A Two-Phase Gradient Projection Algorithm for Solving the Combined Modal Split and Traffic Assignment Problem with Nested Logit Function

Seungkyu Ryu, Anthony Chen, and Songyot Kitthamkesorn

1 Korea Institute Science Technology Information, Daehak-ro, Yusong-gu, Daejeon, Republic of Korea
2 Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong
3 Department of Civil Engineering, Chiang Mai University, 239 Huay Kaew Road, Chiang Mai 50200, Thailand

Correspondence should be addressed to Seungkyu Ryu; ryuseungkyu@gmail.com

Received 25 February 2020; Revised 2 March 2021; Accepted 3 April 2021; Published 19 April 2021

Copyright © 2021 Seungkyu Ryu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This study provides a gradient projection (GP) algorithm to solve the combined modal split and traffic assignment (CMSTA) problem. The nested logit (NL) model is used to consider the mode correlation under the user equilibrium (UE) route choice condition. Specifically, a two-phase GP algorithm is developed to handle the hierarchical structure of the NL model in the CMSTA problem. The Seoul transportation network in Korea is adopted to demonstrate an applicability in a large-scale multimodal transportation network. The results show that the proposed GP solution algorithm outperforms the method of the successive averages (MSA) algorithm and the classical Evan’s algorithm.

1. Introduction

Combined travel demand models (or combined models) have been used as a network equilibrium approach to resolve the inconsistency problem (i.e., flow) in the traditional four-step travel demand model (see, e.g., [1]). Several combined models were provided in the literature including the combined distribution and assignment (CDA) model [2–6], combined trip distribution, modal split and traffic assignment model [7–9], combined modal split and traffic assignment (CMSTA) problem [8, 10–13], and combined trip generation, trip distribution, modal split and traffic assignment model [2, 14, 15].

Various formulations, ranging from mathematical program (MP), nonlinear complementarity problem (NCP), variational inequality (VI), and fixed point (FP), have been provided to represent different combined models listed above. Please refer to [6, 12, 16, 17] for a review of the different formulation approaches for modeling CDA, CMSTA, and combined travel demand models. These combined models have been adopted to represent different emerging technological applications, such as mixed gasoline and electric vehicles with destination, route, and parking choices [16], electric vehicle charging stations using the CDA model [17], and ridesharing as a new mode choice option using the CMSTA model [18]. Most of the above studies adopted either the complete linearization method of the Frank–Wolfe algorithm [19] or the partial linearization method of Evan’s algorithm [2] for solving these emerging technological applications. Computational results conducted by LeBlanc and Farhangian [20] revealed Evan’s partial linearization method performed better than Frank–Wolfe’s complete linearization method, while Ryu et al. [21] recently demonstrated the superiority of the gradient projection (GP) algorithm over Evan’s algorithm. Note that very few studies have focused on developing solution algorithms for solving large-scale problems with multiple modes in a multimodal transportation network [24–26].

For solving traffic assignment problems, numerous algorithms were developed based on link-based, path-based, and bush-based algorithms. The link-based algorithms operate in the space of link flows and do not require path storage, but it requires more computational efforts comparing to other groups of algorithms. On the other hand, path-based algorithms require explicit path storage in order to directly...
compute path flow equilibration and they have faster convergence comparing to link-based algorithms. The reason is that the link-based algorithm exhibits a zigzagging trajectory as it approaches the optimal solution. On the other hand, the path-based algorithm does not exhibit the zigzagging effect for the same initial solution. In the transportation literature [26, 27], the computational performance of link- and path-based algorithms for the traffic assignment problem is compared in medium- to large-scale randomly generated grid networks and real-size networks. Based on the conclusion, a path-based algorithm is consistently faster in approaching the neighborhood of the optimal solution. Bush-based methods are developed to solve an acyclic network problem. These methods have improved the existing state-of-the-art of algorithms and are fast and memory efficient. However, using the algorithms for solving large-scale networks may remain impractical [28]. The solution algorithms for the traffic assignment model can be summarized below. Interested readers may refer to [27, 29] for the detailed descriptions of the algorithm procedure.

(i) **Link-Based Algorithm.** This includes search direction improvement [30]; restricted simplicial decomposition [31]; and nonlinear simplicial decomposition [32]. Mitradjieva and Lindberg [30] introduced conjugate and biconjugate algorithms. These algorithms improve the search direction by using previous direction information. Restricted simplicial decomposition and nonlinear simplicial decomposition algorithms are represented by a convex combination of extreme points. They first generate new extreme points and then optimize the restricted master problem.

(ii) **Path-Based Algorithm.** This includes disaggregated simplicial decomposition [32]; gradient projection [21, 33–35]; projected gradient [36]; improved social pressure [37]; greedy algorithm [38, 39]; and step size improvement [40].

The disaggregated simplicial decomposition algorithm proposed by Larsson and Patriksson [41] is similar to the restricted simplicial decomposition algorithm. However, the extreme points are generated by path flows instead of link flows. The gradient projection algorithm adjusts the O-D demand between nonshortest paths and the shortest path. The projected gradient algorithm is based on the idea of adjusting flow between the set of path with a cost greater than the average path cost and the set of path with a cost less than the average path cost. Improved social pressure is also using the path flow equilibrium between costlier paths and cheaper paths, but the flow updating strategy is more complicated. To solve the restricted subproblem, a greedy algorithm is introduced. Recently, Du et al. [40] introduced Barzilai–Borwein step size to adjust path flows.

(iii) **Bush-Based Algorithm.** This includes origin-based algorithm [42, 43]; modified origin-based algorithm [44]; algorithm B [45]; linear user cost equilibrium [46]; and paired alternative segments [47].

