U$;
(ii) the partial derivative of satisfaction function w.r.t. the systematic utility of an alternative equals to the choice probability of that alternative, i.e., $\frac{\partial W_{ij}}{\partial U_{ij}} = P_i(U)$ $\forall s \in D$;
(iii) the satisfaction function is monotonic w.r.t. the size of the choice set, i.e., $W(U_1, \ldots, U_j, U_{j+1}) \geq W(U_1, \ldots, U_j)$.

Thus, from Eq. (6), the probability that a traveler will choose an alternative in each stage can be derived as follows:

$$P_{ijrm} = \frac{\partial W_{ij}}{\partial U_{ijrm}}, \quad P_{mij} = \frac{\partial W_{ij}}{\partial U_{mij}}, \quad P_{ij} = \frac{\partial W_i}{\partial U_{ij}}, \quad P_{i} = \frac{\partial W_i}{\partial U_i},$$

where $U_{ijrm} = -e_{ijrm}$, $U_{mij} = U_{ij} + W_{mij}$, $U_{i} = U_{ij} + W_{ij}$, and $U_{i} = U_{i} + W_{i}$.

Therefore, given the number of potential travelers $N_i$ at origin $i$ in a given time period (hour, day, etc.), the number of travelers taking route $r$ on mode $m$ from origin $i$ to destination $j$ in the study period can be computed by multiplying the conditional probability at each stage in a nested structure from the traffic assignment stage up to the trip generation stage:

$$T_{ijrm} = N_i \cdot P_{i} \cdot P_{j} \cdot P_{mij} \cdot P_{ijrm} \quad \forall i,j,m,r.$$

Similarly, the solutions for other stages can be derived in the same manner:

$$T_{ijm} = N_i \cdot P_{ij} \cdot P_{mj} \cdot P_{ijm} \quad \forall i,j,m,$$

$$T_{ij} = N_i \cdot P_{ij} \quad \forall i,j,$$

$$T_{i} = N_i \cdot P_{i} \quad \forall i.$$

2.3. Mathematical formulations

According to the description of the combined travel demand model above, we can see that the equilibrium is reached when no traveler can improve his/her received utility by unilaterally changing the multi-dimensional travel choice decision. Mathematically, the equilibrium condition is a feasible flow pattern $T = [T_{ijrm}, T_{ijm}, T_{ij}, T_{i}]^T$ that satisfies

$$T_i / N_i = P_{i} \quad \forall i,$$

$$T_{j} / T_i = P_{i} \quad \forall i,j,$$

$$T_{ijm} / T_{ij} = P_{mj} \quad \forall i,j,m,$$

and

$$T_{ijrm} / T_{ijm} = P_{ijrm} \quad \forall i,j,m,r.$$

Equivalently, there exists a feasible route-flow pattern $(T_{ijrm})$ such that

$$T_{ijrm} = N_i \cdot P_{ijrm} \quad \forall i,j,m,r,$$

where $P_{ijrm} = P_{ij} \cdot P_{i} \cdot P_{mj} \cdot P_{ijm}$.

In addition, the network flow conservation constraints have to hold, i.e.

$$N_i = T_i + T_{ij}, \quad T_i = \sum_j T_{ij}, \quad \forall i;$$

$$T_{ij} = \sum_m T_{ijm}, \quad T_{ij} = \sum_{j} T_{ijm} \quad \forall i,j,m.$$

Furthermore, if the route travel cost is additive, the equilibrium conditions in terms of link-flows by mode can be described as follows:

$$v_{ai}^m = \sum_i N_i \cdot P_{i}^m,$$

where

$$P_{i}^m = \sum_i P_{i} \cdot P_{i} \cdot P_{mij} \cdot P_{ijm} \cdot \delta_{ij}^m.$$

2.3.1. Unconstrained minimization problem

Assumption 1. Assume that the network $[N,A]$ is strongly connected, where $N$ and $A$ denote the sets of nodes and links, respectively. The number of potential travelers $N_i$ at origin $i$ in a given time period is non-negative, the link travel cost function $g_b^m(v)$ is continuous, differentiable, strictly increasing and separable (i.e., it only relies on its own mode and link-flow), and the generalized route cost function $e_{ijrm}$ is additive.

