

An Origin-Based Model for Unique Shortest Path Routing¹

Changyong Zhang^a

^a*Department of Finance and Banking, Faculty of Business, Curtin University Sarawak, Miri, Malaysia
E-mail: changyong.zhang@curtin.edu.my*

Abstract

Link weights are the main parameters of shortest path routing protocols, the most commonly used protocols for IP networks. The problem of optimally setting link weights for unique shortest path routing is addressed. Due to the complexity of the constraints involved, there exist challenges to formulate the problem in such a way based on which a more efficient solution algorithm than the existing ones may be developed. In this paper, an exact formulation is first introduced and then mathematically proved correct. It is further illustrated that the formulation has advantages over a prior one in terms of both constraint structure and model size for a proposed decomposition method to solve the problem.

Keywords: Mathematical modeling, model validation, constraint structure, decomposition, shortest path routing, link weights

1. Introduction

Shortest path routing protocols such as OSPF are the most widely deployed and commonly used protocols for IP networks [Black, 2000, Moy, 1998, Tanenbaum and Wetherall, 2011]. They also find applications in, for example, road networks [Abraham et al., 2010, Zhan and Noon, 1998]. In shortest path routing, each link is assigned a weight and traffic demands are routed through the shortest paths with respect to link weights [Bertsekas and Gallager, 1992], given by a shortest path first algorithm [Bellman, 1958, Dijkstra, 1959, Ford and Fulkerson, 2010]. Link weights are hence the key parameters and an essential problem is then to find an appropriate weight set for shortest path routing.

A simple way to set link weights is the hop-count method, assigning the weight of each link to one. The length of a path is thus the number of hops. Another default approach recommended by Cisco is the inv-cap method [Cisco Systems Inc., 2000, Thomas, 2003], setting the weight of a link inversely proportional to its capacity, without taking traffic conditions into consideration. More generally, the weight of a link may depend on and be related to its transmission capacity and traffic load. Accordingly, a problem of interest is to find an optimal weight set for shortest path routing [Burton and Toint, 1992], given a network topology, a projected traffic matrix [Altm et al., 2010, Applegate and Cohen, 2006, Feldmann et al., 2001, Wang et al., 2006], and an objective function [Balon et al., 2006, Pióro et al., 2002, Pióro and Medhi, 2004, Zhang, 2006].

The problem has two instances, depending on whether multiple shortest paths or only a unique routing path from an origin node to a destination node is allowed [Altm et al., 2013, Bley et al., 2010, Giroire et al., 2015]. For the first instance, a number of heuristic methods have been introduced, each based on, for example, a local search method mostly using an increasing piecewise linear convex cost function or a heap-reduction technique [Buriol et al., 2008, Fortz and Thorup, 2000, 2004, Fortz and Ümit, 2011, Ramalingam and Reys, 1996], a genetic algorithm [Buriol et al., 2005, Ericsson et al., 2002, Mulyana and Killat, 2002], simulated

¹This paper is based on a research project carried out at Imperial College London. Some related results have been presented on the 4th International Conference on Networking, Reunion Island, April 2005, the 8th INFORMS Telecommunications Conference, Dallas, USA, March 2006, INFORMS Annual Meeting 2006, Pittsburgh, PA, USA, November 2006, and the 6th International Congress on Industrial and Applied Mathematics, Zurich, Switzerland, July, 2007, respectively. The author would like to thank the conference attendees of all the corresponding sessions. An earlier version of this paper can be found at <http://arxiv.org/abs/0807.0038>.

annealing [Pióro et al., 2002, Pióro and Medhi, 2004], Lagrangian relaxation [Holmberg and Yuan, 2000, Srivastava et al., 2005], an integrated approach [Wang et al., 2001], or a MILP-based algorithm [Amaldi et al., 2013, Cianfrani et al., 2012]. For the second instance, the Lagrangian relaxation method and local search method have been proposed [Lin and Wang, 1993, Ramakrishnan and Rodrigues, 2001]. These methods have been tested using given data sets and have been verified to result in acceptable routing performance. Meanwhile, with these heuristic methods, the problem is not formulated exactly and is in general not solved optimally. In particular, the resulting performance is not consistently close to the optimal general routing [Bley et al., 2010, Fortz and Thorup, 2000]. It is hence worth looking into the possibility of formulating the problem explicitly, from which optimal solutions may be obtained for data instances with reasonable sizes arising from real-world applications.

From a management perspective, unique-path routing uses simpler routing mechanisms and allows for easier monitoring of traffic flows [Ben-Ameur and Gourdin, 2003, Hock et al., 2010]. Hence, this paper considers the *unique shortest path routing problem*, as specified in Section 2.1. It is a reduction in the integer multi-commodity flow problem [Ahuja et al., 1993], which has been well addressed [Barnhart et al., 2000, Dinitz et al., 1999, Park et al., 1996].

Partially due to the challenges involved in modeling the problem appropriately, most existing solution algorithms are heuristic [Bley, 2009, Kolliopoulos and Stein, 2001, Skutella, 2002]. Efforts have been made to formulate the problem mathematically. For example, a two-phase heuristic has been proposed, to allocate a unique shortest path for each pair of nodes and to compute link weights compatible with the set of routing paths [Ben-Ameur and Gourdin, 2003]. To guarantee the existence of a compatible set of weights in the second subproblem, necessary conditions are provided and discussed in detail. The second problem is also referred to as the inverse shortest paths problem, variants of which have been extensively studied [Ahuja and Orlin, 2001, 2002, Bley, 2007, Burton and Toint, 1992, 1994, Xu and Zhang, 1995, Zhang and Liu, 1996], with or without the uniqueness of the perceived optimal solution and the integrality of the perturbed cost vector being taken into consideration.

Models without the necessity of resorting to the two-phase heuristic have also been proposed [Zhang and Rodošek, 2005a,b]. This avoids considering the compatibility between the two subproblems. In the meantime, as a critical step when a model is introduced, the correctness of the models still remains to be verified rigorously. Mathematical models have also been developed for related problems [Bley and Koch, 2008, Faragó et al., 2003, Holmberg and Yuan, 2004], whereas they are mostly path-based and potentially result in an exponential number of constraints. This leaves space for further exploring the structure properties of the problems, which may provide more flexibility to derive alternative solution methods.

This paper focuses on mathematical modeling of the problem, which may potentially yield a new exact solution approach for real-world applications with average data sizes. In particular, the correctness of the models is mathematically proved rigorously. In Section 2, the problem is specified and two different exact formulations are introduced. The second one is then mathematically proved correct in Section 3. Differences between the two formulations in both constraint structure and model size are discussed in Section 4, followed by the conclusion in Section 5.

The ideas behind the two formulations may be adopted to model related problems in network routing and other fields. It is also hoped that the steps of model formulation, model validation, and model comparison may provide a reference procedure for mathematical modeling.

2. Model Formulation

2.1. Problem Specification

The unique shortest path routing problem is defined as follows. Given

- a network topology, which is a directed graph structure $\mathcal{G} = (\mathcal{N}, \mathcal{L})$, where

- \mathcal{N} is a finite set of nodes, each of which represents a router; and
- \mathcal{L} is a set of directed links, each of which corresponds to a transmission link; (For each $(i, j) \in \mathcal{L}$, i is the starting node, j is the ending node, and $c_{ij} \geq 0$ is the link capacity.)
- a traffic matrix, which is a set of demands \mathcal{D} ; (It is assumed that there is at most one demand between each origin–destination pair. For each demand $k \in \mathcal{D}$, $s_k \in \mathcal{N}$ is the origin node, $t_k \in \mathcal{N}$ is the destination node, and $d_k > 0$ is the required bandwidth. Accordingly, \mathcal{S} is the set of all origin nodes, \mathcal{T}_s is the set of all destination nodes of demands originating from node $s \in \mathcal{S}$, and \mathcal{D}_s is the set of all demands originating from node $s \in \mathcal{S}$.)
- lower and upper bounds of link weights, which are positive real numbers w_{\min} and w_{\max} , respectively; and
- an objective function, e.g., to maximize the sum of the residual capacities,

find an optimal weight set $w_{ij}, (i, j) \in \mathcal{L}$, subject to

- flow conservation constraints: For each demand, at each node, the sum of all incoming flows (including the demand bandwidth at the origin node) is equal to the sum of all outgoing flows (including the demand bandwidth at the destination node);
- link capacity constraints: For each link, the load of traffic flows traversing the link does not exceed the capacity of that link;
- path uniqueness constraints: Each demand has a unique routing path; and
- path length constraints: For each demand, the length of each path assigned to route the demand is strictly less than that of any other possible and unassigned path to route the demand.

By the above definition, the routing path of a demand is the shortest one among all possible paths. For each link, the routing path of a demand either traverses the link or not. The path length and path uniqueness constraints require that the length of the unique shortest path to route a demand is less than that of any other possible path from the origin to the destination.

As shown in Figure 1, concerning the constraints, there are three scenarios to be considered regarding the relationship between the lengths of shortest paths and link weights.

- If the routing path of demand k traverses link (i, j) , the length of the shortest path from node s_k to j is that from s_k to i plus the weight of link (i, j) ;
- If the routing path of demand k does not traverse link (i, j) but transits node j , the length of the shortest path from node s_k to j is strictly less than the sum of that from s_k to i and the weight of link (i, j) ; (otherwise, there would be at least two shortest paths to route demand k .)
- If the routing path of demand k neither traverses link (i, j) nor transits node j , the length of the shortest path from node s_k to j is less than or equal to the sum of that from s_k to i and the weight of link (i, j) .

With the problem being specified, below it is mathematically formulated from two different perspectives, based on the study of the problem properties. For comparison, a demand-based model is first introduced, followed by the origin-based counterpart.

2.2. A Demand-Based Model

Based on the observation on the relationship between the length of a shortest path and the weights of links that it traverses, the problem can be mathematically formulated as a *demand-based* model (DBM) as follows, by defining one routing decision variable for each link–demand pair [Zhang and Rodošek, 2005a].

- *Routing decision variables:*

$$x_{ij}^k \in \{0, 1\}, \forall k \in \mathcal{D}, \forall (i, j) \in \mathcal{L} \quad (1)$$

is equal to 1 if and only if the routing path of demand k traverses link (i, j) . The number of this set of variables is $|\mathcal{D}||\mathcal{L}|$.

- *Link weight variables:*

$$w_{ij} \in [w_{\min}, w_{\max}], \forall (i, j) \in \mathcal{L} \quad (2)$$

denotes the routing cost of link (i, j) . The number of this set of variables is $|\mathcal{L}|$.

- *Path length variables:*

$$l_i^s \in [0, +\infty), \forall s \in \mathcal{S}, \forall i \in \mathcal{N} \quad (3)$$

represents the length of the shortest path from origin node s to node i . Apparently, $l_{t_k}^{s_k}$ is the length of the shortest path to route demand $k \in \mathcal{D}$ and $l_s^s = 0, \forall s \in \mathcal{S}$. The number of this set of variables is $|\mathcal{S}||\mathcal{N}|$.

- *Flow conservation constraints:*

$$\sum_{h:(h,i) \in \mathcal{L}} x_{hi}^k - \sum_{j:(i,j) \in \mathcal{L}} x_{ij}^k = \begin{cases} -1, & \text{if } i = s_k \\ 1, & \text{if } i = t_k \\ 0, & \text{otherwise} \end{cases}, \forall k \in \mathcal{D}, \forall i \in \mathcal{N}. \quad (4)$$

The number of this set of constraints is $|\mathcal{D}||\mathcal{N}|$.

- *Link capacity constraints:*

$$\sum_{k \in \mathcal{D}} d_k x_{ij}^k \leq c_{ij}, \forall (i, j) \in \mathcal{L}. \quad (5)$$

The number of this set of constraints is $|\mathcal{L}|$.

- *Path uniqueness constraints:* Under the combined restriction of the flow conservation and path length constraints, the constraints are satisfied automatically.