The bush-based algorithm sequentially solves a list of node-based subproblems defined on the bush rooted. The algorithm B adjusts flows between the longest paths and the shortest ones in the bush rooted path tree. A paired alternative segment (PAS) is defined as two completely disjoint path segments, and then the flows are adjusted in disjoint path segments.

In this paper, we are interested in developing solution algorithms for solving the MP formulation of the CMSTA problem that can represent large-scale multimodal transportation networks with both private and public modes and correlation among multiple public modes using the nested logit (NL) model under the network user equilibrium (UE) framework. Contrast to the simple structure of the multinomial logit (MNL) model, the NL model has a hierarchical structure that decomposes the choice probability into two levels represented by the marginal and conditional probabilities. A two-phase GP algorithm is developed to solve the CMSTA problem with a hierarchical nested modal split structure and UE traffic assignment, along with various improvement strategies to speed up the convergence of the path-based GP algorithm.

In the literature, a path-based GP algorithm has been adopted for solving various traffic assignment problems such as UE assignment problem, the nonadditive cost assignment problem, side constrained problem, stochastic user equilibrium assignment problem, transit assignment problem, system optimal assignment problem, and freight traffic assignment problem [21]. Although the algorithm was developed for solving various traffic assignment problems, the algorithm was rarely applied for solving CMSTA problem.

Our goal is to demonstrate that the improved GP algorithm can solve large-scale multiclass or multimodal traffic assignment problems in CMSTA with NL function. In addition, the proposed algorithm can apply to other logit models having a hierarchical structure such as the generalized extreme value model by [48] including the cross-nested logit (CNL), paired combinatorial logit (PCL), and generalized nested logit (GNL). To show proof of concept, a large-scale multimodal transport network in Seoul, Korea, will be used to demonstrate the applicability of the two-phase GP algorithm.

This paper is organized as follows. After the introduction, a relevant background of the NL model for the modal split problem and UE conditions for the traffic assignment problem is introduced in Section 2. Section 3 provides an equivalent MP formulation for the NL-UE CMSTA problem along with the equivalence property. Section 4 describes the two-phase GP algorithm for solving the NL-UE CMSTA problem, improvement strategies, and its overall solution procedure. Section 5 provides numerical results to demonstrate the applicability of the two-phase GP algorithm for solving large-scale multimodal transportation networks. Some concluding remarks are provided in Section 6.

## 2. Mode Choice and Route Choice

In this section, we provide some background of the nested logit (NL) model for the modal split problem and the UE
model for the traffic assignment problem. A list of variable is provided first for convenience, followed by the NL model and multiclass UE model.

2.1. List of Variable. Consider a transportation network \( G = (N, A) \), where \( N \) is the set of nodes and \( A \) is the set of links on a given network. We define the list of variables in Table 1.

2.2. Nested Logit Model for Mode Choice. The NL model is widely used to represent mode choice in the literature (see, e.g., [49–51]). It considers the mode similarity through a tree structure as shown in Figure 1, where each mode \( m \in M_u^r \) may share an upper nest \( u \in U^rs \).

The NL mode choice probability can be expressed as

\[
P_{m}^{rs}(NL) = \frac{\exp\left(\left(\frac{q_{rs}}{T_{rs}^{u}} - \frac{\mu_{um}}{T_{um}^{rs}}\right) \left(\psi_{um}^{rs} - \mu_{um}^{rs}\right)\right)}{\sum_{m \in M_u^{rs}} \exp\left(\left(\frac{q_{rs}}{T_{rs}^{u}} - \frac{\mu_{um}}{T_{um}^{rs}}\right) \left(\psi_{um}^{rs} - \mu_{um}^{rs}\right)\right)^{T_{rs}^{u} - 1}}, \quad \forall m \in M_u^{rs}, \quad rs \in RS
\]

With the tree structure, the NL probability can be decomposed as

\[
P_{m}^{rs} = \prod_{m \in M_u^{rs}} p_{m}^{rs}, \quad \forall m \in M_u^{rs}, \quad rs \in RS
\]

\[
p_{m}^{rs} = \frac{\exp\left(\left(\frac{q_{rs}}{T_{rs}^{u}} - \frac{\mu_{um}}{T_{um}^{rs}}\right) \left(\psi_{um}^{rs} - \mu_{um}^{rs}\right)\right)}{\sum_{m \in M_u^{rs}} \exp\left(\left(\frac{q_{rs}}{T_{rs}^{u}} - \frac{\mu_{um}}{T_{um}^{rs}}\right) \left(\psi_{um}^{rs} - \mu_{um}^{rs}\right)\right)}, \quad \forall m \in M_u^{rs}, \quad u \in U^{rs}, \quad rs \in RS
\]

\[
p_{u}^{rs} = \frac{\left[\sum_{m \in M_u^{rs}} \exp\left(\left(\frac{q_{rs}}{T_{rs}^{u}} - \frac{\mu_{um}}{T_{um}^{rs}}\right) \left(\psi_{um}^{rs} - \mu_{um}^{rs}\right)\right)^{T_{rs}^{u}}\right]}{\sum_{m \in M_u^{rs}} \exp\left(\left(\frac{q_{rs}}{T_{rs}^{u}} - \frac{\mu_{um}}{T_{um}^{rs}}\right) \left(\psi_{um}^{rs} - \mu_{um}^{rs}\right)\right)^{T_{rs}^{u}}}, \quad \forall u \in U^{rs}, \quad rs \in RS
\]

Note that if the parameter \((T_{rs}^{u})\) is equal to 1.0, the NL model reduces to the MNL model.

