Research Article

Travel Demand-Based Assignment Model for Multimodal and Multiuser Transportation System

Bingfeng Si, Xuedong Yan, Hunjun Sun, Xiaobao Yang, and Ziyou Gao

MOE Key Laboratory for Urban Transportation Complex Systems Theory and Technology,
Beijing Jiaotong University, Beijing 100044, China

Correspondence should be addressed to Bingfeng Si, bfsi@bjtu.edu.cn

Received 19 July 2012; Accepted 27 October 2012

Academic Editor: Xue-Xiang Huang

In this paper, the structural characteristic of urban multimodal transport system is fully analyzed and then a two-tier network structure is proposed to describe such a system, in which the first-tier network is used to depict the traveller’s mode choice behaviour and the second-tier network is used to depict the vehicle routing when a certain mode has been selected. Subsequently, the generalized travel cost is formulated considering the properties of both traveller and transport mode. A new link impedance function is proposed, in which the interferences between different vehicle flows are taken into account. Simultaneously, the bi-equilibrium patterns for multimodal transport network are proposed by extending Wardrop principle. Correspondingly, a bi-level programming model is then presented to describe the bi-equilibrium based assignment for multi-class multimodal transport network. The solution algorithm is also given. Finally, a numerical example is provided to illustrate the model and algorithm.

1. Introduction

With the rapid development of economy, the transportation infrastructures have been improved significantly in the most cities of China, especially in some metropolitan cities like Beijing and Shanghai, where the integrated urban transportation systems have been established gradually. Synchronously, the modal share for passenger travel has been dramatically changed. Table 1 lists statistics of the trip intensity and mode split of Beijing in 1986, 2000, 2005, and 2010, respectively [1].

It shows that Beijing’s transportation development mode is a typical multimodal transportation system. The system consists of different transportation subsystems or subnets for passenger cars, buses, trains, bicycles, and so forth, in which the multimodal traffic flows are interdependent and interactive. Obviously, the equilibriums between the various
subsystems and within each subsystem are much more complicated than that for a pure private vehicle system. Thus, the following critical issues should be carefully considered for resolving the multimodal network equilibrium problem.

(i) The multimodal transportation network is a superposition or compound of various physical subnets for different transportation modes.
(ii) The performance of each mode depends on both self-demand and the demands in other modes, which means that there are interactions between different modes.
(iii) The traffic flow pattern in a multimodal network involves traveler’s combined choice behaviors in which the travelers not only choose trip modes through the whole multimodal network but also select routes within each subnet.
(iv) The criteria of mode choice and route choice during a trip are usually different. In the mode choice stage, the travelers’ decisions are generally influenced by a combination of travel time, potential expense, and other factors. Once the trip mode has been selected, the travelers only care about how to minimize their travel time through route choice within the specific subnet. Therefore, different types of travelers have different psychological preferences for mode choice while the characteristics of travelers have no impact on their route choices.
(v) There are feedbacks between these two choices. Firstly, the traveler’s mode choice results in the total demands for different transportation modes, which determines the traffic flows through the multimodal network. Secondly, the traffic assignment patterns corresponding to the route choice within subnets determine the travel time in the respective modes, which conversely affect the mode choice.

The user equilibrium (UE) assignment problem for the private vehicle traffic network has been formulated by Beckmann et al. [2], Sheffi [3], Patriksson [4], and so forth. For multimodal networks, the earlier models [5–9] were developed for modal choice using logit type functions to split travel demand for each travel mode. However, these models cannot reflect a multimodal network’s configuration and how the traffic flows are distributed in the network [10]. In order to overcome this, the combined or integrated models have been developed [11–13], in which modal split and flow assignment are incorporated together. Based on the assumption that travel cost structures are either separable or symmetric, the above models were formulated as convex optimization programs. However, the assumption for the cost structure may be not realistic in certain situations [14, 15]. In order to model the asymmetric interactions, some general combined travel demand models were formulated as a variational inequality problem [10, 16–20] and a fixed point problem [21]. Although the above studies have combined the traveler’s modal choice and traffic flow assignment, they usually focus on user equilibrium for path flow assignment within single mode traffic network. However, the multimodal equilibrium issue associated with mode split has rarely been explored, and few studies investigate the relationship between these two equilibriums.

<table>
<thead>
<tr>
<th>Year</th>
<th>Total trips/day</th>
<th>Average trip distance (km)</th>
<th>Bus Mode split (%)</th>
<th>Subway</th>
<th>Taxi</th>
<th>Car</th>
<th>Bicycle</th>
</tr>
</thead>
<tbody>
<tr>
<td>1986</td>
<td>1.61</td>
<td>—</td>
<td>29.31</td>
<td>0.36</td>
<td>5.24</td>
<td>65.09</td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>2.77</td>
<td>8.0</td>
<td>27.33</td>
<td>9.03</td>
<td>23.96</td>
<td>39.68</td>
<td></td>
</tr>
<tr>
<td>2005</td>
<td>2.64</td>
<td>9.3</td>
<td>26.60</td>
<td>5.70</td>
<td>7.60</td>
<td>29.80</td>
<td>30.30</td>
</tr>
<tr>
<td>2010</td>
<td>2.82</td>
<td>10.6</td>
<td>28.20</td>
<td>11.50</td>
<td>6.70</td>
<td>34.20</td>
<td>16.40</td>
</tr>
</tbody>
</table>
From the viewpoint of economics, the transportation service can be measured by the generalized travel cost in addition to the expense charged. The generalized cost is not fixed given a specific trip but is dependent on travel demand. If many individuals choose to use a certain mode, it will get congested and its travel time will increase. In response, some travelers may take alternative modes. Consequently, the alternative modes can also be congested, which will push travelers back to the original mode. Therefore, the UE principle in a multimodal system includes two categorizes. One is the user equilibrium between different modes, and the other is the traditional user equilibrium among different routes in respective subnets. The first equilibrium derives from the monotonically increasing of the generalized cost of each mode with travel demand while the later one derives from travel time in a specific subnet also with the monotonically increasing nature of the link impedance functions. There exists a two-way influence between the two types of equilibriums. If the travel demands of transportation modes are all given, the demands will be assigned in each subnet based on traditional user equilibrium, under which the travel time of various transportation modes will be obtained. Subsequently, such travel times will lead to the changes in the generalized costs of different modes, and then the travelers will reselect transportation modes. Eventually, both user equilibrium between different modes and user equilibrium between different routes in each mode are achieved.