- *Path length constraints:*

$$\left. \begin{aligned} x_{ij}^k = 0 \wedge \sum_{h:(h,j) \in \mathcal{L}} x_{hj}^k = 0 &\Rightarrow l_j^{s_k} \leq l_i^{s_k} + w_{ij} \\ x_{ij}^k = 0 \wedge \sum_{h:(h,j) \in \mathcal{L}} x_{hj}^k = 1 &\Rightarrow l_j^{s_k} < l_i^{s_k} + w_{ij} \\ x_{ij}^k = 1 &\Rightarrow l_j^{s_k} = l_i^{s_k} + w_{ij} \end{aligned} \right\}, \forall k \in \mathcal{D}, \forall (i, j) \in \mathcal{L}, \quad (6)$$

which are in logic form, as interpreted in Section 2.1 and illustrated in Figure 1 with thick lines being routing paths. They can be linearized by introducing two constants ε and M with $0 < \varepsilon \ll M$. The new constraints are

$$\left. \begin{aligned} l_j^{s_k} &\leq l_i^{s_k} + w_{ij} - \varepsilon \left(\sum_{h:(h,j) \in \mathcal{L}} x_{hj}^k - x_{ij}^k \right) \\ l_j^{s_k} &\geq l_i^{s_k} + w_{ij} - M(1 - x_{ij}^k) \end{aligned} \right\}, \forall k \in \mathcal{D}, \forall (i, j) \in \mathcal{L}. \quad (7)$$

The number of this set of constraints is $2|\mathcal{D}||\mathcal{L}|$. By enumerating all possible values of the routing decision variables $x_{ij}^k, k \in \mathcal{D}, (i, j) \in \mathcal{L}$, it can be verified that the linearized constraints are identical to the original ones. For the efficiency in solving the problem, it is worth looking into how to choose the values of ε and M accordingly with respect to the size of the network $\mathcal{G} = (\mathcal{N}, \mathcal{L})$ as well as the values of w_{\min} and w_{\max} .

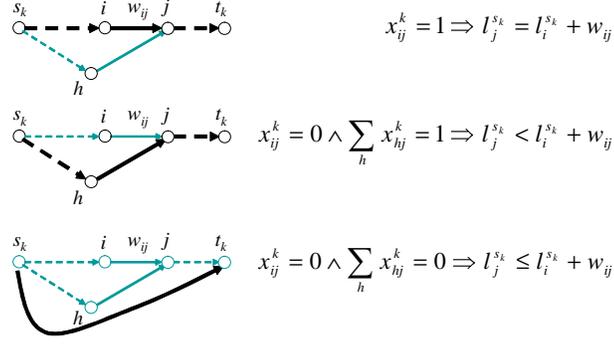


Figure 1: Illustration of the path length constraints

- *Objective function:*

$$\max \sum_{(i,j) \in \mathcal{L}} \left(c_{ij} - \sum_{k \in \mathcal{D}} d_k x_{ij}^k \right),$$

which is equivalent to

$$\min \sum_{(i,j) \in \mathcal{L}} \sum_{k \in \mathcal{D}} d_k x_{ij}^k. \quad (8)$$

As a result, the complete model is

$$\begin{aligned} \text{DBM :} \quad & \text{Optimize} \quad (8) \\ & \text{Subject to} \quad (4), (5), (7), (1), (2), (3) \end{aligned} \quad (9)$$

A necessary condition of the unique shortest path routing problem is the *sub-path optimality* requirement, which will be invoked in the verification of the models in Section 3. The requirement says that a sub-path of a routing path is also a unique shortest path [Ben-Ameur and Gourdin, 2003, Bley and Koch, 2008]. Specifically, given an origin node $s \in \mathcal{S}$ and a node $i \in \mathcal{N}$, $i \neq s$, all demands which originate from s and transit i must go through the same incoming link to i .

Proposition 1. *The path length constraints in the demand-based model (9) imply the sub-path optimality constraints.*

Proof. Let $k_1, k_2 \in \mathcal{D}$ be two demands with $s_{k_1} = s_{k_2} = s$. Assume that they use two disjoint paths to traverse from node u to v . Demand k_1 uses path

$$P_i = (i_1, i_2) \rightarrow (i_2, i_3) \rightarrow \dots \rightarrow (i_{m-1}, i_m), (i_p, i_{p+1}) \in \mathcal{L}, p = 1, \dots, m-1,$$

where $i_1 = u$ and $i_m = v$, and demand k_2 uses path

$$P_j = (j_1, j_2) \rightarrow (j_2, j_3) \rightarrow \dots \rightarrow (j_{n-1}, j_n), (j_q, j_{q+1}) \in \mathcal{L}, q = 1, \dots, n-1,$$

where $j_1 = u$ and $j_n = v$.

By the definition of the routing decision variables,

$$x_{i_p i_{p+1}}^{k_1} = 1 \text{ and } x_{i_p i_{p+1}}^{k_2} = 0, \forall p \in \{1, \dots, m-1\}.$$

Then by constraints (6), on the one hand, since $x_{ij}^k = 1 \Rightarrow l_j^{s_k} = l_i^{s_k} + w_{ij}$, $\forall k \in \mathcal{D}, \forall (i, j) \in \mathcal{L}$, considering demand k_1 , it holds that

$$\begin{aligned} l_v^s &= l_u^s + l_{i_m}^u \\ &= l_u^s + l_{i_{m-1}}^u + w_{i_{m-1} i_m} \end{aligned}$$

$$\begin{aligned}
&= l_u^s + l_{i_{m-2}}^u + w_{i_{m-2}i_{m-1}} + w_{i_{m-1}i_m} \\
&= l_u^s + l_{i_1}^u + w_{i_1i_2} + \cdots + w_{i_{m-1}i_m} \\
&= l_u^s + l_{P_i}.
\end{aligned} \tag{10}$$

On the other hand, since $x_{ij}^k = 0 \wedge \sum_{h:(h,j) \in \mathcal{L}} x_{hj}^k = 1 \Rightarrow l_j^{sk} < l_i^{sk} + w_{ij}$ and $x_{ij}^k = 0 \Rightarrow l_j^{sk} \leq l_i^{sk} + w_{ij}$, $\forall k \in \mathcal{D}, \forall (i, j) \in \mathcal{L}$, it follows from considering k_2 that

$$\begin{aligned}
l_v^s &= l_u^s + l_{i_m}^u \\
&< l_u^s + l_{i_{m-1}}^u + w_{i_{m-1}i_m} \\
&\leq l_u^s + l_{i_{m-2}}^u + w_{i_{m-2}i_{m-1}} + w_{i_{m-1}i_m} \\
&\leq l_u^s + l_{i_1}^u + w_{i_1i_2} + \cdots + w_{i_{m-1}i_m} \\
&= l_u^s + l_{P_i}.
\end{aligned} \tag{11}$$

Clearly, (10) and (11) contradict each other, which means that the two demands cannot be routed over two different paths between two shared nodes. It is hence proved that the sub-path optimality constraints are satisfied. \square

The sub-path optimality constraints are thus not explicitly embedded into DBM. Mathematically, the constraints are represented as

$$\sum_{h:(h,i) \in \mathcal{L}} \max_{k \in \mathcal{D}_s} x_{hi}^k \leq 1, i \neq s, \forall s \in \mathcal{S}, \forall i \in \mathcal{N}, \tag{12}$$

which can be linearized, with a new set of variables $y_{ij}^s, s \in \mathcal{S}, (i, j) \in \mathcal{L}$, as

$$y_{ij}^s \geq x_{ij}^k, \forall k \in \mathcal{D}_s, \forall (i, j) \in \mathcal{L} \quad \text{and} \quad \sum_{h:(h,i) \in \mathcal{L}} y_{hi}^s \leq 1, i \neq s, \forall s \in \mathcal{S}, \forall i \in \mathcal{N}.$$

By Proposition 1, the demand-based model defined in (9) is equivalent to

$$\begin{aligned}
&\text{Optimize} && (8) \\
&\text{Subject to} && (4), (5), (7), (12), (1), (2), (3)
\end{aligned} \tag{13}$$

Model (13) will be invoked to verify the correctness of DBM in Section 3.

2.3. An Origin-Based Model

In Section 2.2, the unique shortest path routing problem is formulated as a demand-based model, which defines one routing decision variable for each link–demand pair. Based on the study of properties associated with the solution, it can be found that all routing paths of demands originating from the same node constitute a tree, rooted at the origin node [Zhang and Rodošek, 2005b]. Accordingly, a more natural formulation is to define one routing decision variable for each link–origin pair. For example, in Figure 2, instead of defining three routing decision variables for link (i, j) , one for each of the three demands sharing the same origin node s , the new formulation defines only one routing decision variable for link (i, j) , paired with origin node s .

Based on the above observation, an *origin-based* model (OBM) for the problem is formulated as follows.

- *Routing decision variables:*

$$y_{ij}^s \in \{0, 1\}, \forall s \in \mathcal{S}, \forall (i, j) \in \mathcal{L} \tag{14}$$

is equal to 1 if and only if the routing path of at least one demand originating from node s traverses link (i, j) . The number of this set of variables is $|\mathcal{S}||\mathcal{L}|$.

- *Auxiliary flow variables:*

$$f_{ij}^s \in [0, +\infty), \forall s \in \mathcal{S}, \forall (i, j) \in \mathcal{L} \tag{15}$$

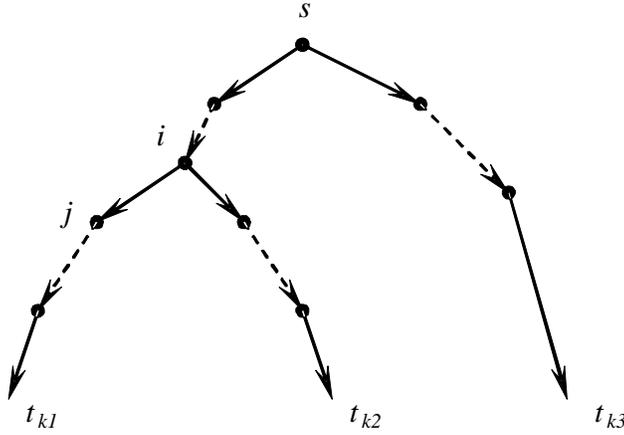


Figure 2: Illustration of the origin-based model

corresponds to the load of traffic flows originating from node s and traversing link (i, j) . The number of this set of variables is $|\mathcal{S}||\mathcal{L}|$.

- *Link weight variables:*

$$w_{ij} \in [w_{\min}, w_{\max}], \forall (i, j) \in \mathcal{L} \quad (16)$$

denotes the routing cost of link (i, j) . The number of this set of variables is $|\mathcal{L}|$.

- *Path length variables:*

$$l_i^s \in [0, +\infty), \forall s \in \mathcal{S}, \forall i \in \mathcal{N} \quad (17)$$

represents the length of the shortest path from origin node s to node i . In particular, $l_s^s = 0, \forall s \in \mathcal{S}$. The number of this set of variables is $|\mathcal{S}||\mathcal{N}|$.

- *Flow conservation constraints:* For each tree, at the root node, the difference between the sum of outgoing flows and that of incoming flows is the sum of bandwidths of all demands originating from the node; at the destination node of each demand originating from the root node, the difference between the sum of incoming flows and that of outgoing flows is the bandwidth of the demand; and the sum of incoming flows is equal to that of outgoing flows at other nodes.

$$\sum_{h:(h,i) \in \mathcal{L}} f_{hi}^s - \sum_{j:(i,j) \in \mathcal{L}} f_{ij}^s = \begin{cases} -d_s, & \text{if } i = s \\ d_k, & \text{if } i = t_k, \forall k \in \mathcal{D}_s \\ 0, & \text{otherwise} \end{cases}, \forall s \in \mathcal{S}, \forall i \in \mathcal{N}, \quad (18)$$

where $d_s = \sum_{k \in \mathcal{D}_s} d_k$. The number of this set of constraints is $|\mathcal{S}||\mathcal{N}|$.

- *Flow bound constraints:* For each tree, the total flow load over each link does not exceed the sum of all demand bandwidths originating from the root node and it is equal to zero if no demand originating from the root node is routed through the link.

$$f_{ij}^s \leq y_{ij}^s \sum_{k \in \mathcal{D}_s} d_k, \forall s \in \mathcal{S}, \forall (i, j) \in \mathcal{L}. \quad (19)$$

The number of this set of constraints is $|\mathcal{S}||\mathcal{L}|$.

- *Link capacity constraints:*

$$\sum_{s \in \mathcal{S}} f_{ij}^s \leq c_{ij}, \forall (i, j) \in \mathcal{L}. \quad (20)$$

The number of this set of constraints is $|\mathcal{L}|$.