2.3. Multiclass User Equilibrium for Route Choice. For the road-based transportation mode such as passenger car and bus, we have the Karush–Kuhn–Tucker (KKT) conditions at equilibrium as follows:

\[
f_{km}^{rs} (c_{km}^{rs} - \pi_{km}^{rs}) = 0, \quad \forall m \in M_u^{rs}, \quad rs \in RS, \quad k \in K_{rs}^m
\]

\[
c_{km}^{rs} - \pi_{km}^{rs} \geq 0, \quad \forall m \in M_u^{rs}, \quad rs \in RS, \quad k \in K_{rs}^m
\]

\[
f_{km}^{rs} \geq 0, \quad \forall m \in M_u^{rs}, \quad rs \in RS, \quad k \in K_{rs}^m
\]

Equations (3) and (4) define a user equilibrium flow pattern in the route choice [52].

\[
\min Z = Z_1 + Z_2 + Z_3 + Z_4 + Z_5 = \sum_{a \in A} \int_{0}^{T_a} t_a(\omega) d\omega + \sum_{rs \in RS} \frac{1}{\theta^T} \left(\ln \left(\frac{\omega}{\theta^T} - \omega\right)\right) d\omega + \sum_{rs \in RS} \sum_{m \in M_u^{rs}} \frac{q_{rs}^m}{\theta^T} q_{rs}^m (\ln q_{rs}^m - 1)
\]

\[
- \sum_{rs \in RS} \frac{T_{rs}^m}{\theta^T} \left(\sum_{m \in M_u^{rs}} q_{rs}^m\right) (\ln \left(\sum_{m \in M_u^{rs}} q_{rs}^m\right) - 1) + \sum_{rs \in RS} \sum_{m \in M_u^{rs}} q_{rs}^m \theta^T c_{rs}^m,
\]

s.t.
<table>
<thead>
<tr>
<th>Notation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>Set of origins, $R \subset \mathbb{N}$</td>
</tr>
<tr>
<td>S</td>
<td>Set of destinations, $S \subset \mathbb{N}$</td>
</tr>
<tr>
<td>RS</td>
<td>Set of O-D pairs rs</td>
</tr>
<tr>
<td>$K_{rs}$</td>
<td>Set of routes connecting O-D pair rs</td>
</tr>
<tr>
<td>$U_{rs}$</td>
<td>Set of upper nests connecting O-D pair rs</td>
</tr>
<tr>
<td>$M_{rs}^u$</td>
<td>Set of lower nests connecting O-D pair rs in upper nest u</td>
</tr>
<tr>
<td>$M_{rs}^T$</td>
<td>Set of transit modes connecting O-D pair rs</td>
</tr>
</tbody>
</table>

**Parameters and inputs**

<table>
<thead>
<tr>
<th>Notation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Psi_{rs}^m$</td>
<td>Given exogenous utility of mode $m$ in nest $u$ connecting O-D pair rs</td>
</tr>
<tr>
<td>$\mu_{rs}^m$</td>
<td>Expected perceived utility of mode $m$ in nest $u$ connecting O-D pair rs</td>
</tr>
<tr>
<td>$\theta^s$</td>
<td>O-D specific parameter</td>
</tr>
<tr>
<td>$\tau_{rs}^u$</td>
<td>Upper nest parameter of each O-D pair rs (i.e., $0 &lt; \tau_{rs}^u &lt; \theta^s$)</td>
</tr>
<tr>
<td>$P_{rs}(m/n)$</td>
<td>Conditional probability</td>
</tr>
<tr>
<td>$P_{rs}^u$</td>
<td>Marginal probability</td>
</tr>
<tr>
<td>$\epsilon_{inner}$</td>
<td>Convergence error in the inner loop</td>
</tr>
<tr>
<td>$\epsilon_{outer}$</td>
<td>Convergence error in the outer loop</td>
</tr>
<tr>
<td>$d_a^B$, $d_a^M$</td>
<td>Dwelling times on link $a$ of bus and metro</td>
</tr>
<tr>
<td>$\eta_u^a$, $\eta_u^T$</td>
<td>Equal to 1 for link $a$ with a bus line, metro line, and bus-metro line and 0 otherwise</td>
</tr>
<tr>
<td>$l_a$</td>
<td>Length on link $a$</td>
</tr>
<tr>
<td>$W_a$</td>
<td>Walking speed (i.e., 4 km/h)</td>
</tr>
<tr>
<td>$W_r$</td>
<td>Weighting time for the transit mode (bus and metro) on link $a$</td>
</tr>
<tr>
<td>$\Phi_{rs}$</td>
<td>Transit fare connecting O-D pair rs</td>
</tr>
<tr>
<td>$P_{rs}$</td>
<td>Inverse value of time for converting transit fare into equivalent time unit</td>
</tr>
<tr>
<td>$\delta_{rs}$</td>
<td>Total travel demand connecting O-D pair rs</td>
</tr>
<tr>
<td>$C_a$</td>
<td>Capacity on link $a$</td>
</tr>
</tbody>
</table>