The generalized travel cost is associated with the properties of transportation mode, such as travel time, travel expense, and convenience. Meanwhile, the traveler’s psychological preference is another important factor leading to different mode-choice behaviors. Different types of travelers may perceive different values of the properties stated above. For example, the high-income travelers will concern more about the factor of time while the low-income travelers will care more about the factor of expense. To account for multiple user classes that can be distinguished by the value of travel time, the multiclass, multicriteria traffic network equilibrium models were developed, in which each class of travelers perceives the travel disutility associated with a route as a subjective weighting for travel time and travel cost [22]. The models allow both travel time and travel cost of a link to depend on the entire link load pattern, rather than on the particular link flow only [23–25]. The multiclass, multicriteria models were further applied for dynamic traffic assignment [26] and multimodal network issues, such as using aggregate hierarchical logit structures for mode choice [17, 27] and extending the fixed point theory to the multimodal network equilibrium model [21]. However, the previous studies considered the factors of travel time and expense in the link impedance function, and the same criteria are adopted for the travelers’ mode choice and route choice. So, the travelers have different preferences not only in the stage of mode choice but also for route choice. In fact, the influence factors of traveler’s mode choice in the multimodal network and route choice in a single modal network are different. The generalized travel cost involving travel time, expense, or other factors should be addressed as user preference or multiclass problems in the stage of mode choice. Once travelers decide which mode they would take, travel time is the only factor for their route choices unless there is an imposed charge for selecting the shortest travel time, for example, toll for urban freeways. Such a special issue in the multimodal transportation system is beyond the scope of this study. Nevertheless, few studies clearly indicate the criteria difference between mode choice and route choice, neither is the issue formalised in previous multimodal network models.

In addition, the structure of multimodal system is generally more complicated than that of private car roadway system. Hierarchical structure is an efficient way to model the multimodal system with multiple subnet levels. Mainguenaud [28] presented a data
model to manage multimodal networks with a Geographical Information System (GIS), which allows definition of a node and link as an abstraction of a subnet. Jing et al. [29] proposed the Hierarchical Encoded Path View (HEPV) model that partitions large graph into smaller subgraphs and organizes them in a hierarchical fashion. Van Nes [30] introduced a strategy for hierarchical multimodal network levels utilizing specific journey functions according to travel distance as well as quality in terms of travel speed and comfort. Jung and Pramanik [31] developed a graph model, called hierarchical multilevel graph, for very large topographical road maps. This graph model provides a tool to structure and abstract a topographic road map in a hierarchical fashion. These studies mainly focused on topology of multimodal transportation network and discussed how to deal with the problem of a very large volume of data, but the travelers’ choice behaviors in the multimodal network were rarely investigated. Not until recent years, researchers have started to integrate traveler’s mode choice and route choice with complex network structure in the multimodal network models. Lo et al. [32] transformed a multimodal network to a so-called State Augmented Multimodal (SAM) network, by which the network equilibrium problem can be resolved directly. Wu and Lam [33] used a multilayer network to represent the multimodal network with combined modes that can facilitate generating feasible routes. García and Marín [34] explored the network equilibrium model in the space of hyperroute flows, which contributes to considering asymmetric costs and modeling multimodal network in a more flexible way. Si et al. [35–37] presented an augmented network model for urban transit system. The route choice in the augmented transit network was defined according to the passengers’ behaviors, and the corresponding network equilibrium model with an improved shortest path algorithm was developed for the urban transit assignment problem.

The objective of this study is to address the aforementioned concerns of the urban multimodal network equilibrium issue, including (1) assigning traffic based on both user equilibrium between different modes and user equilibrium between different routes; (2) adopting different criteria for travelers’ mode choice and route choice behaviors, namely, using multiclass-related general travel cost in the stage of mode choice and traditional link impedance for route choice within each single mode subnet; (3) constructing a hierarchical network to describe the multimodal transportation system, in which the first-tier network is used to depict the travelers’ mode choice behaviors and the second-tier network is used to describe travelers’ route choice behaviors within the single mode subnets. In this paper, the biequilibrium patterns for multimodal transportation network are proposed by extending Wardrop principle. Correspondingly, a bilevel programming model with its solution algorithm is applied for the biequilibrium traffic assignment in the multimodal transportation network. Finally, a numerical example is provided to illustrate the model and algorithm.

2. Hierarchical Network for Multimodal Transportation System

In this paper, the multimodal transportation system is expressed as $G = (N, A, K)$, where $A$ is the set of roads, $N$ is the set of nodes that usually represent the intersections or zones, and $K$ is the set of transportation modes. Clearly, there are $K$ subnets in the multimodal transportation system, and each subnet, represented by $G_k = (A_k, N_k)$, corresponds to transport mode $k (k \in K)$.

Figure 1 illustrates a physical network example for the proposed multimodal network system, which consists of one O-D pair $(r-s)$, nine nodes, twelve roads, and three transportation modes (car, bus, and bike). It shows that the different modes have different network
structures. Figure 2 shows the subnets for different modes separately, in which the traveler can choose the different routes from node 1 to node 9.

Generally, the traveler during a trip from origin to destination should make two successive decisions in the multimodal system. The first one is the mode choice in the whole network, and the second is the route choice in the corresponding subnet once a mode is selected. At the first stage, the multimodal system can be represented as a simple network by the different connections, as shown in Figure 3.

According to the structural features of urban multimodal transportation system demonstrated above, a hierarchical network model can be used to describe such a system. In the model, each node is described by two variables \( (n, k) \), where \( n(n \in N) \) denotes the location in the physical network and \( k(k \in K) \) denotes the transportation mode. Note that the notations of origin and destination nodes require special attention. An origin node is denoted as single variable \( r \) and a destination is denoted as \( s \), where \( r \) and \( s \) designate their physical locations. The set of links connecting the different nodes is divided into two categories. One category includes loading link and unloading link, the end of which is either origin or destination; the other category only includes in-vehicle link that indicates connectivity in each subnet. Both categories are all described by two variables \( (a, k) \), where \( a(a \in A) \) denotes the physical road and \( k(k \in K) \) denotes the transportation mode. The hierarchical multimodal transport system is described in Figure 4.