If Assumption 1 is satisfied, we can formulate the combined travel demand model as an unconstrained minimization problem as follows:

$$\min_{\nu} Z(\nu) = \sum_i N_i \cdot \bar{W}_i + \sum_m \sum_{a} g_a^m(v_a) \cdot v_a^m,$$

$$- \sum_m \sum_a \int_0 v_a^m g_a^m(\omega)d\omega.$$

Note that, the above formulation reduces to the unconstrained stochastic user equilibrium (SUE) traffic assignment problem (Daganzo, 1982; Sheffi and Powell, 1982) if only the route choice dimension of a single mode is considered. Thus, Eq. (17) can be regarded as a more general formulation, which has the ability to handle multi-dimensional travel choices and includes the unconstrained SUE traffic assignment problem as a special case. In the following, we show some qualitative properties of the proposed formulation.

Proposition 1. If Assumption 1 holds, the optimal solution of the unconstrained minimization problem (17) gives the equilibrium flow pattern of the combined travel demand model.

Proof. Assume that $\nu = (\ldots, v_a^m, \ldots)^T$ is an optimal solution of the minimization program (17), then the first-order necessary conditions of the unconstrained optimization problem should be satisfied, i.e.

$$\nabla Z(\nu) = 0.$$

Now, we focus on the term $\frac{\partial Z}{\partial v_a^m}(\nu)$ (i.e., the partial derivative of $Z(\nu)$ w.r.t. the flow on link $b$ of mode $n$). Since the link cost function $g_b^m(\nu)$ is continuous, differentiable and separable, and the route cost function $e_{ijrm}$ is additive, we have

$$\frac{\partial e_{ijrm}}{\partial v_a^m} = \frac{d g_b^m(v_a^m)}{d v_a^m} \cdot \delta_{ij}^b.$$

Then, by using Eqs. (6) and (7), it is easy to obtain

$$\frac{\partial}{\partial v_a^m} \left( \sum_i N_i \cdot \bar{W}_i \right) = - \sum_i N_i \cdot \sum_j P_{ij} \cdot P_{mj} \cdot P_{ijm} \cdot \frac{d g_b^m(v_a^m)}{d v_a^m} \cdot \delta_{ij}^b,$$

Furthermore, we have

$$\frac{\partial}{\partial v_a^m} \left( \sum_m \sum_a g_a^m(v_a^m) \cdot v_a^m - \sum_m \sum_a \int_0 v_a^m g_a^m(\omega)d\omega \right) = v_a^m \cdot \frac{d g_b^m(v_a^m)}{d v_a^m}.$$
Thus,
\[
\frac{\partial \mathcal{Z}(v)}{\partial v_a} = \left( - \sum_i N_i \cdot \sum_j \sum_{r} P_{ij} \cdot P_{ij} \cdot P_{nij} \cdot P_{rijn} \cdot \delta^v_{ijr} + v_b \right) \frac{\partial g^v_a(v^v_a)}{\partial v_a}.
\] (22)
According to the assumption that the link cost function is strictly increasing, Eq. (22) implies that
\[
v_a^{\text{m}} = \sum_i N_i \cdot \sum_j \sum_{r} P_{ij} \cdot P_{ij} \cdot P_{nij} \cdot P_{rijn} \cdot \delta^m_{ijr} \quad \forall m, a.
\] (23)

The above expression is exactly the equilibrium link-flows given in Eq. (15). Moreover, by directly taking the gradient w.r.t. the flow pattern \(T_{pqnk}\), we can obtain
\[
\frac{\partial \mathcal{Z}(v(T))}{\partial T_{pqnk}} = \sum_a \left( -N_p \cdot P_{p/q} \cdot P_{q/p} \cdot P_{n/q} \cdot P_{k/q} \cdot T_{pqnk} \right) \cdot \frac{\partial g^v_a(v^v_a)}{\partial v_a} \cdot \delta^m_{pqnk} \quad \forall p, q, n, k.
\] (24)
Thus, at the equilibrium point, Eq. (13) is naturally satisfied. Furthermore, it is easy to verify that the flow conservation constraints are satisfied since the sum of the probability at each stage is equal to 1, i.e., \(\sum P_{ijm} = \sum P_{mij} = \sum P_{rijn} = 1\).