- *Path uniqueness constraints:* For each tree, the number of incoming links with nonzero flows is equal to zero at the origin node, is equal to one at the destination node of a demand originating from the root node, and does not exceed one at any other node.

$$\sum_{h:(h,i) \in \mathcal{L}} y_{hi}^s \begin{cases} = 0, & \text{if } i = s \\ = 1, & \text{if } i \in \mathcal{T}_s \\ \leq 1, & \text{otherwise} \end{cases}, \forall s \in \mathcal{S}, \forall i \in \mathcal{N}. \quad (21)$$

The number of this set of constraints is $|\mathcal{S}||\mathcal{N}|$.

- *Path length constraints:* For each tree, the length of the unique shortest path to route a demand originating from the root node is less than that of any other possible path from the origin node to the destination node.

$$\left. \begin{aligned} y_{ij}^s = 0 \wedge \sum_{h:(h,j) \in \mathcal{L}} y_{hj}^s = 0 &\Rightarrow l_j^s \leq l_i^s + w_{ij} \\ y_{ij}^s = 0 \wedge \sum_{h:(h,j) \in \mathcal{L}} y_{hj}^s = 1 &\Rightarrow l_j^s < l_i^s + w_{ij} \\ y_{ij}^s = 1 &\Rightarrow l_j^s = l_i^s + w_{ij} \end{aligned} \right\}, \forall s \in \mathcal{S}, \forall (i, j) \in \mathcal{L}.$$

The logic constraints can be linearized as

$$\left. \begin{aligned} l_j^s &\leq l_i^s + w_{ij} - \varepsilon \left(\sum_{h:(h,j) \in \mathcal{L}} y_{hj}^s - y_{ij}^s \right) \\ l_j^s &\geq l_i^s + w_{ij} - M(1 - y_{ij}^s) \end{aligned} \right\}, \forall s \in \mathcal{S}, \forall (i, j) \in \mathcal{L}, \quad (22)$$

where ε and M are appropriate constants with $0 < \varepsilon \ll M$. The number of this set of constraints is $2|\mathcal{S}||\mathcal{L}|$.

- *Objective function:*

$$\max \sum_{(i,j) \in \mathcal{L}} \left(c_{ij} - \sum_{s \in \mathcal{S}} f_{ij}^s \right),$$

which is equivalent to

$$\min \sum_{(i,j) \in \mathcal{L}} \sum_{s \in \mathcal{S}} f_{ij}^s. \quad (23)$$

With all the above specification, the origin-based model is defined as

$$\begin{aligned} \text{OBM:} \quad & \text{Optimize} \quad (23) \\ & \text{Subject to} \quad (18), (19), (20), (21), (22), (14), (15), (16), (17) \end{aligned} \quad (24)$$

3. Model Validation

To verify the correctness of OBM defined in (24), a two-step procedure is adopted. DBM defined in (9) is first proved to be a correct model for the unique shortest path routing problem, by Propositions 2 and 5. Then, the former is shown equivalent to the latter concerning both the feasibility and the optimality of the problem, which implies that OBM is a correct model for the problem as well.

The first step is based on the correctness of the formulation for a relaxed problem, the *integer multi-commodity flow problem* [Ahuja et al., 1993] with sub-path optimality condition, which has been well studied

and a corresponding demand-based formulation has been recognized correct. Accordingly, the equivalence between OBM and DBM is built on the proof of the equivalence between two corresponding models for the relaxation, defined respectively in (25) and (26), by Propositions 8 and 9. The equivalence between OBM and DBM is then completed by showing the equivalence of the additional path length constraints of the two models in Proposition 11 and Proposition 12.

$$\begin{aligned} \mathbf{RDBM} : \quad & \text{Optimize} && (8) \\ & \text{Subject to} && (4), (5), (12), (1) \end{aligned} \tag{25}$$

$$\begin{aligned} \mathbf{ROBM} : \quad & \text{Optimize} && (23) \\ & \text{Subject to} && (18), (19), (20), (21), (14), (15) \end{aligned} \tag{26}$$

RDBM defined in (25) is actually a relaxation of (13), which is equivalent to (9) by Proposition 1.

3.1. Correctness of DBM

The correctness of DBM is verified by showing that of the equivalent model (13). Apparently, the difference between (13) and the relaxation (25) lies in the path length constraints (7) as well as the additional link weight variables (2) and the path length variables (3). To validate that (13) formulates the unique shortest path routing problem correctly, constraints (7) in (13) are proved to represent correctly the additional path length constraints. Specifically, the following two statements are shown correct. In (13), the path length constraints (7), combined with the flow conservation constraints (4), ensure that

1. the routing path of each demand is a shortest path and
2. the routing path of each demand is a unique path.

The two statements are verified in Propositions 2 and 5, respectively. Proposition 2 is validated by showing that the length of the routing path of each demand is less than or equal to that of any other possible path. The proof of Proposition 5 is built on, with two lemmas, the satisfaction of the single-path requirement by the relaxed problem RDBM given in (25), followed by the verification that the uniqueness requirement is satisfied by (13). Since the original logic constraints (6) are identical to the linearized ones (7), the proof is based on (6).

Proposition 2. *A routing path given by the solution to (13) is a shortest one.*

Proof. Assume that for demand k ,

$$P_i = (i_1, i_2) \rightarrow (i_2, i_3) \rightarrow \cdots \rightarrow (i_{m-1}, i_m), (i_p, i_{p+1}) \in \mathcal{L}, p = 1, \dots, m-1,$$

with $i_1 = s_k$ and $i_m = t_k$, is the assigned routing path and

$$P_j = (j_1, j_2) \rightarrow (j_2, j_3) \rightarrow \cdots \rightarrow (j_{n-1}, j_n), (j_q, j_{q+1}) \in \mathcal{L}, q = 1, \dots, n-1,$$

with $j_1 = s_k$ and $j_n = t_k$, is another possible and non-assigned path from s_k to t_k .

By the definition of the routing decision variables, $x_{i_p i_{p+1}}^k = 1$, $p = 1, \dots, m-1$ and there exists $\tilde{q} \in \{1, 2, \dots, n-1\}$ such that $x_{j_{\tilde{q}} j_{\tilde{q}+1}}^k = 0$. As a result, by constraints (6), on the one hand, since $x_{i_j}^k = 1 \Rightarrow l_j^{s_k} = l_i^{s_k} + w_{ij}$, $\forall (i, j) \in \mathcal{L}$,

$$l_{t_k}^{s_k} = l_{i_m}^{s_k} = l_{i_{m-1}}^{s_k} + w_{i_{m-1} i_m} = l_{i_1}^{s_k} + w_{i_1 i_2} + \cdots + w_{i_{m-1} i_m} = l_{P_i}.$$

On the other hand, since $x_{i_j}^k = 0 \Rightarrow l_j^{s_k} \leq l_i^{s_k} + w_{ij}$, $\forall (i, j) \in \mathcal{L}$,

$$l_{t_k}^{s_k} = l_{j_n}^{s_k} \leq l_{j_{n-1}}^{s_k} + w_{j_{n-1} j_n} \leq l_{j_1}^{s_k} + w_{j_1 j_2} + \cdots + w_{j_{n-1} j_n} = l_{P_j}.$$

It then follows that $l_{P_i} \leq l_{P_j}$, which means that the length of the routing path P_i is less than or equal to the length of any other possible path P_j to route demand k . It is hence proved that the routing path P_i is a shortest one. \square

To prove that the path uniqueness constraints are included implicitly in the model (13), they are shown to be satisfied by the relaxed problem (25) first.

Lemma 3. *An optimal solution to (25) contains no flow loops.*

Proof. Let $x_{ij}^{*k}, k \in \mathcal{D}, (i, j) \in \mathcal{L}$ be an optimal solution to (25). Then all constraints (4), (5), and (12) are satisfied by $x_{ij}^{*k}, k \in \mathcal{D}, (i, j) \in \mathcal{L}$. Specifically,

$$\sum_{h:(h,i) \in \mathcal{L}} x_{hi}^{*k} - \sum_{j:(i,j) \in \mathcal{L}} x_{ij}^{*k} = \begin{cases} -1, & \text{if } i = s_k \\ 1, & \text{if } i = t_k \\ 0, & \text{otherwise} \end{cases}, \forall k \in \mathcal{D}, \forall i \in \mathcal{N}, \quad (27)$$

$$\sum_{k \in \mathcal{D}} d_k x_{ij}^{*k} \leq c_{ij}, \forall (i, j) \in \mathcal{L}, \quad (28)$$

and

$$\sum_{h:(h,i) \in \mathcal{L}} \max_{k \in \mathcal{D}_s} x_{hi}^{*k} \leq 1, i \neq s, \forall s \in \mathcal{S}, \forall i \in \mathcal{N}. \quad (29)$$

Suppose that there exist $k_l \in \mathcal{D}$ and a loop $\mathcal{C} : (j_1, j_2) \rightarrow \dots \rightarrow (j_{n-1}, j_n) \rightarrow (j_n, j_{n+1}), (j_i, j_{i+1}) \in \mathcal{L}, i = 1, \dots, n$ with $j_{n+1} = j_1$ such that $x_{j_i j_{i+1}}^{*k_l} = 1, i = 1, \dots, n$. Clearly,

$$\sum_{h:(h,i) \in \mathcal{C}} x_{hi}^{*k_l} - \sum_{j:(i,j) \in \mathcal{C}} x_{ij}^{*k_l} = 0, \forall i \in \{j_1, \dots, j_n\}. \quad (30)$$

Denote

$$y_{ij}^{*k} = \begin{cases} x_{ij}^{*k}, & \text{if } k \neq k_l \text{ or } (i, j) \notin \mathcal{C} \\ 0, & \text{otherwise} \end{cases}, \forall k \in \mathcal{D}, \forall (i, j) \in \mathcal{L}. \quad (31)$$

Then $y_{ij}^{*k} \in \{0, 1\}, \forall k \in \mathcal{D}, \forall (i, j) \in \mathcal{L}$ and $y_{j_i j_{i+1}}^{*k_l} = 0, \forall (j_i, j_{i+1}) \in \mathcal{C}, i = 1, \dots, n$. In addition, it holds that by (27), (30), and (31),

$$\sum_{h:(h,i) \in \mathcal{L}} y_{hi}^{*k} - \sum_{j:(i,j) \in \mathcal{L}} y_{ij}^{*k} = \sum_{h:(h,i) \in \mathcal{L}} x_{hi}^{*k} - \sum_{j:(i,j) \in \mathcal{L}} x_{ij}^{*k} = \begin{cases} -1, & \text{if } i = s_k \\ 1, & \text{if } i = t_k \\ 0, & \text{otherwise} \end{cases}, \forall k \in \mathcal{D}, \forall i \in \mathcal{N},$$

by (28) and (31),

$$\sum_{k \in \mathcal{D}} d_k y_{ij}^{*k} \leq \sum_{k \in \mathcal{D}} d_k x_{ij}^{*k} \leq c_{ij}, \forall (i, j) \in \mathcal{L},$$

and by (29) and (31),

$$\sum_{h:(h,i) \in \mathcal{L}} \max_{k \in \mathcal{D}_s} y_{hi}^{*k} \leq \sum_{h:(h,i) \in \mathcal{L}} \max_{k \in \mathcal{D}_s} x_{hi}^{*k} \leq 1, i \neq s, \forall s \in \mathcal{S}, \forall i \in \mathcal{N}.$$

Hence, $y_{ij}^{*k}, k \in \mathcal{D}, (i, j) \in \mathcal{L}$ satisfies all constraints (4), (5), and (12) and is thus a feasible solution to (25). Furthermore, by (31), since $d_k > 0, \forall k \in \mathcal{D}$,

$$\begin{aligned} \sum_{(i,j) \in \mathcal{L}} \sum_{k \in \mathcal{D}} d_k y_{ij}^{*k} &= \sum_{(i,j) \in \mathcal{L} \setminus \mathcal{C} \vee k \in \mathcal{D}: k \neq k_l} d_k y_{ij}^{*k} + \sum_{(i,j) \in \mathcal{C}} d_{k_l} y_{ij}^{*k_l} \\ &= \sum_{(i,j) \in \mathcal{L} \setminus \mathcal{C} \vee k \in \mathcal{D}: k \neq k_l} d_k y_{ij}^{*k} \end{aligned}$$

$$\begin{aligned}
&= \sum_{(i,j) \in \mathcal{L} \setminus \mathcal{C} \vee k \in \mathcal{D}: k \neq k_l} d_k x_{ij}^{*k} \\
&< \sum_{(i,j) \in \mathcal{L} \setminus \mathcal{C} \vee k \in \mathcal{D}: k \neq k_l} d_k x_{ij}^{*k} + \sum_{(i,j) \in \mathcal{C}} d_{k_l} x_{ij}^{*k_l} \\
&= \sum_{(i,j) \in \mathcal{L}} \sum_{k \in \mathcal{D}} d_k x_{ij}^{*k},
\end{aligned}$$

which implies that $x_{ij}^{*k}, k \in \mathcal{D}, (i,j) \in \mathcal{L}$ is not an optimal solution to (25). This results in contradiction and completes the proof. \square

With Lemma 3, the path uniqueness constraints are proved to be satisfied by (25) in Lemma 4.