**Decision variables**

<table>
<thead>
<tr>
<th>Notation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_{rs}^m$</td>
<td>Route flow on the mode $m$ of route $k$ connecting O-D pair rs</td>
</tr>
<tr>
<td>$c_{rs}^m$</td>
<td>Route cost on the mode $m$ of route $k$ connecting O-D pair rs</td>
</tr>
<tr>
<td>$\pi_{rs}^m$</td>
<td>Minimum cost on the mode $m$ connecting O-D pair rs</td>
</tr>
<tr>
<td>$x_k$</td>
<td>Route flow of the private transport mode on route $k$ connecting O-D pair rs</td>
</tr>
<tr>
<td>$c_k$</td>
<td>Route cost of the private transport mode route $k$ connecting O-D pair rs</td>
</tr>
<tr>
<td>$s_k$</td>
<td>Route cost of the private transport mode shortest route connecting O-D pair rs</td>
</tr>
<tr>
<td>$s_{krs}$</td>
<td>Positive-definite scaling factor between route $k$ and the shortest route $k_{rs}$ of O-D pair rs</td>
</tr>
<tr>
<td>$t_a(\cdot)$</td>
<td>Link flow of private transport mode on link $a$</td>
</tr>
<tr>
<td>$\delta_{rs}^a$</td>
<td>Travel time of private transport mode on link $a$</td>
</tr>
<tr>
<td>$q_{rs}^m$</td>
<td>Total demand of public transport mode connecting O-D pair rs</td>
</tr>
<tr>
<td>$g_{rs}^m$</td>
<td>Demand from public transport mode $m$ in the lower nest connecting O-D pair rs</td>
</tr>
<tr>
<td>$\delta_{rs}^{c_{rs}}$</td>
<td>Generalized cost of public transport mode $m$ connecting O-D pair rs</td>
</tr>
</tbody>
</table>

**Abbreviations**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>GP</td>
<td>Gradient projection</td>
</tr>
<tr>
<td>CMSTA</td>
<td>Combined modal split and traffic assignment</td>
</tr>
<tr>
<td>NL</td>
<td>Nested logit</td>
</tr>
<tr>
<td>UE</td>
<td>User equilibrium</td>
</tr>
<tr>
<td>MSA</td>
<td>Method of successive averages</td>
</tr>
<tr>
<td>CDA</td>
<td>Combined distribution and assignment</td>
</tr>
<tr>
<td>MP</td>
<td>Mathematical program</td>
</tr>
<tr>
<td>NCP</td>
<td>Nonlinear complementarity problem</td>
</tr>
<tr>
<td>VI</td>
<td>Variational inequality</td>
</tr>
<tr>
<td>FP</td>
<td>Fixed point</td>
</tr>
<tr>
<td>MNL</td>
<td>Multinomial logit</td>
</tr>
<tr>
<td>CNL</td>
<td>Cross-nested logit</td>
</tr>
<tr>
<td>PCL</td>
<td>Paired combinatorial logit</td>
</tr>
<tr>
<td>KKT</td>
<td>Karush–Kuhn–Tucker</td>
</tr>
<tr>
<td>MAE</td>
<td>Mean absolute error</td>
</tr>
<tr>
<td>RMSE</td>
<td>Root-mean-square error</td>
</tr>
<tr>
<td>ATT</td>
<td>Average travel time</td>
</tr>
<tr>
<td>IGP</td>
<td>Improved gradient projection</td>
</tr>
</tbody>
</table>
The objective function consists of three terms. $Z_1$ is the well-known Beckmann’s [53] transformation. $Z_2−Z_5$ correspond to the NL model. Equation (5) shows the flow conservation constraints. Note that $g_{rs}^c$ consists of travel time, waiting time, and other personal utilities (e.g., fare, transfer time, and inconvenience).

To set up the generalized travel costs (i.e., the lower nest modes), we adopt a flow-dependent travel time for the auto mode and bus mode and a flow-independent travel time for the metro mode as follows:

\[
\begin{align*}
\text{auto: } c_k^rs &= \sum_{a \in A} t_a(x_a) \delta_k^rs, \\
\text{bus (B): } g_{B}^rs &= \sum_{a \in A} \delta_{Ba}^rs \left( t_a(x_a) + d_a^B \eta_a^B + \frac{(1 - \eta_a^B)I_a}{ws + Wa} \right) + VOT \cdot F^rs, \\
\text{metro (M): } g_{M}^rs &= \sum_{a \in A} \delta_{Ma}^rs \left( t_a(x_a) + d_a^M \eta_a^M + \frac{(1 - \eta_a^M)I_a}{ws + Wa} \right) + VOT \cdot F^rs, \\
\text{B − M: } g_{BM}^rs &= \sum_{a \in A} \delta_{BMA}^rs \left( \min \left( t_a(x_a) + d_a^B, \left( t_a^M + d_a^M \right) \eta_a^T + \frac{(1 - \eta_a^T)I_a}{ws + Wa} \right) \right) + VOT \cdot F^rs.
\end{align*}
\]

Travel times on each link are assumed to follow the Bureau of Public Roads (BPR) function.

\[
t_a(x_a) = t_a^0 \left( 1 + \eta \left( \frac{x_a}{C_a} \right)^\beta \right), \quad \forall a \in A,
\]

where $t_a^0$ is free-flow travel time of link $a$ and $\eta$ and $\beta$ are given parameters.

**Proposition 1.** The MP formulation in equation (4) through equation (5) gives the mode choice solution of the NL model.
the auto mode as follows:

\[
\frac{\partial L}{\partial f^k_{rs}} = c^r_k - \phi^r_k = 0, \quad \forall \, rs \in RS, \quad k \in K_{rs},
\]

\[
\frac{\partial L}{\partial q_{im}} = \frac{1}{\theta} \ln \left( \frac{\sum_{m \in M^r_u} q^r_{im}}{q^r_i - \sum_{m \in M^r_u} q^r_{im}} \right) + \frac{T^r_i}{\theta} \ln \left( \frac{q^r_{im}}{\sum_{m \in M^r_u} q^r_{im}} \right) + g_c^r_m - \phi^r_m = 0, \quad \forall \, rs \in RS, \quad m \in M^r_u.
\]

Note that equation (20) presents the UE condition. From equation (9), we have

\[
\frac{1}{\theta} \ln q^r_i + \frac{T^r_i}{\theta} \ln q^r_{im} - \frac{T^r_i}{\theta} \ln q^r_i + g^r_m = \frac{1}{\theta} \ln q^r_A + c^r_i,
\]

where \( q^r_A \) is auto mode demand \( q^r_A = (q^r - \sum_{m \in M^r_u} q^r_{im}) \).