In such a hierarchical network, the origin is connected with different subnets by the loading links. Similarly, the destination is connected with different subnets by the unloading links. If all travelers are assumed to complete their trips through only one mode, it implies that there should be no connectivity between subnets in the hierarchical network. Based on
the hierarchical network, the multimodal transportation system can be used as a generalized network for traffic assignment or network analysis.

### 3. Travel Costs Based on Traveler’s Characteristics

In this paper, all travelers are divided into $I$ classes by socioeconomic attributes, assuming that the mode-choice decision is homogeneous within each class, but differs among classes. Moreover, the travel time of each mode depends on the travel demands for the mode, and the potential expense of each mode is included in the generalized travel cost for different travelers. The generalized travel costs of different modes for different traveler classes can be expressed as follows:

$$c^{i,k}_w = \alpha^i \mu^k_w(q) + \beta^i \tau^k_w, \quad \forall w, k, i,$$

(3.1)
where $c_{iw}^{ik}$ is the generalized cost of mode $k$ for class $i$ between O-D pair $w$; $\rho_{iw}(q)$ represents the equilibrium travel time for transportation mode $k$ between O-D pair $w$, which is decided by the travel demands (represented by $q$); $\tau_{iw}^{k}$ denotes the potential expense of transportation mode $k$ between O-D pair $w$; $\alpha_i$ and $\beta_i$ are parameters related to socioeconomic attributes of class $i$.

Similar to the general traffic network, the travel time of class $i$ on route $r$ in subnet $k$ between the O-D pair $w$, denoted by $t_{iw,r}^{ik}$, can be obtained by the travel time on the link, that can be expressed as follows:

$$t_{iw,r}^{ik} = \sum_a t_a^{ik} \delta_{a,r}^{k,w}, \quad \forall w,k,i,r,$$

(3.2)

where $t_a^{ik}$ denotes the travel time of class $i$ selecting mode $k$ on road $a$; $\delta_{a,r}^{k,w}$ is route and road incidence variable in the subnet $k$ between O-D pair $w$; if road $a$ is on the route $r$, then $\delta_{a,r}^{k,w} = 1$, otherwise, $\delta_{a,r}^{k,w} = 0$.

Generally, no matter what class of travelers, as long as the transportation mode is selected, the travel time in the corresponding subnet is not relevant to the personal properties. In other words, the travel time on the road network is only related to the characteristics of transportation modes, but not related to the traveler’s personal properties. Let $t_{w,r}^{ik}$ and $t_a^{ik}$ denote the travel time of mode $k$ on route $r$ between the O-D pair $w$ and the travel time of transportation mode $k$ on road $a$, respectively. Then,

$$t_{w,r}^{ik} = t_a^{ik} = \cdots = t_a^{ik}, \quad \forall w,k,r, \forall i \neq j,$$

(3.3a)

$$t_a^{ik} = t_a^{ik} = \cdots = t_a^{ik}, \quad \forall k,a, \forall i \neq j.$$

(3.3b)

Obviously, (3.2) can be rewritten as

$$t_{w,r}^{ik} = \sum_a t_a^{ik} \delta_{a,r}^{k,w}, \quad \forall w,k,r.$$

(3.4)

In the traffic network, the link impedance function mainly describes the relationship between travel time and link flow. It should be noted that the interferences among different modes will occur in the multimodal traffic network if there are no physical barriers between different flows on the road. Therefore, the link impedance function in the multimodal traffic network is very different from that in a single-mode traffic network. The travel time of different modes is decided by not only the road flow of its own mode but also the road flows of the other modes. Accordingly, the link impedance function in multimodal traffic network can be formulated as

$$t_a^k = f\left(t_a^{k(0)}, v_a^1, \ldots, v_a^k, C_a^k\right), \quad \forall k,a,$$

(3.5)

where $t_a^{k(0)}$ is the free-flow travel time of mode $k$ on road $a$; $C_a^k$ is the practical capacity on road $a$; $v_a^k$ is the vehicle flow of mode $k$ on road $a$. Generally, $t_a^{k(0)}$ and $C_a^k$ can be assumed as constants.
In the multimodal traffic network, the link flow is defined as the number of vehicles including cars, buses, and bikes that have traveled over the road sections during a time unit, which is a congregative result by all travelers’ mode choice and route choice behaviors. Therefore, the number of travelers can be looked upon as a variable in the link impedance function, by which the link flows and corresponding travel time can be calculated. Accordingly, the road flow can be represented by the travel demand as follows:

\[ v_a^k = x_a^k \left( \frac{U_k}{A_k} \right), \quad \forall k, a, \]  

where \( x_a^k \) is the travel demand of mode \( k \) on road \( a \); \( U_k \) is the PCU conversion coefficient of mode \( k \); \( A_k \) is the occupancy rate of mode \( k \), which indicates the average number of travelers within each vehicle of mode \( k \).

As stated above, the road flows of different modes on road \( a \) can be expressed by the travel demand of corresponding mode on road \( a \). Consequently, (3.5) can be rewritten as follows:

\[ t_a^k = f_a^k (x_{a1}, \ldots, x_{ak}), \quad \forall k, a. \]  

4. Conservations of Demand in Multimodal Transportation Network

Assuming that the total demands of different travelers between each O-D pair are given and fixed, for a certain class, the sum of demands of different modes equals the total demand between O-D pair, which can be represented as

\[ \sum_k q_{iw}^{i,k} = q_{iw}^i, \quad \forall w, i, \]  

where \( q_{iw}^i \) is the total demand of class \( i \) between O-D pair \( w \); \( q_{iw}^{i,k} \) is the demand of class \( i \) selecting mode \( k \) between O-D pair \( w \).

Secondly, for a certain class selecting a certain mode, the sum of demands on different routes in each subnet equals the demand of the corresponding mode between O-D pair, that is:

\[ \sum_r h_{wr}^{i,k} = q_{iw}^{i,k}, \quad \forall w, k, i, \]  

where \( h_{wr}^{i,k} \) is the demand of class \( i \) on the route \( r \) in subnet \( k \) between O-D pair \( w \).