Therefore, the optimal solution of the unconstrained minimization problem given in Eq. (17) satisfies the equilibrium conditions and gives the equilibrium flow pattern of the combined travel demand model. □

**Proposition 2.** If Assumption 1 holds, the equilibrium link-flow pattern of the combined travel demand model is unique.

**Proof.** Suppose \(v^m_a = (\ldots, v^m_{1a}, \ldots)\) and \(v^m_a = (\ldots, v^m_{2a}, \ldots)\) are two distinct equilibrium link-flow patterns. According to Eq. (15), we have
\[
v^m_a = \sum_i N_i \cdot A^v_i P_i (g^v_i);
\]
where \(P_i (g_i) = (\ldots, P_{ijm} (g_{ij}), \ldots)^\top, g^v_i = (A^v_i)^\top g_m (v^m), \) and \(g_m (v^m) = (\ldots, g^m_i (v^m), \ldots)\) for \(s = 1, 2\).

Thus,
\[
v^m_a - v^m_a = \sum_i N_i \cdot A^v_i (P_i (g^v_i) - P_i (g^m_i)) = \sum_i N_i \cdot A^v_i \nabla g_m P_i (g^m_i) (A^v_i)^\top \nabla g_m (v^m) (v^m - v^m),
\] (25)
where \(g^v_i = (A^v_i)^\top g_m (v^m), v^m \in [v^m_a, v^m_a]\) and Eq. (25) is based on the Mean-Value Theorem. Therefore,
\[
(v^m_a - v^m_a)^\top \nabla g_m (v^m) (v^m - v^m) = \sum_i N_i \cdot (\nabla g_m (v^m) (v^m - v^m))^\top A^v_i \cdot \nabla g_m P_i (g^m_i) (A^v_i)^\top \cdot \nabla g_m (v^m) (v^m - v^m).
\]

According to the properties of the satisfaction functions listed in Eq. (6), it is easy to verify that the matrix \(\nabla g_m P_i (g^m_i)\) is negative semi-definite. Thus, we have
\[
(v^m_a - v^m_a)^\top \nabla g_m (v^m) (v^m - v^m) \leq 0.
\] (26)

Finally, by considering the assumption that \(g_m (v^m)\) is strictly increasing and separable, we have that \(\nabla g_m (v^m)\) is positive definite, thus \(v^m_a = v^m_a\). (27)

This proves that the equilibrium link-flow pattern of the combined travel demand model is unique. □

From Propositions 1 and 2, we can immediately obtain the equivalence of the proposed formulation and the combined travel demand model.

**Proposition 3.** If Assumption 1 holds, the proposed unconstrained minimization program (17) is equivalent to the combined travel demand model.

Note that the distribution of the equilibrium flow pattern of the combined model depends on the probability distribution of the random error terms (i.e., \(e_{ij}, e_{eq}, e_{qm}, e_{gm}\)) in each decision stage. Thus, Eq. (17) is a general formulation that has the ability to handle any reasonable distribution. For example, Oppenheim’s (1995) combined travel demand model with a hierarchical logit structure can be included as a special case of the proposed formulation, where the random error terms are assumed to follow the independently and identically distributed (IID) Gumbel distribution.