Lemma 4. *The path uniqueness constraints are satisfied in the model (25).*

Proof. By Lemma 3, in (25), the flow conservation constraints (4) at origin nodes are equivalent to

$$\sum_{h:(h,i) \in \mathcal{L}} x_{hi}^k = 0 \text{ and } \sum_{j:(i,j) \in \mathcal{L}} x_{ij}^k = 1, i = s_k, \forall k \in \mathcal{D} \quad (32)$$

and at destination nodes are equivalent to

$$\sum_{h:(h,i) \in \mathcal{L}} x_{hi}^k = 1 \text{ and } \sum_{j:(i,j) \in \mathcal{L}} x_{ij}^k = 0, i = t_k, \forall k \in \mathcal{D}. \quad (33)$$

Since $x_{ij}^k \in \{0, 1\}, \forall k \in \mathcal{D}, \forall (i,j) \in \mathcal{L}$, constraints (32) restrict that there is one and only one outgoing link with a nonzero flow from the origin node of demand k . Analogously, constraints (33) say that there is one and only one incoming link with a nonzero flow into the destination node of demand k . In addition, constraints (4) ensure that, at each intermediate node, the number of incoming links with nonzero flows equals that of outgoing links with nonzero flows. Hence, for each demand k , the number of routing paths is no more than one. The path uniqueness constraints are hence satisfied. \square

Proposition 5. *The uniqueness of the resulting shortest path for any demand is satisfied by the path length constraints in (13).*

Proof. Since (13) is a reduction of (25), the solution to the routing decision variables $x_{ij}^k, k \in \mathcal{D}, (i,j) \in \mathcal{L}$ of (13) is a solution to those of (25) as well.

By Lemma 4, there is only one routing path for each demand. Suppose that for demand k ,

$$P_i = (i_1, i_2) \rightarrow (i_2, i_3) \rightarrow \cdots \rightarrow (i_{m-1}, i_m), (i_p, i_{p+1}) \in \mathcal{L}, p = 1, \dots, m-1,$$

with $i_1 = s_k$ and $i_m = t_k$, is the assigned routing path and

$$P_j = (j_1, j_2) \rightarrow (j_2, j_3) \rightarrow \cdots \rightarrow (j_{n-1}, j_n), (j_q, j_{q+1}) \in \mathcal{L}, q = 1, \dots, n-1,$$

with $j_1 = s_k$ and $j_n = t_k$, is another possible and non-assigned path from s_k to t_k .

By the definition of the routing decision variables, $x_{i_p i_{p+1}}^k = 1, p = 1, \dots, m-1$. Hence, by (6), since $x_{ij}^k = 1 \Rightarrow l_j^{s_k} = l_i^{s_k} + w_{ij}, \forall (i,j) \in \mathcal{L}$,

$$l_{t_k}^{s_k} = l_{i_m}^{s_k} = l_{i_{m-1}}^{s_k} + w_{i_{m-1} i_m} = l_{i_1}^{s_k} + w_{i_1 i_2} + \cdots + w_{i_{m-1} i_m} = l_{P_i}. \quad (34)$$

As both P_i and P_j are paths from s_k to t_k , they finally merge at a node r with $r = i_{\tilde{p}} = j_{\tilde{q}}, \tilde{p} \in \{2, 3, \dots, m-1, m\}, \tilde{q} \in \{2, 3, \dots, n-1, n\}$. Then by the definition of the routing decision variables, $x_{i_{\tilde{p}-1} r}^k = 1$ and $x_{j_{\tilde{q}-1} r}^k = 0$. In addition, on the one hand, $\sum_{h:(h,r) \in \mathcal{L}} x_{hr}^k \leq 1$ and on the other hand, $\sum_{h:(h,r) \in \mathcal{L}} x_{hr}^k \geq x_{i_{\tilde{p}-1} r}^k + x_{j_{\tilde{q}-1} r}^k = 1$.

Hence, $\sum_{h:(h,r) \in \mathcal{L}} x_{hr}^k = 1$. As a result, by (6),

$$\begin{aligned}
l_{t_k}^{sk} &= l_{j_n}^{sk} \\
&= l_{j_{n-1}}^{sk} + w_{j_{n-1}j_n} \\
&= l_r^{sk} + w_{rj_{\bar{q}+1}} + \cdots + w_{j_{n-1}j_n} \\
&< l_{j_{\bar{q}-1}}^{sk} + w_{j_{\bar{q}-1}r} + w_{rj_{\bar{q}+1}} + \cdots + w_{j_{n-1}j_n} \\
&\leq l_{j_1}^{sk} + w_{j_1j_2} + \cdots + w_{j_{n-1}j_n} \\
&= l_{P_j}.
\end{aligned} \tag{35}$$

It then follows from (34) and (35) that $l_{P_i} < l_{P_j}$. This proves that path P_i is the unique shortest path to route demand k . \square

Corollary 6. *The demand-based model (9) is a correct one for the unique shortest path routing problem.*

Proof. By Propositions 2 and 5, (13) is a correct model for the unique shortest path routing problem. Hence, as an equivalent model to (13), (9) is a correct model for the problem as well. \square

3.2. Correctness of OBM

By the proof of Lemma 4, the flow conservation constraints (4) are identical to

$$\left. \begin{aligned}
\sum_{h:(h,i) \in \mathcal{L}} x_{hi}^k = 0, \quad \sum_{j:(i,j) \in \mathcal{L}} x_{ij}^k = 1, \quad &\text{if } i = s_k \\
\sum_{h:(h,i) \in \mathcal{L}} x_{hi}^k = 1, \quad \sum_{j:(i,j) \in \mathcal{L}} x_{ij}^k = 0, \quad &\text{if } i = t_k \\
\sum_{h:(h,i) \in \mathcal{L}} x_{hi}^k = \sum_{j:(i,j) \in \mathcal{L}} x_{ij}^k, \quad &\text{otherwise}
\end{aligned} \right\}, \forall k \in \mathcal{D}, \forall i \in \mathcal{N}. \tag{36}$$

Hence, RDBM defined in (25) is equivalent to

$$\begin{aligned}
&\text{Optimize} && (8) \\
&\text{Subject to} && (36), (5), (12), (1)
\end{aligned} \tag{37}$$

Lemma 7. *In (26), the flow conservation constraints at origin nodes are equivalent to*

$$\sum_{h:(h,i) \in \mathcal{L}} f_{hi}^s = 0 \text{ and } \sum_{j:(i,j) \in \mathcal{L}} f_{ij}^s = \sum_{k \in \mathcal{D}_s} d_k, i = s, \forall s \in \mathcal{S}.$$

Proof. On the one hand, by (15), if $i = s$,

$$\sum_{h:(h,i) \in \mathcal{L}} f_{hi}^s \geq 0, \forall s \in \mathcal{S}.$$

On the other hand, by (19) and (21), if $i = s$,

$$\sum_{h:(h,i) \in \mathcal{L}} f_{hi}^s \leq \sum_{h:(h,i) \in \mathcal{L}} \left(y_{hi}^s \sum_{k \in \mathcal{D}_s} d_k \right) = \sum_{k \in \mathcal{D}_s} d_k \sum_{h:(h,i) \in \mathcal{L}} y_{hi}^s = 0, \forall s \in \mathcal{S}.$$

Hence,

$$\sum_{h:(h,i) \in \mathcal{L}} f_{hi}^s = 0, i = s, \forall s \in \mathcal{S}.$$

It can then be derived directly from (18) that

$$\sum_{j:(i,j) \in \mathcal{L}} f_{ij}^s = \sum_{k \in \mathcal{D}_s} d_k, i = s, \forall s \in \mathcal{S}.$$

\square

By Lemma 7, the flow conservation constraints (18) are identical to

$$\left. \begin{aligned} \sum_{h:(h,i) \in \mathcal{L}} f_{hi}^s &= 0, \quad \sum_{j:(i,j) \in \mathcal{L}} f_{ij}^s = d_s, \quad \text{if } i = s \\ \sum_{h:(h,i) \in \mathcal{L}} f_{hi}^s - \sum_{j:(i,j) \in \mathcal{L}} f_{ij}^s &= d_k, \quad \text{if } i = t_k, \forall k \in \mathcal{D}_s \\ \sum_{h:(h,i) \in \mathcal{L}} f_{hi}^s - \sum_{j:(i,j) \in \mathcal{L}} f_{ij}^s &= 0, \quad \text{otherwise} \end{aligned} \right\}, \forall s \in \mathcal{S}, \forall i \in \mathcal{N}, \quad (38)$$

where $d_s = \sum_{k \in \mathcal{D}_s} d_k$. As a result, ROBM defined in (26) is equivalent to

$$\begin{aligned} &\text{Optimize} \quad (23) \\ &\text{Subject to} \quad (38), (19), (20), (21), (14), (15) \end{aligned} \quad (39)$$

In the following, largely based on validating the equivalence between the two identical models (39) and (37), it is shown that ROBM (26) and RDBM (25) are equivalent to each other concerning the feasibility of the relaxed problem.

Proposition 8. *There is a solution to (26) if (25) is feasible.*

Proof. Let $x_{ij}^{*k}, k \in \mathcal{D}, (i, j) \in \mathcal{L}$ be a feasible solution to (25), and so to (37). All constraints (36), (5), and (12) are hence satisfied by $x_{ij}^{*k}, k \in \mathcal{D}, (i, j) \in \mathcal{L}$. Denote

$$y_{ij}^{*s} = \max_{k \in \mathcal{D}_s} x_{ij}^{*k}, \forall s \in \mathcal{S}, \forall (i, j) \in \mathcal{L} \quad (40)$$

and

$$f_{ij}^{*s} = \sum_{k \in \mathcal{D}_s} d_k x_{ij}^{*k}, \forall s \in \mathcal{S}, \forall (i, j) \in \mathcal{L}. \quad (41)$$

Apparently, $y_{ij}^{*s} \in \{0, 1\}$ and $f_{ij}^{*s} \in [0, +\infty)$, $\forall s \in \mathcal{S}, \forall (i, j) \in \mathcal{L}$.

By (36), for any $k \in \mathcal{D}$, if $i = s_k$, $\sum_{h:(h,i) \in \mathcal{L}} x_{hi}^{*k} = 0$ and $\sum_{j:(i,j) \in \mathcal{L}} x_{ij}^{*k} = 1$. Then by (41), for any $s \in \mathcal{S}$, if $i = s$,

$$\sum_{h:(h,i) \in \mathcal{L}} f_{hi}^{*s} = \sum_{h:(h,i) \in \mathcal{L}} \sum_{k \in \mathcal{D}_s} d_k x_{hi}^{*k} = \sum_{k \in \mathcal{D}_s} d_k \sum_{h:(h,i) \in \mathcal{L}} x_{hi}^{*k} = 0$$

and

$$\sum_{j:(i,j) \in \mathcal{L}} f_{ij}^{*s} = \sum_{j:(i,j) \in \mathcal{L}} \sum_{k \in \mathcal{D}_s} d_k x_{ij}^{*k} = \sum_{k \in \mathcal{D}_s} d_k \sum_{j:(i,j) \in \mathcal{L}} x_{ij}^{*k} = \sum_{k \in \mathcal{D}_s} d_k.$$

Hence, at origin nodes, constraints (38) in (39) are satisfied by $f_{ij}^{*s}, s \in \mathcal{S}, (i, j) \in \mathcal{L}$.