Obviously, the following formulation can be obtained according to (4.2):

\[ \sum_i \sum_r h_{wr}^{i,k} = \sum_i q_{iw}^{i,k}, \quad \forall w, k. \]
Let \( q^k_w \) and \( h^k_{w,r} \) denote the total demand of mode \( k \) between O-D pair \( w \) and the demand on the route \( r \) in subnet \( k \) between O-D pair \( w \), respectively. Then the following two equations can be obtained easily:

\[
\sum_i q^i_{w} = q^k_w, \quad \forall w,k, \tag{4.4}
\]

\[
\sum_i h^i_{w,r} = h^k_{w,r}, \quad \forall w,k,r. \tag{4.5}
\]

Then, (4.3) can be rewritten as

\[
\sum_r h^k_{w,r} = \sum_i q^i_{w} = q^k_w, \quad \forall w,k. \tag{4.6}
\]

In addition, for class \( i \) in subnet \( k \) between O-D pair \( w \), the demand on road \( a \) can be represented by the demand on the routes passing through the road, that is:

\[
x^i_{a,k} = \sum_w \sum_r h^i_{w,r} \delta_{a,r}^k, \quad \forall a,i,k, \tag{4.7}
\]

where \( x^i_{a,k} \) is the demand of class \( i \) selecting mode \( k \) on road \( a \).

Similarly, the following formulation can be obtained according to (4.7):

\[
\sum_i x^i_{a,k} = \sum_i \sum_w \sum_r h^i_{w,r} \delta_{a,r}^k = \sum_w \sum_r \sum_i h^i_{w,r} \delta_{a,r}^k, \quad \forall a,i,k. \tag{4.8}
\]

Thus, the total demand of mode \( k \) on road \( a \) is the sum of demand of different classes selecting mode \( k \) on road \( a \), that is:

\[
x^k_{a} = \sum_i x^i_{a,k}, \quad \forall a,k. \tag{4.9}
\]

The following formulation can be gotten easily according to (4.5) and (4.9):

\[
x^k_{a} = \sum_w \sum_r h^k_{w,r} \delta_{a,r}^k, \quad \forall a,k. \tag{4.10}
\]

5. Biequilibrium Model for Multimodal Transport Network

Equilibrium is a central concept in numerous disciplines from economics and regional science to operational research/management science [38]. The example in transportation science is the famous Wardrop equilibrium. In the conventional equilibrium of transportation, the single-mode traffic network with purely automobile flow is considered, and only the motorists’ route choices are examined, while the traveler’s mode and route combined choices and the resulting complicated equilibrium in the multimodal traffic network have not been explored substantially.
As aforementioned, the UE principle in multimodal transportation system can be divided into two categories in order to be consistent with the travelers’ combined choice behaviors. One category of equilibrium exists between different modes, where the generalized travel cost for a certain class selecting a certain mode is the same and the minimum generalized travel costs of unselected transportation modes must not be less than the minimum cost between O-D pair. The other category is the traditional equilibrium among different routes in each single-mode subnet between O-D pair. The biequilibriums in the multimodal transportation system can be described as

\[
\begin{align*}
 c_{i,k}^{i,k} &= \eta_{i}^{i,k}, & q_{i,k}^{i,k} \geq 0, & \forall w,k,i, \\
 t_{w,r}^{k} &= \mu_{w,r}^{k}, & h_{w,r}^{k} \geq 0, & \forall w,k,r,
\end{align*}
\]

(5.1)

(5.2)

where \(\eta_{i}^{i,k}\) and \(\mu_{w,r}^{k}\) are the generalized travel cost for class \(i\) and the travel time for mode \(k\) between O-D pair \(w\) at equilibrium.

In this paper, the following bilevel programming model is proposed to describe the combined equilibrium assignment through the multimodal transportation network.

The upper-level problem is to find \(\tilde{q} \in \Omega = \{q \mid \sum_{k} q_{i,k}^{i,k} = q_{i}^{i,k}, q_{i,k}^{i,k} \geq 0, \forall w,i,k\}\) such that

\[
\sum_{w} \sum_{i} \sum_{k} \left\{ a^{i} \cdot \mu_{w}^{k} (\tilde{q}) + \beta^{i} \cdot \tau_{w}^{k} \right\} \times \left( q_{i,k}^{i,k} - \tilde{q}_{i,k}^{i,k} \right) \geq 0,
\]

(5.3)

where \(q\) is the vector of \(q_{i,k}^{i,k}\); the function \(\mu_{w}^{k}(q)\) is decided by the following lower-level problem. The lower-level problem is to find

\[
\tilde{x}(q) \in \Psi = \left\{ x \mid \sum_{r} h_{w,r}^{k} = \sum_{i} q_{i,k}^{i,k}, x_{a}^{k} = \sum_{w} \sum_{r} h_{w,r}^{k} \delta_{a,r}^{k}, h_{w,r}^{k} \geq 0, \forall w,k,r,a \right\}
\]

(5.4)

such that

\[
\sum_{a} \sum_{k} \tilde{f}_{a}^{k} (\tilde{x}(q)) \times \left( x_{a}^{k} - \tilde{x}_{a}^{k}(q) \right) \geq 0,
\]

(5.5)

where \(x\) is the vector of \(x_{a}^{k}\).

It can be seen that the variational inequality (VI) model for upper-level problem is to find equilibrium demand of class \(i\) selecting mode \(k\) between O-D pair \(w\), that is, \(q_{i,k}^{i,k}\), to meet the first equilibrium principle in (5.1). The travelers’ generalized costs are partially decided by the equilibrium flow patterns and the corresponding travel time through the different subnets. The relationship between them is described by the lower-level VI model with parameters in (5.5). The lower-level problem represents the equilibrium assignment reflecting travelers’ route choice behaviors within each subnet, and the goal is to find the equilibrium flows and corresponding travel time under the condition that the demands
of different classes and selected different modes are all given. The variables $q$ and $x$ can be regarded as decision variables for the bilevel problem. The biequilibrium for the urban multimodal network can be achieved by solving the bilevel problem.

The equivalence between the solution to the previous model and the equilibrium conditions for multimodal transportation network is given as follows.