Similar to other combined travel demand models formulated as a constrained convex optimization program (e.g., Beckmann et al., 1956; Florian et al., 1975; Evans, 1976; Florian and Nguyen, 1978; Boyce et al., 1983, 1988; Safwat and Magnanti, 1988; Lam and Huang, 1992; Boyce and Daskin, 1997; Boyce and Bar-Gera, 2001), the unconstrained minimization formulation is valid only for situations without asymmetric link interactions (e.g., the link travel cost of one mode is not affected by the flow of another mode on the same link). To model the more general situation that allows asymmetric link interactions, we formulate the combined travel demand model as a variational inequality problem.
It can also be formulated as another equivalent form: Find \((T_{ijmr}, T_{ijmr}, T_{pq}, T_{qr}) \in \Omega, \) such that
\[
\sum_i \sum_j \sum_m \sum_r \left( T_{ijmr} - P_{ijmr} \cdot N_i \right) (T_{ijmr} - T_{ijmr}) \geq 0, \tag{30}
\]
where the corresponding standard form could be written as
\[
F(x) = (\ldots (T_{ijmr} - P_{ijmr} \cdot N_i), \ldots)^T \text{ and } x = (\ldots T_{ijmr}, \ldots)^T.
\]

**Assumption 2.** Assume that the network \([N, A]\) is strongly connected, where \(N\) and \(A\) denote the sets of nodes and links, respectively. The number of potential travelers \(N_i\) at origin \(i\) in a given time period is non-negative, the generalized link cost function \(g^m_{ij}(v)\) is continuous and the generalized route cost function \(g^m_{ij}(v)\) is additive. The probability distribution function at each decision stage satisfies the conditional probability of each alternative being greater than zero.

Note that this probability distribution assumption is typically satisfied by most of the commonly applied distributions in practice, such as the Gumbel distribution, the log-normal distribution and the multivariate normal distribution.

**Proposition 4.** Suppose Assumption 2 holds, the VI formulations (28) and (30) are equivalent to the combined travel demand model.

**Proof.** To show the equivalence, we only need to show that \(x^\ast\) is an optimal solution of the problems if and only if it is an equilibrium flow pattern of the combined travel demand model. First, if \(x^\ast\) is an equilibrium flow pattern of the combined travel demand model, the equilibrium conditions Eqs. (12), (28), and (30) are satisfied naturally. Second, suppose \(x^\ast\) is an optimal solution of the VI formulations (28) and (30). Without loss of generality, we can assume \(x^\ast\) is greater than zero. Otherwise, if any \(T_{ijmr}^\ast\) equal to zero, we will have \(N_i\) equal to zero due to Assumption 2, which naturally satisfies the equilibrium conditions. Then, according to Corollary 1.3 in Nagurney (1993), we have \(F(x) = 0\), which implies the equilibrium conditions are satisfied. Therefore, the proposed VI formulations are equivalent to the combined travel demand model. □

**Assumption 3.** Assume that the network \([N, A]\) is strongly connected, where \(N\) and \(A\) denote the sets of nodes and links, respectively. The number of potential travelers \(N_i\) at origin \(i\) in a given time period is non-negative, the generalized link cost function \(g^m_{ij}(v)\) is continuous differentiable and the generalized route cost function \(g^m_{ij}(v)\) is additive.

**Proposition 5.** Suppose Assumption 3 holds, the VI formulations (28) and (30) have at least one solution.

**Proof.** Note that the choice probabilities can be represented as a function of flow pattern. Thus, according to the continuous and additive assumption of the cost functions and definitions of the choice probabilities in Eq. (7), we can see \(F(\cdot)\) is a continuous mapping. Since \(\Omega\) is non-empty, convex and compact, the solution existence is obtained (Theorem 1.4, Nagurney, 1993). □

From the formulation, we can observe that it can handle general distributions due to the direct utilization of the probability expression in each decision stage. Moreover, it can also handle more practical situations (e.g., non-separable link cost functions with asymmetric interactions, link interactions by modes or by user classes, non-additive route cost functions, etc.) since it only assumes \(g^m_{ij}(v)\) is continuous. According to Proposition 2 under the assumption that link travel cost function is strictly increasing, for VI formulations (28) and (29), we can also obtain unique link-flow pattern \(v^\ast\) at the optimal point. However, the solutions \((T_{ijmr}, T_{ijmr}, T_{pq}, T_{qr})\) may not be unique. Thus, in the following, we consider a special case that the random error terms at each decision stage (i.e., \(e_{it}, e_{ijmr}, e_{ijmr}\)) follow the IID Gumbel distribution, which results in a hierarchical multinomial logit model. Though the logit model has been criticized for its inherent weakness due to the independence from irrelevant alternative (IIA) property (e.g., Daganzo and Sheffi, 1977), it is still of special interest since the probability can be expressed in a closed form, which makes it much more convenient for numerical computations and applications. Under this special case, we propose another equivalent VI formulation, and show the existence and uniqueness of the equilibrium flow pattern under certain conditions.