Also by (36), for any $k_l \in \mathcal{D}$, if $i = t_{k_l}$, $\sum_{h:(h,i) \in \mathcal{L}} x_{hi}^{*k_l} - \sum_{j:(i,j) \in \mathcal{L}} x_{ij}^{*k_l} = 1$ and $\sum_{h:(h,i) \in \mathcal{L}} x_{hi}^{*k_l} - \sum_{j:(i,j) \in \mathcal{L}} x_{ij}^{*k_l} = 0, \forall k \in \mathcal{D}_{s_{k_l}}, k \neq k_l$. Then by (41), for any $k_l \in \mathcal{D}$, if $s = s_{k_l}$ and $i = t_{k_l}$,

$$\begin{aligned} \sum_{h:(h,i) \in \mathcal{L}} f_{hi}^{*s} - \sum_{j:(i,j) \in \mathcal{L}} f_{ij}^{*s} &= \sum_{h:(h,i) \in \mathcal{L}} \sum_{k \in \mathcal{D}_s} d_k x_{hi}^{*k} - \sum_{j:(i,j) \in \mathcal{L}} \sum_{k \in \mathcal{D}_s} d_k x_{ij}^{*k} \\ &= \sum_{k \in \mathcal{D}_s} \left(d_k \sum_{h:(h,i) \in \mathcal{L}} x_{hi}^{*k} \right) - \sum_{k \in \mathcal{D}_s} \left(d_k \sum_{j:(i,j) \in \mathcal{L}} x_{ij}^{*k} \right) \\ &= \sum_{k \in \mathcal{D}_s} d_k \left(\sum_{h:(h,i) \in \mathcal{L}} x_{hi}^{*k} - \sum_{j:(i,j) \in \mathcal{L}} x_{ij}^{*k} \right) \\ &= d_{k_l} \left(\sum_{h:(h,i) \in \mathcal{L}} x_{hi}^{*k_l} - \sum_{j:(i,j) \in \mathcal{L}} x_{ij}^{*k_l} \right) \\ &\quad + \sum_{k \in \mathcal{D}_s, k \neq k_l} d_k \left(\sum_{h:(h,i) \in \mathcal{L}} x_{hi}^{*k} - \sum_{j:(i,j) \in \mathcal{L}} x_{ij}^{*k} \right) \\ &= d_{k_l}. \end{aligned}$$

Hence, at destination nodes, constraints (38) in (39) are satisfied.

Similarly, by (36), for any $k \in \mathcal{D}$, if $i \neq s_k$ and $i \neq t_k$, $\sum_{h:(h,i) \in \mathcal{L}} x_{hi}^{*k} - \sum_{j:(i,j) \in \mathcal{L}} x_{ij}^{*k} = 0$. Then by (41), for any $s \in \mathcal{D}$ and $i \in \mathcal{N}$, if $i \neq s$ and $i \notin \mathcal{D}_s$,

$$\begin{aligned} \sum_{h:(h,i) \in \mathcal{L}} f_{hi}^{*s} - \sum_{j:(i,j) \in \mathcal{L}} f_{ij}^{*s} &= \sum_{h:(h,i) \in \mathcal{L}} \sum_{k \in \mathcal{D}_s} d_k x_{hi}^{*k} - \sum_{j:(i,j) \in \mathcal{L}} \sum_{k \in \mathcal{D}_s} d_k x_{ij}^{*k} \\ &= \sum_{k \in \mathcal{D}_s} \left(d_k \sum_{h:(h,i) \in \mathcal{L}} x_{hi}^{*k} \right) - \sum_{k \in \mathcal{D}_s} \left(d_k \sum_{j:(i,j) \in \mathcal{L}} x_{ij}^{*k} \right) \\ &= \sum_{k \in \mathcal{D}_s} d_k \left(\sum_{h:(h,i) \in \mathcal{L}} x_{hi}^{*k} - \sum_{j:(i,j) \in \mathcal{L}} x_{ij}^{*k} \right) \\ &= 0. \end{aligned}$$

Hence, at other nodes, constraints (38) in (39) are satisfied as well.

By (40) and (41), it holds that for any $s \in \mathcal{S}$ and $(i, j) \in \mathcal{L}$,

$$f_{ij}^{*s} = \sum_{k \in \mathcal{D}_s} d_k x_{ij}^{*k} \leq \sum_{k \in \mathcal{D}_s} d_k \max_{k \in \mathcal{D}_s} x_{ij}^{*k} = \max_{k \in \mathcal{D}_s} x_{ij}^{*k} \sum_{k \in \mathcal{D}_s} d_k = y_{ij}^{*s} \sum_{k \in \mathcal{D}_s} d_k.$$

Constraints (19) in (39) are then satisfied by $y_{ij}^{*s}, f_{ij}^{*s}, s \in \mathcal{S}, (i, j) \in \mathcal{L}$.

By (41) and (5),

$$\sum_{s \in \mathcal{S}} f_{ij}^{*s} = \sum_{s \in \mathcal{S}} \sum_{k \in \mathcal{D}_s} d_k x_{ij}^{*k} = \sum_{k \in \mathcal{D}} d_k x_{ij}^{*k} \leq c_{ij}, \forall (i, j) \in \mathcal{L}.$$

Constraints (20) in (39) are thus satisfied by $f_{ij}^{*s}, s \in \mathcal{S}, (i, j) \in \mathcal{L}$.

By (36), for any $k \in \mathcal{D}$, if $i = s_k$, $\sum_{h:(h,i) \in \mathcal{L}} x_{hi}^{*k} = 0$ and so for any $k \in \mathcal{D}$ and $(h, i) \in \mathcal{L}$, if $i = s_k$, $x_{hi}^{*k} = 0$. Then by (40), for any $s \in \mathcal{S}$, if $i = s$,

$$\sum_{h:(h,i) \in \mathcal{L}} y_{hi}^{*s} = \sum_{h:(h,i) \in \mathcal{L}} \max_{k \in \mathcal{D}_s} x_{hi}^{*k} = \sum_{h:(h,i) \in \mathcal{L}} \max_{k \in \mathcal{D}_s} 0 = 0.$$

Hence, at origin nodes, constraints (21) in (39) are satisfied by $y_{ij}^{*s}, s \in \mathcal{S}, (i, j) \in \mathcal{L}$.

Also by (36), for any $k_l \in \mathcal{D}$, if $i = t_{k_l}$, $\sum_{h:(h,i) \in \mathcal{L}} x_{hi}^{*k_l} = 1$. Then by (40), for any $k_l \in \mathcal{D}$, if $s = s_{k_l}$ and $i = t_{k_l}$,

$$\sum_{h:(h,i) \in \mathcal{L}} y_{hi}^{*s} = \sum_{h:(h,i) \in \mathcal{L}} \max_{k \in \mathcal{D}_s} x_{hi}^{*k} \geq \sum_{h:(h,i) \in \mathcal{L}} x_{hi}^{*k_l} = 1.$$

In addition, by (12), for any $k_l \in \mathcal{D}$, if $i = t_{k_l}$, $\sum_{h:(h,i) \in \mathcal{L}} \max_{k \in \mathcal{D}_{s_{k_l}}} x_{hi}^{*k} \leq 1$. Then by (40), for any $k_l \in \mathcal{D}$, if $s = s_{k_l}$ and $i = t_{k_l}$,

$$\sum_{h:(h,i) \in \mathcal{L}} y_{hi}^{*s} = \sum_{h:(h,i) \in \mathcal{L}} \max_{k \in \mathcal{D}_s} x_{hi}^{*k} \leq 1.$$

Thus, for any $k \in \mathcal{D}$, if $s = s_k$ and $i = t_k$,

$$\sum_{h:(h,i) \in \mathcal{L}} y_{hi}^{*s} = 1.$$

Hence, at destination nodes, constraints (21) in (39) are satisfied.

Similarly, by (12), for any $s \in \mathcal{S}$ and $i \in \mathcal{N}$, if $i \neq s$, $\sum_{h:(h,i) \in \mathcal{L}} \max_{k \in \mathcal{D}_s} x_{hi}^{*k} \leq 1$. Then by (40), for any $s \in \mathcal{S}$ and $i \in \mathcal{N}$, if $i \neq s$ and $i \notin \mathcal{T}_s$,

$$\sum_{h:(h,i) \in \mathcal{L}} y_{hi}^{*s} = \sum_{h:(h,i) \in \mathcal{L}} \max_{k \in \mathcal{D}_s} x_{hi}^{*k} \leq 1.$$

Hence, at other nodes, constraints (21) in (39) are satisfied as well.

Since all constraints (38), (19), (20), and (21) in (39) are satisfied by y_{ij}^{*s} and $f_{ij}^{*s}, s \in \mathcal{S}, (i, j) \in \mathcal{L}$, it is a corresponding feasible solution to (39) and so to (26), of the feasible solution to (25), i.e., $x_{ij}^{*k}, k \in \mathcal{D}, (i, j) \in \mathcal{L}$.

□

Proposition 9. *There is a solution to (25) if (26) is feasible.*

Proof. Let $y_{ij}^{*s}, f_{ij}^{*s}, s \in \mathcal{S}, (i, j) \in \mathcal{L}$ be a feasible solution to (26), and so to (39). Then all constraints (38), (19), (20), and (21) are satisfied by $y_{ij}^{*s}, f_{ij}^{*s}, s \in \mathcal{S}, (i, j) \in \mathcal{L}$. Denote

$$x_{ht_k}^{*k} = y_{ht_k}^{*s_k}, \forall k \in \mathcal{D}, \forall (h, t_k) \in \mathcal{L}.$$

By constraints (21) at destination nodes, $\sum_{h:(h,t_k) \in \mathcal{L}} y_{ht_k}^{*s_k} = 1, \forall k \in \mathcal{D}$. Hence, for any $k \in \mathcal{D}$, there exists $h : (h, t_k) \in \mathcal{L}$ such that $y_{ht_k}^{*s_k} = 1$.

For any $k \in \mathcal{D}$ and $(i, j) \in \mathcal{L}$, if $(i, j) \neq (i, t_k)$, x_{ij}^{*k} is assigned as follows:

$$\begin{aligned} \textbf{Initialize} \quad & x_{ij}^{*k} \leftarrow 0, \quad \forall k \in \mathcal{D}, \forall (i, j) \in \mathcal{L}, (i, j) \neq (i, t_k) \\ \textbf{For} \quad & k \in \mathcal{D} \\ \quad \mathbf{0} \quad & i \leftarrow t_k \\ \quad \textbf{Do} \quad & \text{find } h : (h, i) \in \mathcal{L}, \text{ such that } y_{hi}^{*s_k} = 1 \\ & x_{hi}^{*k} \leftarrow 1 \\ & i \leftarrow h \\ \quad \textbf{Until} \quad & i = s_k \end{aligned} \tag{42}$$

In (42), at each iteration of the inner loop, on the one hand, by (38),

$$\sum_{h:(h,i) \in \mathcal{L}} f_{hi}^{*s_k} \geq \sum_{j:(t_k,j) \in \mathcal{L}} f_{t_kj}^{*s_k} + d_k \geq d_k > 0, \forall k \in \mathcal{D}$$

and by (19),

$$\sum_{h:(h,i) \in \mathcal{L}} f_{hi}^{*s_k} \leq \sum_{h:(h,i) \in \mathcal{L}} \left(y_{hi}^{*s_k} \sum_{\bar{k} \in \mathcal{D}_{s_k}} d_{\bar{k}} \right) = \sum_{\bar{k} \in \mathcal{D}_{s_k}} d_{\bar{k}} \sum_{h:(h,i) \in \mathcal{L}} y_{hi}^{*s_k}, \forall k \in \mathcal{D}.$$

Hence,

$$\sum_{h:(h,i) \in \mathcal{L}} y_{hi}^{*s_k} > 0, \forall k \in \mathcal{D}.$$

On the other hand, by (21),

$$\sum_{h:(h,i) \in \mathcal{L}} y_{hi}^{*s_k} \leq 1, \forall k \in \mathcal{D}.$$

Then at node i in each iteration of the inner loop,

$$\sum_{h:(h,i) \in \mathcal{L}} y_{hi}^{*s_k} = 1, \forall k \in \mathcal{D}.$$

Therefore, at node i , there exists h with $(h, i) \in \mathcal{L}$ such that $x_{hi}^{*k} = y_{hi}^{*s_k} = 1$. Moreover, by (38), for any $k \in \mathcal{D}$, the assigning process terminates at node s_k .