Assuming that $\tilde{q} \in \Omega$ is a solution to VI problem in (5.3), then $\tilde{q}$ is bound to meet the following conditions:

1. $\tilde{q}_{iw}^k (c_{iw}^k - \eta_{iw}^k) = 0$, $\forall w, k, i$, (5.6a)
2. $c_{iw}^k - \eta_{iw}^k \geq 0$, $\forall w, k, i$, (5.6b)

where $\eta_{iw}^k$ is the dual multiplier of the constraint condition in (4.1).

Similarly, assuming that $\tilde{x} \in \Psi$ is a solution to VI problem in (5.5), then $\tilde{x}$ is bound to meet the following conditions:

1. $h_{iw,r}^k (t_{iw,r}^k - \mu_{iw}^k) = 0$, $\forall w, k, r$, (5.7a)
2. $t_{iw,r}^k - \mu_{iw}^k \geq 0$, $\forall w, k, r$, (5.7b)

where $\mu_{iw}^k$ is the dual multiplier of the constraint condition in (4.6).

Obviously, the first equilibrium condition in (5.1) can be gotten from (5.6a) and (5.6b), and the second equilibrium condition in (5.2) can be gotten by (5.7a) and (5.7b).

6. Solution Algorithm

Due to the intrinsic complexity of model formulation, the bilevel programming problem has been recognized as one of the most difficult, yet challenging, problems for global optimality in transportation system. In the past decades, researchers [35, 36, 39–42] developed alternative solution algorithms for this problem. The sensitivity analysis-based method proposed by Tobin and Friesz [43] is used to solve the bilevel programming model proposed in this paper.

It is necessary to derive the derivatives of the decision variables with respect to the parameters for the lower-level problem in the sensitivity analysis approach. In our proposed problem, we need to calculate the derivatives of the optimal dual multiplier of the constraint condition in (4.6), that is, the equilibrium of O-D travel time (represented by $\mu$), with respect to the travel demand (represented by $q$). By assuming that the initial $q^{(0)}$ is given and other conditions are fixed, the equilibrium O-D travel time matrix for a multimodal traffic network, $\tilde{\mu}(q^{(0)})$ can be obtained by solving the lower level of the model. Through conducting a sensitivity analysis of VI model in (5.5) (see appendix), the approximate differential coefficient, $\nabla_q \mu$, can be obtained. Then the response function can be approximated by the Taylor expansions. That is,

$$\mu(q) \approx \tilde{\mu}(q^{(0)}) + (\nabla_q \mu)^T (q - q^{(0)}).$$ (6.1)

By substituting (6.1) into the upper-level problem, the whole optimization model can be simplified as one-level optimization problem. The solution of this one-level optimization
will then be input into the lower level of the model to run the next iteration. By repeating the iteration process, it is possible to obtain an optimum solution for the above bilevel programming model. This process can be summarized as the following steps.

**Step 1.** Set the initial value \( q^{(0)} \), and set the number of iterations to \( i = 1 \).

**Step 2.** Find the solution of the lower-level model, \( \bar{\mu}^{(i)} \).

**Step 3.** Find the linear equation of the matrix, \( \mu(q) \), through sensitivity analysis and Taylor expansion.

**Step 4.** Put the linear equation of the matrix into the upper-level model to update the value of \( q^{(i)} \) by solving upper-level problem.

**Step 5.** Examine the convergence. If \( q^{(i)} \approx q^{(i-1)} \) or \( i = N \), then iteration stops, where \( N \) is the maximum number of iterations. Otherwise, set \( i = i + 1 \) and start a new iteration.

Note that both Steps 2 and 4 solve different VI models. The approach most commonly used to solve VI model is the popularly known “diagonalization” method, which mimics the Jacobi (resp., Gauss-Seidel) decomposition approach used for solving systems of equations [44]. The idea behind the method is to fix flows for all but one group of variables and to iteratively solve a sequence of separable subproblems which can be described as mathematical programs. As for VI model in (5.3), the vector function \( f^i_w(q) \) is “diagonalized” by the current solution in \( n \)th iteration, yielding a symmetric assignment problem, which can be represented by the following mathematical program:

\[
\min_{q \in \Omega} Z(q) = \sum_{w} \sum_{k} \sum_{i} \int_{0}^{q_{w}} \epsilon^{i,k}(w) \left( q^{1,1}_{w(n-1)}, \ldots, q^{1,k}_{w(n-1)}, \ldots, q^{i,k}_{w(n-1)}, \ldots, \omega \right) d\omega. \tag{6.2}
\]

Similarly, as for VI model in (5.5), the vector function \( f^i_a(x) \) is “diagonalized” at the current solution, yielding the following mathematical program:

\[
\min_{x \in \Psi} F(x) = \sum_{a} \sum_{k} \int_{0}^{x_{a}} f^{i,k}(a) \left( x_{a(n-1)}, x^{2}_{a(n-1)}, \ldots, \omega \right) d\omega. \tag{6.3}
\]

The Frank-Wolfe method or MSA method can be employed to solve the diagonalization problem (6.2) and (6.3). Due to the limited space, the detailed process MSA method for (6.3) is given as follows here.

**Step 1.** Initialization: set \( x_{a}^{k} = 0 \) and compute \( f^{k(0)}_{a} \) for any \( k \) and \( a \). Find the shortest route in subnet \( k \) between O-D pair \( w \). Then perform all-or-nothing assignment to load \( q_{w}^{k} \) for subnet \( k \) and obtain \( x_{a}^{k(1)} \) for any \( k \) and \( a \). Set iteration \( n = 1 \).

**Step 2.** Compute \( f^{k(n)}_{a} \) based on \( x_{a}^{k(n)} \).

**Step 3.** Find the shortest route in subnet \( k \) between O-D pair \( w \). Perform all-or-nothing assignment to load \( q_{w}^{k} \) and obtain \( y_{a}^{k(n)} \) for any \( k \) and \( a \).
Step 4. Compute

\[
x_a^{k(n+1)} = x_a^{k(n)} + \frac{1}{n} \left( y_a^{k(n)} - x_a^{k(n)} \right), \quad \forall k, a.
\]  

(6.4)

Step 5. Convergence test: if a convergence criterion is met, stop. The current solutions, \( \{x_a^{k(n+1)}\} \), are the sets of equilibrium solutions; otherwise, set \( n = n + 1 \) and go to Step 2.