2.3.3. Special case: Hierarchical multinomial logit model

Since the random error terms follow the IID Gumbel distribution, according to random utility theory (Eqs. (6) and (7)), it is easy to derive the traveler’s choice probability at each stage will follow a hierarchical multinomial logit model (McFadden, 1981):
\[ c_{ijm} = g_{ijm}(T_{ijm}) - \frac{1}{\beta_r} \ln T_{ijm} \quad \forall i,j,m,r; \]
\[ c_{ij} = \frac{1}{\beta_m} \ln T_{ij} - h_{ijm} \quad \forall i,j; \]
\[ c_0 = \frac{1}{\beta_t} \ln T_{ij} - h_{ijm} \quad \forall i; \]
\[ c_0 = \frac{1}{\beta_t} \ln T_{ij} \quad \forall i; \]
\[ \ln \left( \frac{x_{v}}{\lambda_{v}} \right) \geq 0 \quad \forall x_{v} \in \Omega, \quad (41) \]
where \( f(x) = [c^T, e^T, e^T, e^T] \) and \( x = [T_{ijm}, T_{ij}, T_i, T_0]^T \).

Note that the proposed VI formulation is based on the generalized link cost function, which depends on the entire link-flow pattern. That means the proposed VI formulation has the ability to deal with asymmetric link interactions, which includes Dornheim’s (1995) convex programming formulation of the combined travel demand model with separable link cost structure as a special case.

In the following, we give some qualitative properties of the proposed formulation:

**Proposition 6.** Suppose Assumption 3 holds, \( x^* \in \Omega \) is a solution of the VI formulation (39), it is an equilibrium demand pattern of the combined travel demand model.

**Proof.** See Appendix A. \( \square \)

**Proposition 7.** Suppose Assumption 3 holds, the VI formulation (39) has at least one solution.

Now, we explore the uniqueness property of the equilibrium flow pattern \( (T_{ijm}, T_{ij}, T_i, T_0) \).

**Assumption 4.** Assume that the network \([N,A]\) is strongly connected, where \( N \) and \( A \) denote the sets of nodes and links, respectively. The number of potential travelers \( N_i \) at origin \( i \) in a given time period is non-negative and the generalized route cost function \( g_{ijm}(v) \) is additive.

**Proposition 8.** Suppose Assumption 4 holds, the VI formulation (39) has a unique solution.

**Proof.** See Appendix B.

Note that the assumption of Eq. (42) has been widely accepted in the literature (e.g., Ahn (1978) for the market equilibrium problem; Dafermos (1982) Florian and Spiess (1982), Sheffi (1985) and Patriksson (1994) for the asymmetric traffic assignment problem), which is required to establish the convergence of the optimal link cost function (i.e., diagonalization) algorithm. Since the verification of Eq. (42) may not be straightforward, some stronger conditions (e.g., the sufficiency condition) could be adopted as follows:

\[ \left| \frac{\partial g_i(v)}{\partial v_j} \right| > \max \left\{ \sum_{i,j} \left| \frac{\partial g_i(v)}{\partial v_j} \right| , \sum_{i,j} \left| \frac{\partial g_i(v)}{\partial v_j} \right| \right\}. \quad (43) \]

The above condition implies that the Jacobian matrix \( \nabla g(v) \) is diagonally dominant. It means the link cost of a particular class (i.e., a mode in this paper) depends mainly on the link-flow of its own class (mode).