Clearly, by (42),

$$x_{ij}^{*k} \in \{0, 1\} \text{ and } x_{ij}^{*k} \leq y_{ij}^{*s_k}, \forall k \in \mathcal{D}, \forall (i, j) \in \mathcal{L}.$$

Then by (21) at origin nodes, if $i = s_k$,

$$\sum_{h:(h,i) \in \mathcal{L}} x_{hi}^{*k} \leq \sum_{h:(h,i) \in \mathcal{L}} y_{hi}^{*s_k} = 0, \forall k \in \mathcal{D}.$$

In addition, the assigning process terminates at node $s_k, \forall k \in \mathcal{D}$ and so if $i = s_k$,

$$\sum_{j:(i,j) \in \mathcal{L}} x_{ij}^{*k} = 1, \forall k \in \mathcal{D}.$$

Hence, at origin nodes, constraints (36) in (37) are satisfied by $x_{ij}^{*k}, k \in \mathcal{D}, (i, j) \in \mathcal{L}$. Similarly, by (21) at destination nodes, if $i = t_k$,

$$\sum_{h:(h,i) \in \mathcal{L}} x_{hi}^{*k} = \sum_{h:(h,i) \in \mathcal{L}} y_{hi}^{*s_k} = 1, \forall k \in \mathcal{D}.$$

Moreover, by (42), for any $k \in \mathcal{D}$ and $(i, j) \in \mathcal{L}$, if $i = t_k, x_{ij}^{*k} = 0$. Then, if $i = t_k$,

$$\sum_{j:(i,j) \in \mathcal{L}} x_{ij}^{*k} = 0, \forall k \in \mathcal{D}.$$

Hence, at destination nodes, constraints (36) in (37) are satisfied by $x_{ij}^{*k}, k \in \mathcal{D}, (i, j) \in \mathcal{L}$. Also by (42), at node i in each iteration of the inner loop,

$$\sum_{h:(h,i) \in \mathcal{L}} x_{hi}^{*k} = \sum_{j:(i,j) \in \mathcal{L}} x_{ij}^{*k} = 1, \forall k \in \mathcal{D},$$

and at any other node $\tilde{i} \in \mathcal{N}, \tilde{i} \neq s_k, \tilde{i} \neq t_k$,

$$\sum_{h:(h,\tilde{i}) \in \mathcal{L}} x_{h\tilde{i}}^{*k} = \sum_{j:(\tilde{i},j) \in \mathcal{L}} x_{\tilde{i}j}^{*k} = 0, \forall k \in \mathcal{D}.$$

Then if $i \neq s_k$ and $i \neq t_k$,

$$\sum_{h:(h,i) \in \mathcal{L}} x_{hi}^{*k} - \sum_{j:(i,j) \in \mathcal{L}} x_{ij}^{*k} = 0, \forall k \in \mathcal{D}.$$

Hence, at other nodes, constraints (36) in (37) are satisfied by $x_{ij}^{*k}, k \in \mathcal{D}, (i, j) \in \mathcal{L}$.

By (38) and (20),

$$\sum_{k \in \mathcal{D}} d_k x_{ij}^{*k} = \sum_{s \in \mathcal{S}} \sum_{k \in \mathcal{D}_s} d_k x_{ij}^{*k} \leq \sum_{s \in \mathcal{S}} f_{ij}^{*s} \leq c_{ij}, \forall (i, j) \in \mathcal{L}.$$

Hence, constraints (5) in (37) are satisfied by $x_{ij}^{*k}, k \in \mathcal{D}, (i, j) \in \mathcal{L}$.

By (42), $x_{ij}^{*k} \leq y_{ij}^{*s_k}, \forall k \in \mathcal{D}, \forall (i, j) \in \mathcal{L}$. Then it follows from (21) that if $i \neq s$,

$$\sum_{h:(h,i) \in \mathcal{L}} \max_{k \in \mathcal{D}_s} x_{hi}^{*k} \leq \sum_{h:(h,i) \in \mathcal{L}} \max_{k \in \mathcal{D}_s} y_{hi}^{*s_k} = \sum_{h:(h,i) \in \mathcal{L}} y_{hi}^{*s} \leq 1, \forall s \in \mathcal{S}, \forall i \in \mathcal{N}.$$

Hence, constraints (12) in (37) are satisfied by $x_{ij}^{*k}, k \in \mathcal{D}, (i, j) \in \mathcal{L}$.

Since all constraints (36), (5), and (12) in (37) are satisfied by $x_{ij}^{*k}, k \in \mathcal{D}, (i, j) \in \mathcal{L}$, it is a corresponding feasible solution to (37) and so to (25), of the feasible solution to (26), i.e., $y_{ij}^{*s}, f_{ij}^{*s}, s \in \mathcal{S}, (i, j) \in \mathcal{L}$. \square

Theorem 10. *The two models (26) and (25) are equivalent to each other concerning the feasibility of the relaxed problem.*

Proof. The statement follows directly from Propositions 8 and 9. \square

Based on the proof of the equivalence between the two models (26) and (25) for the relaxed problem, the equivalence between (24) and (9), concerning the feasibility of the unique shortest path routing problem, is to be verified.

Proposition 11. *For each solution satisfying the path length constraints in (9), there is a corresponding solution satisfying the path length constraints in (24).*

Proof. Let $x_{ij}^{*k}, k \in \mathcal{D}, (i, j) \in \mathcal{L}, w_{ij}^*, (i, j) \in \mathcal{L}$, and $l_i^{*s}, s \in \mathcal{S}, i \in \mathcal{N}$ be a feasible solution to (9) and so to (13) by Proposition 1. Since (25) is a relaxation of (13), $x_{ij}^{*k}, k \in \mathcal{D}, (i, j) \in \mathcal{L}$ is a feasible solution to (25) and satisfies all the corresponding constraints. Denote

$$y_{ij}^{*s} = \max_{k \in \mathcal{D}_s} x_{ij}^{*k}, \forall s \in \mathcal{S}, \forall (i, j) \in \mathcal{L}.$$

Clearly, $y_{ij}^{*s} \in \{0, 1\}, \forall s \in \mathcal{S}, \forall (i, j) \in \mathcal{L}$. Also for any $s \in \mathcal{S}$ and $(i, j) \in \mathcal{L}$, there are three cases:

- Case 1: $x_{ij}^{*k} = 0$ and $\sum_{h:(h,j) \in \mathcal{L}} x_{hj}^{*k} = 0, \forall k \in \mathcal{D}_s$;
- Case 2: $x_{ij}^{*k} = 0, \forall k \in \mathcal{D}_s$ and $\exists k_l \in \mathcal{D}_s, \sum_{h:(h,j) \in \mathcal{L}} x_{hj}^{*k_l} = 1$;
- Case 3: $\exists k_l \in \mathcal{D}_s, x_{ij}^{*k_l} = 1$.

For Case 1, on the one hand, for any $k \in \mathcal{D}_s$, (7) can be simplified as

$$l_j^{sk} \leq l_i^{sk} + w_{ij} \quad \text{and} \quad l_j^{sk} \geq l_i^{sk} + w_{ij} - M.$$

On the other hand, since $y_{ij}^{*s} = \max_{k \in \mathcal{D}_s} x_{ij}^{*k} = 0$ and $\sum_{h:(h,j) \in \mathcal{L}} y_{hj}^{*s} = 0$, (22) can be simplified as

$$l_j^s \leq l_i^s + w_{ij} \quad \text{and} \quad l_j^s \geq l_i^s + w_{ij} - M.$$

Hence, the simplified constraints of (22) are identical to those of (7) in this case.

For Case 2, on the one hand, (7) can be simplified as

$$l_j^{sk_l} \leq l_i^{sk_l} + w_{ij} - \varepsilon \quad \text{and} \quad l_j^{sk_l} \geq l_i^{sk_l} + w_{ij} - M$$

and for any $k \in \mathcal{D}_s$ such that $\sum_{h:(h,j) \in \mathcal{L}} x_{hj}^{*k} = 0$,

$$l_j^{sk} \leq l_i^{sk} + w_{ij} \quad \text{and} \quad l_j^{sk} \geq l_i^{sk} + w_{ij} - M.$$

On the other hand, since $y_{ij}^{*s} = \max_{k \in \mathcal{D}_s} x_{ij}^{*k} = 0$ but $\sum_{h:(h,j) \in \mathcal{L}} y_{hj}^{*s} = 1$, (22) can be simplified as

$$l_j^s \leq l_i^s + w_{ij} - \varepsilon \quad \text{and} \quad l_j^s \geq l_i^s + w_{ij} - M.$$

Hence, the simplified constraints of (22) are identical to those of (7) in this case.

For Case 3, on the one hand, by the sub-path optimality constraints, for any $k \in \mathcal{D}_s$, if $x_{ij}^{*k} = 0$, $\sum_{h:(h,j) \in \mathcal{L}} x_{hj}^{*k} = 0$. Then (7) can be simplified as

$$l_j^{sk_l} \leq l_i^{sk_l} + w_{ij} \quad \text{and} \quad l_j^{sk_l} \geq l_i^{sk_l} + w_{ij}$$

and for any $k \in \mathcal{D}_s$ such that $x_{ij}^{*k} = 0$,

$$l_j^{sk} \leq l_i^{sk} + w_{ij} \quad \text{and} \quad l_j^{sk} \geq l_i^{sk} + w_{ij} - M.$$

On the other hand, since $y_{ij}^{*s} = \max_{k \in \mathcal{D}_s} x_{ij}^{*k} = 1$ and $\sum_{h:(h,j) \in \mathcal{L}} y_{hj}^{*s} = 1$, (22) can be simplified as

$$l_j^s \leq l_i^s + w_{ij} \quad \text{and} \quad l_j^s \geq l_i^s + w_{ij}.$$

Hence, the simplified constraints of (22) are identical to those of (7) as well.

Since in all the three cases, $y_{ij}^{*s}, s \in \mathcal{S}, (i, j) \in \mathcal{L}$ results in the same path length constraints for OBM defined in (24) as those resulting from $x_{ij}^{*k}, k \in \mathcal{D}, (i, j) \in \mathcal{L}$ for DBM given in (9), there is a corresponding feasible solution satisfying the path length constraints (22) in (24), provided that there is a feasible solution satisfying the path length constraints (7) in (9). \square

Proposition 12. *For each solution satisfying the path length constraints in (24), there is a corresponding solution satisfying the path length constraints in (9).*

Proof. Let $y_{ij}^{*s}, f_{ij}^{*s}, s \in \mathcal{S}, (i, j) \in \mathcal{L}, w_{ij}^*, (i, j) \in \mathcal{L}$, and $l_i^{*s}, s \in \mathcal{S}, i \in \mathcal{N}$ be a feasible solution to (24). Then since (26) is a relaxation of (24), $y_{ij}^{*s}, f_{ij}^{*s}, s \in \mathcal{S}, (i, j) \in \mathcal{L}$ is a feasible solution to (26) as well and so satisfies all the corresponding constraints. By the assigning process (42), let $x_{ij}^{*k}, k \in \mathcal{D}, (i, j) \in \mathcal{L}$ be the corresponding solution to (25).

for any $s \in \mathcal{S}$ and $(i, j) \in \mathcal{L}$, there are three cases:

- Case 1: $y_{ij}^{*s} = 0$ and $\sum_{h:(h,j) \in \mathcal{L}} y_{hj}^{*s} = 0$;
- Case 2: $y_{ij}^{*s} = 0$ and $\sum_{h:(h,j) \in \mathcal{L}} y_{hj}^{*s} = 1$;
- Case 3: $y_{ij}^{*s} = 1$.

For Case 1, (22) can be simplified as

$$l_j^s \leq l_i^s + w_{ij} \quad \text{and} \quad l_j^s \geq l_i^s + w_{ij} - M.$$

By (42), for any $k \in \mathcal{D}_s$, $x_{ij}^{*k} = 0$ and $\sum_{h:(h,j) \in \mathcal{L}} x_{hj}^{*k} = 0$. Then for any $k \in \mathcal{D}_s$, (7) can be simplified as

$$l_j^{sk} \leq l_i^{sk} + w_{ij} \quad \text{and} \quad l_j^{sk} \geq l_i^{sk} + w_{ij} - M.$$

Hence, the simplified constraints of (7) are identical to those of (22) in this case.