7. Numerical Example

A simple numerical example is used to illustrate the effectiveness of the proposed model and algorithm. The multimodal transportation system and the corresponding hierarchical network structure are, respectively, given by Figures 1 and 4.

The following impedance functions are used in this example [18, 19]:

\[
t_a^k = t_a^{(0)} \prod_m \left[ 1 + \gamma \left( \frac{U_m \cdot x_a^m}{A_m \cdot C_a^m} \right)^\lambda \right], \quad \forall k, a.
\]  

(7.1)

The relevant data of different roads are given in Table 2, where the PCU conversion coefficient, the average occupancy rate, and potential expense, which are pertinent to different modes, are illustrated in Table 3.

The values of \( \gamma = 0.15, \lambda = 4 \) are set for the parameters in (7.1). In this example, the travelers are divided into two classes: (i) for the first class, \( \alpha^1 = 2.5 \) and \( \beta^1 = 0.5 \) indicate that this class is sensitive to the travel time; (ii) for the second class, \( \alpha^2 = 1.5 \) and \( \beta^2 = 0.5 \) indicate that this class is sensitive to the potential expense. The demands of these two classes are all assumed as 5000/Ph⁻¹.

The convergences of the diagonalization method for the lower-level problem and the upper-level problem are, respectively, analyzed using the gap measure proposed by Boyce et al. [45]. The gaps at iteration \( n \) for the assignment models can be defined as

\[
gap(n) = - \sum_a \sum_k t_a^k \cdot (y_a^k(n) - x_a^k(n)),
\]

\[
gap(n) = - \sum_{w} \sum_i \sum_k c_{w}^{i,k} \cdot (v_{w}^{i,k}(n) - q_{w}^{i,k}(n)),
\]  

(7.2)

where \( y_a^k(n) \) is the auxiliary flow of mode \( k \) on link \( a \) at iteration \( n \) given by an all-or-nothing assignment based on link travel time, \( t_a^{k(n)} \), and \( v_{w}^{i,k}(n) \) is the auxiliary demand of class \( i \) selecting mode \( k \) between O-D pair \( w \) at iteration \( n \) given by an all-or-nothing assignment based on the generalized costs, \( c_{w}^{i,k}(n) \).

Figure 5 shows the gaps against the iteration number for the lower-level problem and upper-level problem, respectively. It can be seen that the solution algorithm has a good convergence especially for the upper-level problem. It can be explained that the network structure of the traveler’s mode choice in the upper-level problem is simpler than that of the traveler’s route choice in the lower-level problem.
Table 2: The relevant data of different roads.

<table>
<thead>
<tr>
<th>Road</th>
<th>$t_{a1}^{(1)}/(h)$</th>
<th>$t_{a2}^{(2)}/(h)$</th>
<th>$t_{a3}^{(3)}/(h)$</th>
<th>$C_{a1}^2/(Ph^{-1})$</th>
<th>$C_{a2}^2/(Ph^{-1})$</th>
<th>$C_{a3}^2/(Ph^{-1})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1,2)</td>
<td>0.111</td>
<td>0.178</td>
<td>0.361</td>
<td>1000</td>
<td>1000</td>
<td>600</td>
</tr>
<tr>
<td>(2,3)</td>
<td>0.128</td>
<td>—</td>
<td>0.378</td>
<td>700</td>
<td>—</td>
<td>400</td>
</tr>
<tr>
<td>(1,4)</td>
<td>0.100</td>
<td>0.167</td>
<td>0.350</td>
<td>1500</td>
<td>1500</td>
<td>800</td>
</tr>
<tr>
<td>(2,5)</td>
<td>0.106</td>
<td>0.172</td>
<td>0.356</td>
<td>700</td>
<td>700</td>
<td>400</td>
</tr>
<tr>
<td>(3,6)</td>
<td>0.089</td>
<td>—</td>
<td>0.339</td>
<td>700</td>
<td>—</td>
<td>400</td>
</tr>
<tr>
<td>(4,5)</td>
<td>—</td>
<td>0.144</td>
<td>0.328</td>
<td>—</td>
<td>1000</td>
<td>600</td>
</tr>
<tr>
<td>(5,6)</td>
<td>—</td>
<td>—</td>
<td>0.344</td>
<td>—</td>
<td>—</td>
<td>600</td>
</tr>
<tr>
<td>(4,7)</td>
<td>0.133</td>
<td>0.200</td>
<td>0.383</td>
<td>900</td>
<td>900</td>
<td>500</td>
</tr>
<tr>
<td>(5,8)</td>
<td>0.111</td>
<td>0.178</td>
<td>0.361</td>
<td>700</td>
<td>700</td>
<td>400</td>
</tr>
<tr>
<td>(6,9)</td>
<td>0.144</td>
<td>—</td>
<td>0.394</td>
<td>700</td>
<td>—</td>
<td>400</td>
</tr>
<tr>
<td>(7,8)</td>
<td>0.094</td>
<td>0.161</td>
<td>0.344</td>
<td>900</td>
<td>900</td>
<td>500</td>
</tr>
<tr>
<td>(8,9)</td>
<td>0.100</td>
<td>0.167</td>
<td>0.350</td>
<td>900</td>
<td>900</td>
<td>500</td>
</tr>
</tbody>
</table>

Table 3: The relevant data of different modes.

<table>
<thead>
<tr>
<th>Mode</th>
<th>$U_k$</th>
<th>$A_k$</th>
<th>$\tau_{w}^k$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Car</td>
<td>1</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>Bus</td>
<td>1.5</td>
<td>20</td>
<td>4</td>
</tr>
<tr>
<td>Bike</td>
<td>0.25</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 4 shows the equilibrium results of mode demand and the corresponding generalized costs of different classes. Table 5 shows the equilibrium results of road demand and the corresponding travel time of different modes.