We should also note that the assumption of Eqs. (42) and (43) requires a stronger condition to ensure the strictly monotone of \( g(T_{ijm}) \). However, note that the block matrices, \( \nabla g(v)^T \nabla g(v) \) and \( \nabla g(v)^T \nabla g(v) \) and \( R \), are all positive definite. The uniqueness of the solution in our formulation can be guaranteed under a much milder condition that only requires \( g(T_{ijm}) \) to be monotone. In other words, a unique flow pattern can be obtained as long as \( \nabla g(v) \) is positive semi-definite. As an example, consider the special case that there are no link interactions (i.e., separable link costs). This implies that there exists an equivalent convex optimization formulation (e.g., Oppenheim’s combined model (1995)). In this case, it is clear that monotonicity of \( g_m(v_m^*) \) is sufficient to guarantee a unique optimal flow pattern. \( \square \)

### 3. Solution procedure

To find the equilibrium flow pattern for the special case of the combined travel demand model (i.e., hierarchical multinomial logit model), many iterative methods may be adopted for solving the proposed VI formulations, such as the projection method, the non-linear Jacobian method, the successive over-relaxation method, the proximal point method and the Newton-type method (see Harker and Pang (1988)). Among these iterative methods, the projection method has received much attention due to its global convergence and simplicity of implementation.

Goldstein (1964), Levitin and Polyak (1965) proposed a gradient projection algorithm, i.e., GLP algorithm as follows: given an initial point \( x(0) \in \Omega \), the algorithm generates a sequence \( \{x(k)\} \) according to the following recursive equation:

\[ x(k+1) = P_{\Omega}(x(k) - \beta_k f(x(k))), \quad k = 0, 1, \ldots \quad (44) \]

where \( P_{\Omega}[.] \) denotes a unique projection on \( \Omega \), and \( \beta_k > 0 \) is a judiciously chosen positive stepsize. The efficiency of the GLP algorithm is highly dependent on the proper selection of the stepsize \( \beta_k \). A large stepsize would lead to divergence, while a small stepsize would slow down the convergence. However, to determine an appropriate stepsize \( \beta_k \) is not a trivial task in general.

To overcome this difficulty, several different stepsize strategies have been suggested (e.g., Bertsekas, 1976; Nagurney and Zhang, 1996). Recently, a new self-adaptive Goldstein–Levitin–Polyak (GLP) projection algorithm has been proposed by Han and Sun (2004) and successfully implemented by Zhou and Chen (2003) for solving the asymmetric traffic equilibrium problem. A unique feature of this algorithm is that the stepsize is self-adaptive using the information derived from the previous iterations. This feature is designed to effectively minimize the expensive line searches and to guarantee global convergence. For solving the VI formulation (39), which corresponds to the special logit-based combined travel demand model with link interactions, this new self-adaptive GLP projection algorithm (Han and Sun, 2004; Zhou and Chen, 2003) can also be adopted. The main steps of the self-adaptive GLP projection algorithm are summarized as follows.

**Step 0: Initialization:**
- Set parameters: stepsize \( \beta_0 \), \( \beta_{\text{max}} > 0 \), \( \delta \in (0,1) \), \( \mu \in [0.5,1] \), stopping criteria \( \sigma, \gamma_0 = \beta_0 \), and iteration counter \( k = 0 \).
- Set the initial solution vector.
\[ x(0) = (T_{ym}(0), T_{ym}(0), T_{iy}(0), T_{iy}(0), T_{ai}(0))^T \]

- compute initial mapping
- Compute initial mapping
  \[ F(x(0)) = (c'(0), c''(0), c'(0), c''(0), c''(0)) \]

**Step 2: Termination:**
- Compute the residual error:
  \[ e(x(k), \hat{b}_k) = x(k) - P_{ai}[x(k) - \beta_k F(x(k))]. \]
- If \( \|e(x(k), \hat{b}_k)\|_2 \leq \sigma \), then stop. Output
  \[ x(k) = (T_{ym}(k), T_{ym}(k), T_{iy}(k), T_{iy}(k), T_{ai}(k))^T. \]