From equation (12), we have the marginal probability for the auto mode as follows:

\[
P^r_A = \frac{q^r_A}{q^r_i + \sum_{m \in M^r_u} q^r_{im}} = \frac{\exp \left( \frac{\theta^r_i}{\tau^r_i} \left( \lambda - g^r_m \right) \right) - \exp \left( -\theta^r_i c^r_i \right)}{\exp \left( \frac{\theta^r_i}{\tau^r_i} \left( \lambda - g^r_m \right) \right) + \left( \sum_{m \in M^r_u} \exp \left( \frac{\theta^r_i}{\tau^r_i} \left( \lambda - g^r_m \right) \right) \right)^{r^r_i}}.
\]

For the proof of Proposition 2, we need to prove that objective function in equation (5) is strictly convex with respect to mode-route flow variables and the convexity of the feasible region. For the objective function, taking the Hessian matrix in equation (4) with respect to the mode-route flow variables gives

\[
\frac{\partial^2 Z}{\partial f^r_k \partial f^r_j} = \left\{ \begin{array}{ll}
\frac{d_t(x_a) \delta^r_k}{dx_a} & j = k \\
0 & \text{otherwise}
\end{array} \right.,
\]

which implies the positive-definite matrix when the link travel time \((t_a(x_a))\) is a strictly monotonic increasing function.

\[\text{Proposition 2. The MP formulation in equation (4) through equation (5) yields a unique solution.}\]
which also implies the positive-definite matrix. Therefore, the MP formulation model has a unique solution. This completes the proof.

From equation (9), the excess demand function of the lower nest modes in the NL modal split function can be defined as follows (see Appendix for the detailed derivation):

\[
\begin{align*}
\frac{\partial^2 L}{\partial q_m q_n} & = \begin{cases} 
1 \frac{1}{\theta^2} \left( \frac{1}{\sum_{m \in M^c} q_m} + \frac{1}{\theta^2 - \sum_{m \in M^c} q_m} \right) - \frac{\tau^2}{\theta^2} \left( \frac{1}{q_m} - \frac{1}{\sum_{m \in M^c} q_m} \right) \geq 0; & \text{if } m = n, \\
0; & \text{otherwise}
\end{cases}
\end{align*}
\]

(16)

4. Path-Based Gradient Projection Method

The gradient projection (GP) algorithm is a well-known path-based algorithm for solving various traffic assignment problems. Before considering the gradient projection (GP) algorithm for solving the CMSTA model (or NL-UE model), we provide a brief review on the ordinary GP algorithm for solving the UE traffic assignment problem with a fixed demand. Equation (18) shows the flow update equations [35] for solving a fixed demand traffic assignment problem. In each iteration, the algorithm finds a descent search direction by solving a shortest path problem and updates the solution by scaling it with the second derivative information.

\[
\begin{align*}
& f^r_k (n + 1) = \max \left\{ \left[ f^r_k (n) - \frac{\alpha}{\delta^2_k} (c^r_k (n) - c^r_{K^r} (n)) \right], 0 \right\}, \quad \forall rs \in RS, \quad k \in K^r, \quad k \neq K^r, \\
& f^r_{K^r} (n + 1) = q_{rs} - \sum_{k \in K^r} f^r_k (n + 1), \quad \forall rs \in RS,
\end{align*}
\]

(18)

where \( n \) is the iteration number; \( \alpha \) is the step size; \( \max \{ f, 0 \} \) denotes the projection of the argument onto the non-negative orthant of the decision variables (i.e., nonshortest route flows); and \( f^r_k (n + 1) \) and \( f^r_{K^r} (n + 1) \) are the flows on route \( k \) and shortest route \( K^r \) O-D pair \( rs \) at iteration \( n + 1 \).

Equation (19) shows the route travel time difference \( (c^r_k (n) - c^r_{K^r} (n)) \) and the scaling factor \( (s^r_k (n)) \) to find the search direction.

\[
\begin{align*}
& c_k^r - c_{K^r}^r = \sum_{a \in A} t_a (x_a) \delta^r_{ka} - \sum_{a \in A} t_a (x_a) \delta^r_{K^r a} , \\
& s_k^r = \sum_{a \in A} t_a (x_a) \left( \delta^r_{ka} - \delta^r_{K^r a} \right)^2 ,
\end{align*}
\]

(19)

where \( t_a (x_a) \) and \( t_a (x_a) \) are the travel time and its derivative of the auto mode on link \( a \).

For a detailed description of the GP algorithm, readers should refer to [26,35]. The modifications for solving the
CMSTA model with NL modal split function and the improvement strategies for solving large-scale multimodal transportation networks are described as follows.