For Case 2, (22) can be simplified as

$$l_j^s \leq l_i^s + w_{ij} - \varepsilon \quad \text{and} \quad l_j^s \geq l_i^s + w_{ij} - M.$$

By (42), for any $k \in \mathcal{D}_s$, $x_{ij}^{*k} = 0$ and $\exists k_l \in \mathcal{D}_s$, $\sum_{h:(h,j) \in \mathcal{L}} x_{hj}^{*k_l} = 1$. Then (7) can be simplified as

$$l_j^{sk_l} \leq l_i^{sk_l} + w_{ij} - \varepsilon \quad \text{and} \quad l_j^{sk_l} \geq l_i^{sk_l} + w_{ij} - M$$

and for any $k \in \mathcal{D}_s$ such that $\sum_{h:(h,j) \in \mathcal{L}} x_{hj}^{*k} = 0$,

$$l_j^{sk} \leq l_i^{sk} + w_{ij} \quad \text{and} \quad l_j^{sk} \geq l_i^{sk} + w_{ij} - M.$$

Hence, the simplified constraints of (7) are identical to those of (22) in this case.

For Case 3, (22) can be simplified as

$$l_j^s \leq l_i^s + w_{ij} \quad \text{and} \quad l_j^s \geq l_i^s + w_{ij}.$$

By (42), there exists $k_l \in \mathcal{D}_s$ such that $x_{ij}^{*k_l} = 1$ and $\sum_{h:(h,j) \in \mathcal{L}} x_{hj}^{*k_l} = 1$. In addition, by the sub-path optimality constraints, for any $k \in \mathcal{D}_s$, if $x_{ij}^{*k} = 0$, $\sum_{h:(h,j) \in \mathcal{L}} x_{hj}^{*k} = 0$. Then (7) can be simplified as

$$l_j^{sk_l} \leq l_i^{sk_l} + w_{ij} \quad \text{and} \quad l_j^{sk_l} \geq l_i^{sk_l} + w_{ij}$$

and for any $k \in \mathcal{D}_s$ such that $x_{ij}^{*k} = 0$,

$$l_j^{sk} \leq l_i^{sk} + w_{ij} \quad \text{and} \quad l_j^{sk} \geq l_i^{sk} + w_{ij} - M.$$

Hence, the simplified constraints of (7) are identical to those of (22) as well.

Since in all the three cases, x_{ij}^{*k} , $k \in \mathcal{D}$, $(i, j) \in \mathcal{L}$ results in the same path length constraints for DBM defined in (9) as those resulting from y_{ij}^{*s} , $s \in \mathcal{S}$, $(i, j) \in \mathcal{L}$ for OBM given in (24), there is a corresponding feasible solution satisfying the path length constraints (7) in (9), provided that there is a feasible solution satisfying the path length constraints (22) in (24). \square

Corollary 13. *Concerning the feasibility of the unique shortest path routing problem, the path length constraints in (24) are equivalent to those in (9).*

Theorem 14. *The two models (24) and (9) are equivalent to each other concerning the feasibility of the unique shortest path routing problem.*

Proof. By Proposition 1, (9) is equivalent to (13).

In addition, as discussed at the beginning of Section 3, (24) is a reduction of (26) and (13) is a reduction of (25). Besides the additional link weight and path length variables, the difference between (24) and (26) exists in the path length constraints (22) and that between (13) and (25) exists in the counterparts (7).

By Theorem 10, (26) and (25) are equivalent concerning the feasibility of the relaxed problem. By Corollary 13, the path length constraints in (24) are equivalent to the counterparts in (9) and so those in (13). Therefore, (24) is equivalent to (13) and so to (9), concerning the feasibility of the unique shortest path routing problem. \square

Theorem 15. *The two models (24) and (9) are equivalent to each other concerning the optimality of the unique shortest path routing problem.*

Proof. The statement follows directly from Theorem 14 by constructing the corresponding optimal solutions between (24) and (9). \square

Corollary 16. *The origin-based model defined in (24) is a correct one for the unique shortest path routing problem.*

4. Model Comparison

Concerning the unique shortest path routing problem, it is shown that the routing performance resulting from the proposed exact formulations is considerably better than that from using the default methods, by testing on 30 randomly generated data instances with combinations of different parameter scenarios, given in Table 1. As shown in Table 2, the resulting average maximum utilization is 30.94% of that from the hop-count method and 45.54% of that from the inv-cap method. This hence indicates the significant gain obtained by formulating the problem exactly and solving it optimally.

Table 1: Details of data sets tested

ID	$ \mathcal{N} $	$ \mathcal{L} $	$ \mathcal{D} $	ID	$ \mathcal{N} $	$ \mathcal{L} $	$ \mathcal{D} $	ID	$ \mathcal{N} $	$ \mathcal{L} $	$ \mathcal{D} $
1	10	22	3	11	10	48	20	21	30	236	29
2	10	26	5	12	10	44	50	22	50	128	3
3	10	24	9	13	30	80	3	23	50	132	25
4	10	24	10	14	30	78	15	24	50	130	49
5	10	26	20	15	30	82	29	25	50	238	3
6	10	24	50	16	30	136	3	26	50	238	25
7	10	46	3	17	30	144	15	27	50	238	49
8	10	46	5	18	30	142	29	28	50	644	3
9	10	46	9	19	30	236	3	29	50	648	25
10	10	46	10	20	30	234	15	30	50	642	49

Table 2: Maximum utilization of three link weight setting methods on 30 data instances

ID	Hop-Count	Inv-Cap	DBM/OBM	ID	Hop-Count	Inv-Cap	DBM/OBM	ID	Hop-Count	Inv-Cap	DBM/OBM
1	0.9388	0.9388	0.7001	11	3.5157	1.4149	0.8910	21	6.1026	3.2431	0.9793
2	1.0699	1.0699	0.6999	12	5.0635	3.1937	0.9415	22	0.6998	0.6998	0.6998
3	2.9715	2.8623	0.6999	13	3.1295	1.9148	0.7000	23	2.3427	2.3427	0.7000
4	1.0082	1.2481	0.7001	14	3.1675	3.2731	0.9485	24	1.0511	0.7980	0.9847
5	3.1982	2.0476	0.9612	15	2.1683	1.3442	0.7171	25	1.5016	0.9174	0.9870
6	1.3049	1.0275	0.7613	16	10.0219	7.1661	0.7000	26	2.9699	2.9282	0.7363
7	1.6335	1.0111	0.7000	17	2.3227	1.3396	0.7150	27	2.4900	2.1997	0.9741
8	4.4814	1.2053	0.7000	18	2.0908	1.3664	0.9816	28	1.1029	0.4797	0.7261
9	2.3639	2.3668	0.9454	19	0.7944	0.5500	0.8684	29	2.3573	1.8183	0.9986
10	1.1693	1.1187	0.7000	20	3.9783	1.2725	0.9699	30	3.0459	1.2322	0.9843

Between the two formulations, compared with DBM (9), OBM (24) has advantages on both constraint structure for applying constraint generation algorithms and model size, which may be taken into consideration in developing algorithms to solve the problem.

4.1. Constraint Structure

The constraint structures of the two models are illustrated in Figures 3 and 4, respectively.

In Figure 3, the first row corresponds to the link capacity constraints (5), the next four rows to the flow conservation constraints (4), and the last four rows to the path length constraints (7). Accordingly, columns represent variables.

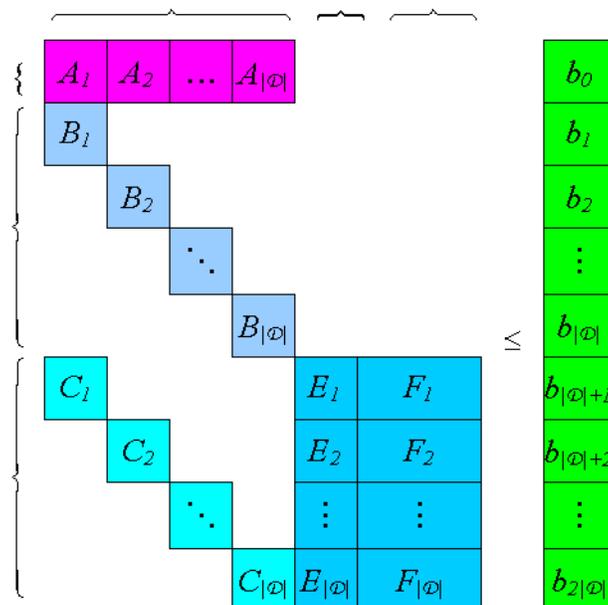


Figure 3: Constraint structure of DBM

As can be noted, among the three sets of constraints, the flow conservation constraints and the link capacity constraints contain only the routing decision variables, whereas the path length constraints couple the routing decision variables with the link weight variables and the path length variables. Hence, constraint generation algorithms such as Benders decomposition method [Benders, 1962] may be adopted to solve the problem. The problem is then decomposed into one integer-programming master problem and one linear-programming subproblem. The master problem deals with the flow conservation constraints and the link capacity constraints, and so contains the routing decision variables only. Accordingly, the subproblem copes with the path length constraints.

Analogously, in Figure 4, the first four rows correspond to the path uniqueness constraints (21), the next four rows to the flow bound constraints (19), the third four rows to the flow conservation constraints (18), the next row to the link capacity constraints (20), and the last four rows to the path length constraints (22). Columns represent variables.

As can be observed, although DBM has a simpler constraint structure, OBM has more flexibility to apply decomposition algorithms to solve the problem.

As shown in Figure 4, with OBM, the problem can be globally decomposed into one master problem and two subproblems, instead of one master problem and one subproblem as with DBM. The master problem contains only the routing decision variables and the path uniqueness constraints. The first subproblem deals with the auxiliary flow variables and the second with the link weight and the path length variables. In addition, the master problem can be further decomposed, with one independent subproblem corresponding to each origin node.

4.2. Model Size

Compared with DBM, OBM defines explicitly the auxiliary flow variables and the flow bound constraints accordingly. However, in general, $|\mathcal{S}| \ll |\mathcal{D}|$ and the size of the latter is significantly smaller than that of

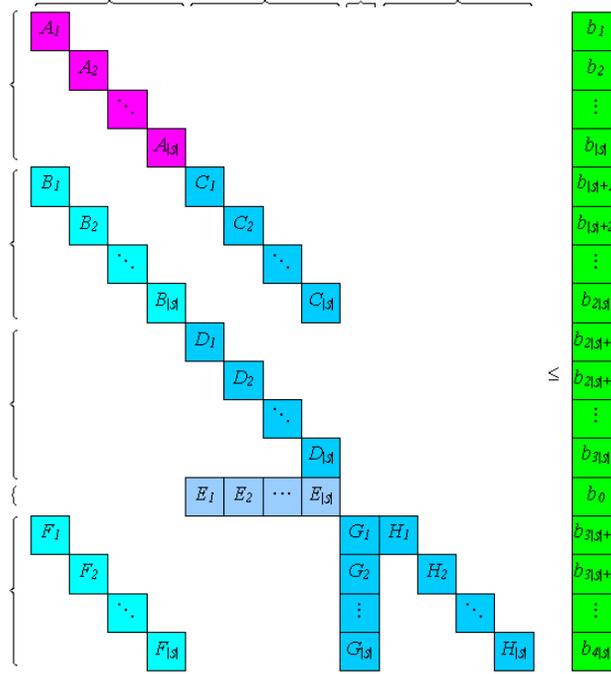


Figure 4: Constraint structure of OBM

the former. The model sizes of the two formulations are shown in Table 3, where $\#Variables$ represents the number of variables and $\#Constraints$ denotes the number of constraints.

More concretely, the model sizes of both the original and the master problems of the two formulations on a randomly generated data instance with $|\mathcal{N}| = 50$, $|\mathcal{L}| = 642$, $|\mathcal{D}| = 1000$, and $|\mathcal{S}| = 50$ are provided in Table 4.