**Step 3: Self-adaptive scaling procedure**
- Determine an appropriate stepsize \( \hat{b}_{k+1} \) and obtain new solution vector:
  Given \( x(k) \) is not a solution point, find
  the smallest non-negative integer \( l_k \) such that
  \[ \hat{b}_{k+1} = l_k \hat{b}_k. \]
  \[ x(k + 1) = P_{ai}[x(k) - \beta_k F(x(k))]. \]

satisfy
\[ (2 - \delta) \hat{b}_{k+1} (x(k) - x(k + 1))^T (F(x(k)) - F(x(k + 1))) \]
\[ - \frac{\beta_{k+1}}{\beta_k} \|F(x(k)) - F(x(k + 1))\|^2 \]
\[ \geq \max \left\{ \frac{\hat{b}_{k+1} - \hat{b}_k}{\hat{b}_k} \|x(k) - \bar{x}(k + 1)\|^2, 0 \right\}. \]

- If
  \[ 0.5 \hat{b}_{k+1} (x(k) - x(k + 1))^T (F(x(k)) - F(x(k + 1))) \]
  \[ - \frac{\beta_{k+1}}{\beta_k} \|F(x(k)) - F(x(k + 1))\|^2 \]
  \[ \geq \max \left\{ \frac{\hat{b}_{k+1} - \hat{b}_k}{\hat{b}_k} \|x(k) - \bar{x}(k + 1)\|^2, 0 \right\}, \]

then \( \hat{b}_{k+1} = \min\{\hat{b}_{k+1}/\beta_k, \hat{b}_k\} \); otherwise \( \hat{b}_{k+1} = \hat{b}_k \).

- Increment iteration counter \( k = k + 1 \) and go to Step 2.

The self-adaptive stepsize rule above is reminiscent to Bertsekas’s Armijo rule (1976). However, it is more practical and robust since we allow the sequence \( \hat{b}_k \) to be non-monotone (i.e., \( \hat{b}_k \) can decrease as well as increase). The results reported in Zhou and Chen (2003) showed that the self-adaptive scheme is significantly better than the strategies that use a fixed stepsize (may be empirical determined), or a predetermined stepsize sequence (Nagurney and Zhang, 1996). Furthermore, under the assumptions of the strongly monotone and Lipschitz continuous mapping \( F(x) \), the self-adaptive GLP method is globally convergent. Readers may refer to Han and Sun (2004) and Zhou and Chen (2003) for the convergence proof, detailed analysis of the self-adaptive stepsize scheme, and numerical results of the self-adaptive GLP projection method. Note that, in the procedure described above, the projection operator (i.e., \( P_{ai} \)) is performed on the feasible flow set, which is a polyhedron. Therefore, a convex quadratic program has to be solved, which may increase the computational overhead when the network size is large. To deal with this issue, a decomposition approach could be considered (Chen et al., 2002; Zhou and Chen, 2006).

4. Conclusions

New alternative formulations have been developed in this paper for a combined travel demand model, where each individual traveler is regarded as a consumer of urban trips and his (or her) travel behavior can be interpreted by random utility theory. The proposed formulations allow a systematic and consistent treatment of travel choice over different dimensions as well as behavioral richness. Specifically, the proposed unconstrained minimization program can be regarded as an extension of the unconstrained minimization program of the SUE traffic assignment problem, and the VI formulation has the ability to handle asymmetric link interactions and non-additive route cost structures. Both of them allow general distribution assumptions of the random error term in each travel choice dimension. In addition, an equivalent VI formulation is developed for a special and useful case where the random error terms follow the IID Gumbel distribution, which results in a hierarchical logit-based decision structure. This VI formulation generalizes Oppenheim’s convex programming model. Qualitative properties of the formulations are also given to verify the equivalence of the optimal solution and equilibrium conditions of the combined travel demand model and to ensure the existence and uniqueness of the solution. A self-adaptive GLP projection algorithm is provided for solving the special logit-based combined travel demand model. For future research, solution algorithm for the VI formulation with general distributions should be developed. One possibility is to combine the diagonalization (or relaxation) algorithm of Dafyantis (1982) with Sheffi and Powell’s method of successive averages with Monte Carlo simulation for solving the multi-dimensional travel choice problem with general distributions. It would also be of interest to further study other equivalent formulations of the combined travel demand model under a more flexible structure of the random error terms (e.g., paired combinatorial logit, cross-nested logit, logit kernel, etc.), to test the proposed formulations and algorithm with real-world applications, and to explore other effects of the combined travel demand model under different situations.