4.1. Two-Phase Equilibration. Unlike the UE traffic assignment problem, the CMSTA problem considered in this paper (i.e., UE for route choice and NL for mode choice) requires not only equilibrating route flows of the auto mode for route choice but also equilibrating modal splits among multiple available modes (i.e., both private and public transport modes) following the NL model for mode choice. Based on the hierarchical choice structure given in Figure 1, a two-phase equilibration method is developed to perform the equilibration for the two travel choices. In the first phase, GP equilibrates between the private auto mode and public transit mode using the minimum route travel time of the auto mode and the excess demand cost of the transit mode. In the second phase, the lower nested probability function is applied to determine the mode demands among the transit modes. From the lower nested probability function, the mode demands among the transit modes are derived as follows:

Figure 2: Illustration of the mode choice equilibration under the nested logit modal split function; (a) network, demand, and cost functions; (b) upper level equilibrium (marginal prob.); (c) lower level equilibrium (conditional prob.); (d) travel time change by flows; (e) equilibrium solution; (f) probabilities verification; (g) cost verification.
on the other hand, if \( w(q^r_m) > c_{rs} \) (see Figure 3(b)), route flows from the auto mode should be decreased accordingly to balance the two auto and public modal costs. On the other hand, if \( w(q^r_m) < c_{rs} \) (see Figure 3(a)), route flows from the auto mode should be increased accordingly to balance the two auto and public modal costs.

Once the upper nest is equilibrated in the first phase (i.e., the auto demand \( (\sum_{k \in K_r} f^r_k) \) and the public transit demand \( (q^r_m) \) are determined), the second phase is performed to split the public transit demand \( (q^r_m) \) into multiple transit modes \( (q^m_n) \) in the lower nest to achieve equilibration in the mode choice. For the lower nest, the modal split probability function can be directly calculated as follows:

\[
q^m_n(n+1) = q^m_n(n+1) \frac{\exp((\theta^r/\tau^r) - g^m_n(n+1))}{\sum_{m \in M} \exp((\theta^r/\tau^r) - g^m_n(n+1))},
\]

with the minimum route travel time of the auto mode \( (c_{rs}^{rs}) \) and the excess demand cost for transit mode in equation (20).

Figure 3 provides a graphical illustration of the first phase equilibration in the two-phase GP algorithm, which involves the modal splits of the upper nest (i.e., private car and public transit). If \( w(q^r_m) < c_{rs}^{rs} \) (see Figure 3(a)), route flows from the auto mode should be decreased accordingly to equilibrate the modal costs of both auto and public modes. On the other hand, if \( w(q^r_m) > c_{rs}^{rs} \) (see Figure 3(b)), route flows from the auto mode should be increased accordingly to balance the two auto and public modal costs.

4.2. Column Generation Scheme. The disaggregate simplicial decomposition (DSD) algorithm proposed by Larsson and Patriksson [41] is one of the earlier path-based traffic assignment algorithms. The algorithm consists of a linearized subproblem and a master problem to determine the next solution point. From the subproblem, extreme points (i.e., route generation) are resulted and then the master problem is solved with extreme points generated in the subproblem, which is equilibration. In the algorithm, the master problem is iterated until the route flows stabilized before generating new paths in the column generation procedure. Korpelevich [54] introduced a double projection method to relax the strongly monotone condition in the mapping in the variational inequality problem. From the predictor step, the strongly monotone condition is relaxed and the solution is adjusted to obtain a more accurate solution in the corrector step. Galligari and Scandarone [55] recently proposed a fast path equilibration algorithm with an adaptive column generation scheme for each O-D pair. Lu et al. [56] used column generation scheme under the dynamic user equilibrium traffic assignment problem. Wang et al. [57] introduced the tolerance-based strategies for the column generation scheme in the dynamic user equilibrium traffic assignment problem. Ameli et al. [22, 58] recently proposed trip-based dynamic traffic assignment and simulation-based dynamic traffic assignment. These two algorithms adopt the column generation scheme for path generation. In addition, they assessed the computational performance of the inner loop methods in the first paper and proposed algorithm for solving large-scale dynamic traffic assignment in the second paper.

In this study, we apply a similar adaptive column generation scheme in the gradient projection algorithm. To improve the computational efficacy in the few iterations, we adopt the following condition in the inner loop (master problem) as the termination step:

\[
e_{inner} < e_{outer} \cdot \gamma,
\]

where \( 0 < \gamma \leq 1.0 \) is the given parameter.

First, the convergence error in the outer loop is computed after the column generation step. Then, the inner loop (equilibration) step is performed without the column generation step until it satisfies the termination condition in equation (22). The overall flowchart is shown in Section 5.

Another implementation improvement is the column dropping scheme. If the path set \( K_r(n) \) for each O-D pair \( rs \) contains relatively few paths, the path-based algorithms will perform satisfactorily without column dropping. However, the number of routes in \( K_r(n) \) can be very large for some situations (e.g., problems with mostly simple paths connecting each O-D pair, heavily congested networks, or multimodal transportation networks). Under these situations, it may be necessary to remove the inactive (zero flow or near-zero flow) paths from the route set to reduce the amount of calculation and bookkeeping needed in each iteration. Chen and Jayakrishnan [59] have shown that GP, which automatically achieves column dropping during the projection step, reduces computational time if the number of active (positive flow) paths is considerably less than the number of generated paths. In this study, we adopt the column dropping step suggested by Chen and Jayakrishnan [59].

4.3. Solution Procedure. Figure 4 shows the overall flowchart of the solution procedure. After the initialization step, column (or route) generation is performed. With a new path set, flow updated is active, and then restrict equilibration is performed without new path generation. If convergence error computed in restrict equilibration satisfies the inner loop convergence criteria, they go to column dropping step. The process is iterated until convergence error satisfies the outer loop convergence criteria.

Figure 5 shows the detailed flow update procedure following the two-phase gradient projection algorithm.

As described above, modified GP extends the equilibration of private vehicles (i.e., traffic assignment) to also consider equilibration of multiple transport modes (i.e., modal splits) by using a two-phase procedure which is the equilibration between public and private modes as the first phase and equilibration of modal splits among the public transport modes as the second phase.