Next, we analyze the impacts of the pertinent parameters in this example on the modal share and the performance of the whole network. Here, the shares of different modes, denoted by $P_{w}^k$, can be computed by

$$P_{w}^k = \frac{q_{w}^k}{\sum_i q_{w}^i}, \quad \forall w, k.$$  \hfill (7.3)

The total travel time of the network, denoted by $T$, is used to represent the performance of the whole network, that is:

$$T = \sum_a \sum_k \frac{U_k \cdot \lambda^k}{A_k} \cdot t_{a}^k.$$  \hfill (7.4)

Figures 6(a)–6(d), respectively, show the changes in modal share and the total travel time of the whole network with the changes of the parameters ($\alpha^1, \alpha^2$), which indicate the travelers' sensitivity to the factor of travel time (or congestion). It shows that the share of bike and the total travel time of whole network will decrease while the shares of bus and car will increase symmetrically with the increasing of $\alpha^1$ or $\alpha^2$. The travelers who select the bike mode with longer travel time will shift into the car or bus mode when such travelers become more sensitive to travel time.

Figures 7(a)–7(d), respectively, display the changes in modal share and the total travel time with the changes of the parameters ($\beta^1, \beta^2$), which indicate the travelers' sensitivity to the potential travel expense. It can be found that the share of bike and the total travel time
Figure 5: The convergences of the algorithms for upper problem (a) and lower problem (b).

Table 4: The equilibrium results of demand and the corresponding costs of different classes.

<table>
<thead>
<tr>
<th>Roads</th>
<th>Mode 1 (car)</th>
<th>Mode 2 (bus)</th>
<th>Mode 3 (bike)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Class 1</td>
<td>Class 2</td>
<td>Class 1</td>
</tr>
<tr>
<td>Demand/Ph^-1</td>
<td>609</td>
<td>284</td>
<td>1809</td>
</tr>
</tbody>
</table>

Table 5: The equilibrium results of road demands and its travel time of different modes.

<table>
<thead>
<tr>
<th>Roads</th>
<th>x_i^1 (Ph^-1)</th>
<th>x_i^2 (Ph^-1)</th>
<th>x_i^3 (Ph^-1)</th>
<th>t_i^1 (h)</th>
<th>t_i^2 (h)</th>
<th>t_i^3 (h)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1,2)</td>
<td>166.72</td>
<td>0</td>
<td>2148.55</td>
<td>0.1147</td>
<td>0.1835</td>
<td>0.3960</td>
</tr>
<tr>
<td>(2,3)</td>
<td>0</td>
<td>—</td>
<td>900.00</td>
<td>0.1284</td>
<td>—</td>
<td>0.3835</td>
</tr>
<tr>
<td>(1,4)</td>
<td>726.28</td>
<td>3493.00</td>
<td>3465.45</td>
<td>0.1070</td>
<td>0.1783</td>
<td>0.4225</td>
</tr>
<tr>
<td>(2,5)</td>
<td>166.72</td>
<td>0</td>
<td>1248.55</td>
<td>0.1075</td>
<td>0.1754</td>
<td>0.3754</td>
</tr>
<tr>
<td>(3,6)</td>
<td>0</td>
<td>—</td>
<td>900.00</td>
<td>0.0893</td>
<td>—</td>
<td>0.3440</td>
</tr>
<tr>
<td>(4,5)</td>
<td>—</td>
<td>3443.00</td>
<td>1940.48</td>
<td>—</td>
<td>0.1476</td>
<td>0.3489</td>
</tr>
<tr>
<td>(5,6)</td>
<td>—</td>
<td>—</td>
<td>1572.41</td>
<td>—</td>
<td>—</td>
<td>0.3540</td>
</tr>
<tr>
<td>(4,7)</td>
<td>726.28</td>
<td>50.00</td>
<td>1524.97</td>
<td>0.1356</td>
<td>0.2034</td>
<td>0.4028</td>
</tr>
<tr>
<td>(5,8)</td>
<td>166.72</td>
<td>3443.00</td>
<td>1616.62</td>
<td>0.1175</td>
<td>0.1880</td>
<td>0.4183</td>
</tr>
<tr>
<td>(6,9)</td>
<td>0</td>
<td>—</td>
<td>2472.41</td>
<td>0.1858</td>
<td>—</td>
<td>0.7331</td>
</tr>
<tr>
<td>(7,8)</td>
<td>726.28</td>
<td>50.00</td>
<td>1524.97</td>
<td>0.0961</td>
<td>0.1639</td>
<td>0.3619</td>
</tr>
<tr>
<td>(8,9)</td>
<td>893.00</td>
<td>3493.00</td>
<td>3141.59</td>
<td>0.1321</td>
<td>0.2202</td>
<td>0.6725</td>
</tr>
</tbody>
</table>
will increase while the shares of bus and car will decrease with the increment of $\beta^1$ or $\beta^2$. The travelers who select the bus or car mode will shift into the bike mode without any potential expense when such travelers become more sensitive to the potential expense. The previous results imply that the performance of the whole network in terms of total travel time would be better when the travelers are more sensitive to the travel time and less sensitive to potential expense.

The shares of various modes and the total travel time can be dramatically changed with the changes of $(\alpha^1, \alpha^2)$ or $(\beta^1, \beta^2)$ in a certain range. However, these values will not change significantly when these parameters reach a certain value. When the travelers are all excessive time-sensitive or cost-sensitive, the other factor can affect their mode choice behaviors to a very small extent. For example, when the travelers are very sensitive to the potential expense, they would not consider the factor of travel time. In such a condition, most of the travelers would choose bike as their traffic tools since they do not bear any costs. As the speed of bike is slowest, the total cost of network will reach the maximum. On the contrary, when the traveler is very sensitive to the travel time, they will not consider the factor of money. Therefore, the travelers always tend to choose the mode with the shortest travel time (such as car). Simultaneously, the travel time of such mode will become longer.

Figure 6: The changes of shares of different modes and the total travel time with the changes of $(\alpha^1, \alpha^2)$. 
and longer with its increasing demand. The equilibrium between different modes will be achieved ultimately, and the shares and the total travel time of the whole network will not be changed at such equilibrium.

Assuming that the total number of travelers between O-D remains unchanged (take the value of 10000 persons each hour), Figures 8(a) and 8(b), respectively, show the change trends of the shares and the total travel time with the proportion of class I which is sensitive to the travel time. It can be shown that the share of bike mode will decrease and the share of car will go up slightly, while the share of bus remained unchanged. Meanwhile, the total travel time of whole network will increase with the increment of the proportion of class I. The results also imply that in the multiclass multimodal transportation network, the more travelers who focus on the factor of travel time, the lower the total travel time of network is.