Acknowledgement

The work described in this paper is supported by a CAREER grant from the National Science Foundation of the United States (CMS-0134161).

Appendix 1

**Proposition 6.** Suppose Assumption 3 holds, \( x^* \in \Omega \) is a solution of the VI formulation (39), it is an equilibrium demand pattern of the combined travel demand model.

**Proof.** Rearrange the left-hand-side of (39), we obtain:
\[ \sum_i \sum_j \sum_m \sum_i \left[ g_{ym}(T_{ym}) - \frac{1}{f_k} \ln T_{ym} - \frac{1}{f_k} \ln T_{ym} \right] (T_{ym} - T_{ym}) \]
\[ + \sum_i \sum_j \sum_m \left[ \frac{1}{f_m} \ln T_{ym} - h_m - \frac{1}{f_m} \ln T_{ym} \right] (T_{ym} - T_{ym}) \]
\[ + \sum_i \sum_j \left[ \frac{1}{f_i} \ln T_{iy} - h_i - \frac{1}{f_i} \ln T_{iy} \right] (T_{iy} - T_{iy}) \]
\[ + \sum_i \sum_j \left[ \frac{1}{f_i} \ln T_{ai} - h_i - \frac{1}{f_i} \ln T_{ai} \right] (T_{ai} - T_{ai}) \geq 0. \]

Let \( \pi_{ym}, \lambda_{iy}, \phi_i, \phi_0 \) be the dual variables associated with the flow conservation constraints at each stage in (14). The KKT conditions of VI can be given as follows (Proposition 1.3.4, Facchinei and Pang, 2003):
Proposition 8. Suppose Assumption 4 holds, the VI formulation (39) has a unique solution.

Proof. Based on the standard form of proposed VI formulation (41), the gradient of $F(x)$ w.r.t. to $x$ can be written as the following block diagonal matrix:

$$
\nabla F(x) = \begin{bmatrix}
0 & \cdots & 0 \\
\vdots & \ddots & \vdots \\
0 & \cdots & 0
\end{bmatrix}
$$

It is easy to verify that these blocks, $\nabla v_1, \nabla v_2, \cdots, \nabla v_s$ and $\nabla v_s^0$, are diagonal matrices with positive diagonal entries, $\frac{1}{\lambda^k}$ and $\frac{1}{\lambda^k}$, respectively. The first block, $\nabla v_1$, can be written as the sum of two matrices:

$$
\nabla v_1 = \nabla v_1^0 + \nabla v_1^1
$$

where $g(T_{ij}) = (\cdots, g_{ij}, \cdots)^T$ and $R$ is a diagonal matrix with positive diagonal entries $\frac{1}{\lambda^k}$. From chain rule, the first term of Eq. (A17) can be expressed as

$$
\nabla v_1^0 \equiv \nabla F(T_{ij}) = (\frac{\partial g_{ij}}{\partial T_{ij}}) \cdot B(v)(\frac{d v_{ij}}{d T_{ij}})
$$

According to Assumption 4, (A2), $\nabla g(v)$ is positive definite (Ahn, 1978). Thus, from the analysis of the diagonal block matrices above, the positive definite of the Jacobian matrix $\nabla F(x)$ follows easily, which implies that the mapping $F(x)$ is strictly monotone (Facchinei and Pang, 2003). In addition, from Proposition 7, we know that there exists at least one solution, and the solution is unique for the proposed VI formulation (Nagurney, 1993).

References