In the column generation step, we adopt origin-based column generation to improve computation efficiency. See [60] for more detailed results. Based on the literatures (see, e.g., [26, 35, 61, 62]), a unit stepsize (i.e., \( \alpha = 1.0 \)) is adopted for all iteration \( n \) since the second derivative information for an automatic scaling and the one O-D at-a-time flow update
strategy is used in the equilibration procedure. This scheme has been found to be helpful in reducing computational efforts. In this paper, we also used unit step size (i.e., $\alpha = 1.0$) with “one-at-a-time” flow update strategy. However, a line search step (e.g., self-adaptive strategies) can be used to determine a suitable stepsize to help better convergence, especially when highly accurate solutions are sought.

5. Numerical Experiments

The transportation network in the city of Seoul, Korea, is used to investigate the performance of the proposed two-stage GP algorithm in solving the NL-UE model. The convergence characteristics are tested, and then parameter sensitivity analysis is examined.

5.1. Description of Seoul Network and Experimental Set Up.

The Seoul transportation network shown in Figure 6 consists of 425 zones, 7,512 nodes, 11,154 links, and 107,434 O-D pairs with a total of 2,874,441 trips. Table 2 summarizes some details of the multimodal transportation network in Seoul, Korea. The nested logit structure and modal split functions for modeling mode choice are shown in Figure 7.

The occupancy of auto mode and bus mode is set 1.3 per passenger car and 19.27 per bus, respectively (MTA, 2015). The transit has the distance-based fare structure.
Figure 5: Two-phase flow update strategy.

Figure 6: Continued.
loss of generality, the metro mode is assumed to provide enough operation capacity such that there is no congestion in this mode. One hundred maximum number of inner loop is set, and $\gamma$ value is set 0.1. The tolerance error of the relative gap (RG) shown in equation (23) is set $1E-8$:

$$
\text{RG} = \sum_{k \in K_{rs}} \frac{f_k^r(n) \cdot \left( c_k^r(n) - \min \{ c_k^{rs}(n+1), w_T^r(n+1) \} \right)}{f_k^r(n) \cdot c_k^r(n) + w_T^r(n) \cdot q_T^r(n)}
$$

(23)

The improved GP algorithm is coded in Intel Visual FORTRAN XE and runs on a 3.60 GHz processor and 16.00 GB of RAM.

5.2. Convergence Characteristics. Figure 8 shows the convergence characteristics using the method of successive averages (MSA), Evan’s algorithm, ordinary GP algorithm, and improved GP (IGP) algorithm. The method of successive averages (MSA) is the most widely used method in the predetermined line search scheme [52]. It predetermines a diminishing step size sequence such as $\{1, 1/2, 1/3, \ldots, 1/n\}$. On the other hand, Evan’s algorithm [2] is a classical algorithm for solving the combined distribution and assignment (CDA) problem as well as the combined modal split and traffic

![Figure 6: Multimodal transportation network in Seoul, Korea: (a) auto network; (b) bus network, and (c) metro network.](image_url)

![Figure 7: Nested mode choice structure and its modal splits.](image_url)

<table>
<thead>
<tr>
<th>Total length (km)</th>
<th>Auto</th>
<th>Bus</th>
<th>Metro</th>
<th>Bus-metro</th>
</tr>
</thead>
<tbody>
<tr>
<td>2,790</td>
<td>1,202</td>
<td>467</td>
<td>1,669</td>
<td></td>
</tr>
<tr>
<td>Number of stations</td>
<td>—</td>
<td>3,635</td>
<td>354</td>
<td>3,989</td>
</tr>
<tr>
<td>Number of lines</td>
<td>—</td>
<td>1,765</td>
<td>14</td>
<td>1,779</td>
</tr>
</tbody>
</table>

Table 2: Summary of the multimodal transportation network in Seoul, Korea.
assignment problem (see, e.g., [20]). The results reveal the slow convergence of the MSA and Evan’s algorithm. The GP algorithm presents a slower convergence than the IGP algorithm. With the GP algorithm, the solution cannot reach the highly accurate level of a RG of 1E-8, while the solution with IGP is reached to RG of 1E-8 in 3,100 seconds. Note that the tolerance error of 1E-8 is much stricter than the typical one (i.e., 1E-4) required in practice [64].

In the following analyses, we conduct the sensitivity analysis with respect to the inner loop parameters $\gamma$. The parameter adjusts the stability of path flows before generating new paths in the column generation step. Figure 9(a) shows the convergence with different values (0.1, 0.5, 0.8), and Figure 9(b) provides active number of inner iterations. Note that the maximum number of inner iteration is 100. As $\gamma$ value increases, the computational effort is significantly increased. When a higher $\gamma$ value is used, the required number of inner iteration is relatively smaller, but it requires more outer iterations. On the other hand, using a small $\gamma$ value requires more inner iterations with the stabilized route flows, but it requires less the number of outer iteration. In Figure 9(b), we can observe that a smaller $\gamma$ requires more inner iterations. The algorithm is converged in the smaller outer iteration. After a few outer iterations, the inner iteration reaches the maximum inner loop. On the other hand, using higher $\gamma$ value is converged well in a few inner iterations, but it requires more outer iteration.
Figure 10: Effect of logit parameters on modal splits.