8. Conclusions

This paper presents a biequilibrium traffic assignment model for multimodal transportation networks using the bilevel programming method. The model development is based on several important concepts that are not explored by the previous studies.
First, a two-tier hierarchical multimodal network is proposed for the model, in which the first-tier network is used for mode choice and the second-tier network is used to for route choice in the single mode subnets.

Second, the model distinguishes the criteria between mode choice and route choice. The mode choice behavior is based on the multiclass generalized travel cost while the route choice behavior is based on the travel time only. The generalized cost functions of different modes and the link impedance functions are formulated while the interferences between different modes are considered. The approach can better reflect traveler’s preference and decision-making process in a multimodal transportation system.

Third, the biequilibrium pattern of traffic assignment is firstly proposed for multimodal traffic network modeling. Its major advantage is integrating the separated two steps of mode split and traffic assignment in the traditional transportation planning method into a unified process.

The solution algorithm for the bilevel programming model is illustrated by a simple numerical example. The sensitivity analysis shows that as travelers are more sensitive to travel time, they are more likely to choose the mode with less travel time, which will mitigate the congestion of whole network. In contrast, as travelers are more sensitive to travel expense, they are more likely to choose the mode without expense, such as bike, which will aggravate the congestion of whole network. As for the travelers who are more sensitive to travel time, the changes of their choice behaviors will impact on the performance of whole network.
significantly. Additionally, with the increment of the proportion of traveler class that is more sensitive to travel time, the network congestion will be mitigated gradually.

It should be noted that there are some limitations in this paper. For example, all travelers are assumed to complete their trips through only one mode; in other words, they are assumed not to change modes during their journey. Such reasonable assumption will preclude the possibility of park-and-ride or similar mode-change mechanisms. In addition, the case of fixed demand is considered, while the case of elastic demand or demand uncertainty is not taken into account. So, the promising future work would be extension to reliability analysis for these situations.

Appendix

Sensitivity Analysis of VI Model (5.5)

Assume that the solutions to \( \tilde{x}(q^{(0)}) \) and \( \mu(q^{(0)}) \) of the variational inequality problem in (5.5) at \( q = q^{(0)} \) have been obtained and that \( f_k^a(x) \) is strongly monotone in \( x \), so that the solutions are unique. According to Tobin and Friesz [43], the necessary conditions excluding the nonbinding constraints for solution at \( q = q^{(0)} \) for VI problem in (5.5) can be expressed as follows:

\[
\sum_a \tilde{f}_k^a(x) \cdot \delta_{w,r}^{k,w} - \tilde{\mu}_w^k = 0, \quad \forall w, k, r,
\]

\[
\sum_r \tilde{h}_{w,r}^k - \sum_i q_i^k w = 0, \quad \forall w, k.
\]  \( \text{(A.1)} \)

Let \( y = [h, \mu]^T \), where \( h \) is the vector of \( h_{w,r}^k \). Let \( J_y \) and \( J_q \) denote the Jacobian matrixes of (A.1) with respect to \( y \) and \( q \) at the point \( q = q^{(0)} \), respectively:

\[
J_y = \begin{bmatrix}
\nabla h^T & \Lambda^T \\
\Lambda & 0
\end{bmatrix},
\]  \( \text{(A.2)} \)

where \( t \) is the vector of \( t_{w,r}^k \), and \( \Lambda \) is the O-D and path incidence matrix. Suppose

\[
[J_y]^{-1} = \begin{bmatrix}
B_{11} & B_{12} \\
B_{21} & B_{22}
\end{bmatrix}.
\]  \( \text{(A.3)} \)

The following results can be obtained:

\[
B_{22} = \left[ \Lambda \cdot \nabla h^{-1} \cdot \Lambda^T \right]^{-1},
\]

\[
B_{12} = \nabla h^{-1} \cdot \Lambda^T \cdot B_{22} = \nabla h^{-1} \cdot \Lambda^T \left[ \Lambda \cdot \nabla h^{-1} \cdot \Lambda^T \right]^{-1},
\]

\[
B_{21} = -B_{22} \cdot \Lambda \cdot \nabla h^{-1} = -\left[ \Lambda \cdot \nabla h^{-1} \cdot \Lambda^T \right]^{-1} \cdot \Lambda \cdot \nabla h^{-1},
\]  \( \text{(A.4)} \)

\[
B_{11} = \nabla h^{-1} \cdot \left[ I + \Lambda^T \cdot B_{21} \right] = \nabla h^{-1} \cdot \left[ I - \Lambda^T \cdot \left[ \Lambda \cdot \nabla h^{-1} \cdot \Lambda^T \right]^{-1} \cdot \Lambda \cdot \nabla h^{-1} \right],
\]
where $I$ is unit matrix:

$$J_q = \begin{bmatrix} \nabla_q t \\ -I \end{bmatrix}. \quad (A.5)$$

From theorems in Tobin and Friesz [43], the following result can be obtained:

$$\begin{bmatrix} \nabla_q h \\ \nabla_q \mu \end{bmatrix} = (J_q)^{-1} \cdot \begin{bmatrix} -J_q \\ -\begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \begin{bmatrix} \nabla_q t \\ -I \end{bmatrix} \end{bmatrix}. \quad (A.6)$$

Thus, the approximate differential coefficient, $\nabla_q \mu$, can be obtained:

$$\nabla_q \mu = -B_{21} \cdot \nabla_q t + B_{22} = \begin{bmatrix} \Lambda \cdot \nabla_h t^{-1} \cdot \Lambda^T \end{bmatrix}^{-1} \cdot \Lambda \cdot \nabla_h t^{-1} \cdot \nabla_q t + \left( \Lambda \cdot \nabla_h t^{-1} \cdot \Lambda^T \right)^{-1}. \quad (A.7)$$

**Acknowledgments**

The work described in this paper is mainly supported by the grants from the National Natural Science Foundation of China (Project nos. 71071016 and 71131001) and the National Basic Research Program of China (Project nos. 2012CB725400 and 2012JBZ005).

**References**