Figure 11: Effect of logit parameter on flow allocation: (a) auto flow difference; (b) bus flow difference; (c) metro flow difference; (d) bus-metro flow difference.
5.3. Application of the CMSTA Problem. This section examines the sensitivity of the IGP algorithm with respect to the nested logit parameters (i.e., $\theta_s$ and $\tau_T^{rs}$). Figures 10 and 11 present the effects of these two parameters on modal splits and average modal travel costs, respectively. Figure 10 shows that $\theta_s$ is sensitive to the modal splits between private and public modes (i.e., a 3% increase in the private car mode when $\theta_s$ increases from 0.1 to 0.133), while $\tau_T^{rs}$ is fairly insensitive to the modal splits. When the transit demand transfers to auto mode, the bus mode is affected more than metro and bus-metro mode. This is because the travel time of the bus and auto is dependent. The result also shows that the modal splits are different albeit the ratio ($\theta_s/\tau_T^{rs}=2.0$) is the same.

Figure 11 shows the link flow differences between case 1 (i.e., $\theta_s=0.1$ and $\tau_T^{rs}=0.05$) and case 2 (i.e., $\theta_s=0.133$ and $\tau_T^{rs}=0.067$). The green color shows link flows using case 1 are larger than those of case 2, while the red color shows the reverse.

Specifically, the figure shows the mean absolute error (MAE) and the root-mean-square error (RMSE) for assessing the link flow difference and average travel time (ATT) (i.e., total travel time/mode demand). Recall that the modal splits for case 1 are 72.19% and 27.81% for auto and transit mode, respectively. Figure 10 shows the effects of these two parameters on modal splits. When the transit demand transfers to auto mode, the bus mode is affected more than metro and bus-metro mode. This result also shows that the modal splits are different even though the ratio ($\theta_s/\tau_T^{rs}=2.0$) is the same.

6. Concluding Remarks

In this paper, we presented (1) the CMSTA model with the nested logit (NL) model for the mode choice and the user equilibrium (UE) model for the route choice and (2) a modified-and-improved gradient projection (IGP) algorithm to solve the proposed NL-UE model. The proposed NL-UE model considered the mode similarity through the NL model under the congested network obtained by the UE condition. Specifically, an equivalent MP formulation for the NL-UE model was provided using a modified access demand to incorporate the two-level nested tree structure. On the other hand, the IGP algorithm was developed in such a way to perform a two-level equilibration. In the first phase, the upper nest of the NL model is solved. Then, the lower nest of the NL model is applied with updated flows from the upper nest in the second phase. For the solution algorithm, the improved GP algorithm with inner loop equilibration procedure was introduced and shown the effectiveness in solving the combined NL modal split and UE traffic assignment (NL-UE) problem. This is because the excess demand cost function consists of two logarithm terms. These terms are sensitive to a small traffic flow change. A more stabilized path flow is required before generating new paths.

A real-size transportation network in Seoul, Korea, was used to demonstrate the applicability of the IGP algorithm for solving the proposed NL-UE model. The algorithm can conduct 1E-8 convergence criteria. The convergence time of the IGP algorithm seemed to be better than that of the ordinary GP algorithm. Further, the NL parameters have a significant impact on the convergence time and the mode share results.

For future research, the improved GP algorithm should be extended to consider mode (or vehicle) interactions with asymmetric cost functions. In addition, the route choice model for transit modes should be considered with transit information systems (e.g., arrival information). It would be interesting to see how the improved GP algorithm performs when more realistic features are incorporated into the CMSTA problem. In addition, a time-dependent model can be considered with travel activity patterns.

Appendix

Equation (A.1) is arranged as follows:

\[
\exp \left( \frac{\theta_s}{\tau_T^{rs}} \left(-g^{rs}_{m} \right) \right) = \exp \left( \ln \left( \frac{q_{m}^{rs}}{\sum_{m \in M_T^{rs}} q_{m}^{rs}} \right) \right) = \frac{q_{m}^{rs}}{\sum_{m \in M_T^{rs}} q_{m}^{rs}}
\]

Equation (A.1) is expressed with equation (A.2) as follows:

\[
\frac{1}{\sum_{m \in M_T^{rs}} \exp \left( \frac{\theta_s}{\tau_T^{rs}} \left(-g^{rs}_{m} \right) \right)} = \exp \left( \ln \left( \frac{q_{m}^{rs}}{\sum_{m \in M_T^{rs}} q_{m}^{rs}} \right) \right) \cdot \exp \left( \frac{\theta_s}{\tau_T^{rs}} g^{rs}_{m} \right) - \ln \left( \sum_{m \in M_T^{rs}} \exp \left( \frac{\theta_s}{\tau_T^{rs}} \left(-g^{rs}_{m} \right) \right) \right)
\]

\[
= \ln \left( \frac{q_{m}^{rs}}{\sum_{m \in M_T^{rs}} q_{m}^{rs}} \right) \cdot \frac{\theta_s}{\tau_T^{rs}} \left(-g^{rs}_{m} \right)
\]

Hence, equation (A.3) is expressed with equation (A.2) as follows:
Data Availability

The basic data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors’ Contributions

Seungkyu Ryu collected and analyzed the data, interpreted the results, and prepared the manuscript draft. All authors contributed to study conception and design, reviewed the results, and approved the submitted version of the manuscript.

Acknowledgments

This research was supported by the Basic Science Research Program through the National Research Foundation (NRF) of Korea by the Ministry of Science (grant no. NRF-2016R1C1B2016254), the Ministry of Science, ICT, Republic of Korea (project no. K-21-L01-C05-S01), the Research Grants Council of the Hong Kong Special Administrative Region (project no. 15212217), the Research Committee of the Hong Kong Polytechnic University (project no. 1-ZVJV), and the Chiang Mai University.

References

[26] G. Wang, A. Chen, S. Kittamkesorn et al., “A multi-modal network equilibrium model with captive mode choice and...