As can be noted from Table 4 that, with OBM, the number of variables of the original problem decreases from over 600000 to 64200 and that of constraints drops from over 1000000 to less than 38000. In addition, with OBM, the numbers of both variables and constraints of the master problem reduce by 20 times.

Table 3: Model sizes of the two formulations

Model	$\#Variables$	$\#Constraints$
DBM	$ \mathcal{D} \mathcal{L} + \mathcal{S} \mathcal{N} + \mathcal{L} $	$ \mathcal{D} \mathcal{N} + 2 \mathcal{D} \mathcal{L} + \mathcal{L} $
OBM	$2 \mathcal{S} \mathcal{L} + \mathcal{S} \mathcal{N} + \mathcal{L} $	$2 \mathcal{S} \mathcal{N} + 3 \mathcal{S} \mathcal{L} + \mathcal{L} $

Table 4: Model sizes of the two formulations on a reasonably large data instance

	Original Problem		Master Problem	
	$\#Variables$	$\#Constraints$	$\#Variables$	$\#Constraints$
DBM	645142	1334642	642000	50642
OBM	64200	37742	32100	2500

As a conclusion, compared with DBM, OBM has a smaller model size and a more flexible constraint structure for decomposition algorithms such as the Benders decomposition method to solve the problem.

5. Conclusion

With the aim of an exact solution approach to the unique shortest path routing problem on average data instances arising from real-world applications, two mathematical formulations with a polynomial number of constraints are developed explicitly. A demand-based formulation is first introduced, based on the study of

the relationship between the length of a shortest path and the weights of links that the path traverses. The problem is further formulated as an origin-based model by analyzing solution properties of the problem. The two formulations are then mathematically proved correct and equivalent to each other concerning both the feasibility and the optimality of the problem. Based on the study of the constraint structures and model sizes of the two formulations, the origin-based formulation is identified to be a more appropriate one for decomposition algorithms such as the Benders decomposition method to solve the problem.

The two formulations may be generalized to other network flow and network routing problems such as transportation and energy supply. Future work may exist in looking into further improvements concerning both problem formulation, including strength of formulation, and solution algorithm on the efficiency of the solution approach proposed. In particular, the focus may be on possible enhancements from three perspectives: the closeness between the initial solution and the final solution to the master problem, the strength of cuts generated at each iteration, and the efficiency of an algorithm to solve the integer-programming master problem. For instance, redundant constraints may be generated to tighten the feasible region of the initial master problem, strategies such as the active set method may be applied to strengthen the cuts generated from the subproblems, and schemes such as the Lagrangian relaxation method may be embedded into the solution algorithm to improve the efficiency of solving the master problem at each iteration.

Acknowledgements

This research has been supported in part by the Curtin Sarawak FoB Research Fund (008018). The author is grateful to Ken Darby-Dowman, Eleni Hadjiconstantinou, and István Maros for carefully reviewing a thesis, on which this paper is based. The author greatly appreciates Fernando Ordóñez for helping to go through an earlier version of the paper and for providing valuable feedback. The author wishes to thank Professor Tom Archibald and Professor Jonathan Crook for all the communications in the process of handling the paper. The author would also like to extend his sincere gratitude to the associate editor and two anonymous referees for taking time to review an earlier version of this paper and to provide constructive comments and suggestions, which helped to improve the paper further.

References

- Ittai Abraham, Amos Fiat, Andrew V. Goldberg, and Renato F. Werneck. Highway dimension, shortest paths, and provably efficient algorithms. In *Proceedings of the Twenty-First Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA'10, pages 782–793. Society for Industrial and Applied Mathematics, 2010.
- Ravindra K. Ahuja and James B. Orlin. Inverse optimization. *Operations Research*, 49(5):771–783, 2001.
- Ravindra K. Ahuja and James B. Orlin. Combinatorial algorithms for inverse network flow problems. *Networks*, 40(4):181–187, 2002.
- Ravindra K. Ahuja, Thomas L. Magnanti, and James B. Orlin. *Network Flows: Theory, Algorithms, and Applications*. Prentice Hall, Inc., Englewood Cliffs, NJ, 1993.
- Ayşegül Altın, Bernard Fortz, Mikkel Thorup, and Hakan Ümit. Intra-domain traffic engineering with shortest path routing protocols. *Annals of Operations Research*, 204(1):65–95, 2013.
- Ayşegül Altın, Pietro Belotti, and Mustafa Ç. Pınar. OSPF routing with optimal oblivious performance ratio under polyhedral demand uncertainty. *Optimization and Engineering*, 11(3):395–422, 2010.
- E. Amaldi, A. Capone, and L.G. Gianoli. Energy-aware IP traffic engineering with shortest path routing. *Computer Networks*, 57(6):1503–1517, 2013.

- D. Applegate and E. Cohen. Making routing robust to changing traffic demands: Algorithms and evaluation. *IEEE/ACM Transactions on Networking*, 14(6):1193–1206, 2006.
- Simon Balon, Fabian Skivée, and Guy Leduc. How well do traffic engineering objective functions meet TE requirements? In Fernando Boavida, Thomas Plagemann, Burkhard Stiller, Cedric Westphal, and Edmundo Monteiro, editors, *NETWORKING 2006. Networking Technologies, Services, and Protocols; Performance of Computer and Communication Networks; Mobile and Wireless Communications Systems: 5th International IFIP-TC6 Networking Conference, Coimbra, Portugal, May 15-19, 2006. Proceedings*, pages 75–86. Springer Berlin Heidelberg, 2006.
- Cynthia Barnhart, Christopher A. Hane, and Pamela H. Vance. Using branch-and-price-and-cut to solve origin-destination integer multicommodity flow problems. *Operations Research*, 48(2):318–326, 2000.
- Richard Bellman. On a routing problem. *Quarterly of Applied Mathematics*, 16(1):87–90, 1958.
- Walid Ben-Ameur and Eric Gourdin. Internet routing and related topology issues. *SIAM Journal on Discrete Mathematics*, 17(1):18–49, 2003.
- J. F. Benders. Partitioning procedures for solving mixed-variables programming problems. *Numerische Mathematik*, 4(1):238–252, 1962.
- Dimitri Bertsekas and Robert Gallager. *Data Networks*. Prentice-Hall, Inc., Upper Saddle River, NJ, 1992.
- Uyless Black. *IP Routing Protocols: RIP, OSPF, BGP, PNNI and Cisco Routing Protocols*. Prentice Hall PTR, Upper Saddle River, NJ, 2000.
- Andreas Bley. Inapproximability results for the inverse shortest paths problem with integer lengths and unique shortest paths. *Networks*, 50(1):29–36, 2007.
- Andreas Bley. Approximability of unsplittable shortest path routing problems. *Networks*, 54(1):23–46, 2009.
- Andreas Bley and Thorsten Koch. Integer programming approaches to access and backbone IP network planning. In *Modeling, Simulation and Optimization of Complex Processes*, pages 87–110. Springer, Berlin, 2008.
- Andreas Bley, Bernard Fortz, Eric Gourdin, Kaj Holmberg, Olivier Klopfenstein, Michał Pióro, Artur Tomaszewski, and Hakan Ümit. Optimization of OSPF routing in IP networks. In Arie Koster and Xavier Muñoz, editors, *Graphs and Algorithms in Communication Networks*, Texts in Theoretical Computer Science. An EATCS Series, pages 199–240. Springer Berlin Heidelberg, 2010.
- L.S. Buriol, M.G.C. Resende, C.C. Ribeiro, and M. Thorup. A hybrid genetic algorithm for the weight setting problem in OSPF/IS-IS routing. *Networks*, 46(1):36–56, 2005.
- L.S. Buriol, M.G.C. Resende, and M. Thorup. Speeding up dynamic shortest-path algorithms. *INFORMS Journal on Computing*, 20(2):191–204, 2008.
- D. Burton and Ph. L. Toint. On an instance of the inverse shortest paths problem. *Mathematical Programming*, 53(1):45–61, 1992.
- D. Burton and Ph. L. Toint. On the use of an inverse shortest paths algorithm for recovering linearly correlated costs. *Mathematical Programming*, 63(1):1–22, 1994.
- A. Cianfrani, V. Eramo, M. Listanti, M. Polverini, and A. V. Vasilakos. An OSPF-integrated routing strategy for QoS-aware energy saving in IP backbone networks. *IEEE Transactions on Network and Service Management*, 9(3):254–267, 2012.

- Cisco Systems Inc. *Internetworking Technologies Handbook*. Cisco Press, Indianapolis, IN, 3rd edition, 2000.
- E. W. Dijkstra. A note on two problems in connexion with graphs. *Numerische Mathematik*, 1(1):269–271, 1959.
- Yefim Dinitz, Naveen Garg, and X. Michel Goemans. On the single-source unsplittable flow problem. *Combinatorica*, 19(1):17–41, 1999.
- M. Ericsson, M. G. C. Resende, and P. M. Pardalos. A genetic algorithm for the weight setting problem in OSPF routing. *Journal of Combinatorial Optimization*, 6(3):299–333, 2002.
- Andras Farago, Aron Szentesi, and Balazs Szvatovszki. Inverse optimization in high-speed networks. *Discrete Applied Mathematics*, 129(1):83 – 98, 2003.
- Anja Feldmann, Albert Greenberg, Carsten Lund, Nick Reingold, Jennifer Rexford, and Fred True. Deriving traffic demands for operational IP networks: Methodology and experience. *IEEE/ACM Transactions on Networking*, 9(3):265–280, 2001.
- D. R. Ford and D. R. Fulkerson. *Flows in Networks*. Princeton University Press, Princeton, NJ, 2010.
- B. Fortz and M. Thorup. Internet traffic engineering by optimizing OSPF weights. In *INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, volume 2, pages 519–528, 2000.
- Bernard Fortz and Mikkel Thorup. Increasing internet capacity using local search. *Computational Optimization and Applications*, 29(1):13–48, 2004.
- Bernard Fortz and Hakan Umit. Efficient techniques and tools for intra-domain traffic engineering. *International Transactions in Operational Research*, 18(3):359–376, 2011.
- Frederic Giroire, Stephane Perennes, and Issam Tahiri. On the complexity of equal shortest path routing. *Networks*, 65(4):344–352, 2015.
- D. Hock, M. Hartmann, M. Menth, and C. Schwartz. Optimizing unique shortest paths for resilient routing and fast reroute in IP-based networks. In *2010 IEEE Network Operations and Management Symposium (NOMS)*, pages 309–316, 2010.
- Kaj Holmberg and Di Yuan. A Lagrangian heuristic based branch-and-bound approach for the capacitated network design problem. *Operations Research*, 48(3):461–481, 2000.
- Kaj Holmberg and Di Yuan. Optimization of internet protocol network design and routing. *Networks*, 43(1):39–53, 2004.
- Stavros G. Kolliopoulos and Clifford Stein. Approximation algorithms for single-source unsplittable flow. *SIAM Journal on Computing*, 31(3):919–946, 2001.
- F.Y.-S. Lin and J.L. Wang. Minimax open shortest path first routing algorithms in networks supporting the SMDS service. In *Proceedings of IEEE International Conference on Communications*, volume 2, pages 666–670, 1993.
- John T. Moy. *OSPF: Anatomy of an Internet Routing Protocol*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, 1998.
- Eueung Mulyana and Ulrich Killat. An alternative genetic algorithm to optimize OSPF weights. In *Proceedings of Internet Traffic Engineering and Traffic Management, 15th ITC Specialist Seminar*, 2002.

- Kyungchul Park, Seokhoon Kang, and Sungsoo Park. An integer programming approach to the bandwidth packing problem. *Management Science*, 42(9):1277–1291, 1996.
- M. Pióro, Á. Szentesi, J. Harmatos, A. Jüttner, P. Gajowniczek, and S. Kozdrowski. On open shortest path first related network optimisation problems. *Performance Evaluation*, 48(1-4):201–223, 2002.
- Michal Pióro and Deepankar Medhi. *Routing, Flow, and Capacity Design in Communication and Computer Networks*. Morgan Kaufmann Publishers Inc., San Francisco, CA, 2004.
- K. G. Ramakrishnan and Manoel A. Rodrigues. Optimal routing in shortest-path data networks. *Bell Labs Technical Journal*, 6(1):117–138, 2001.
- G. Ramalingam and Thomas Reps. An incremental algorithm for a generalization of the shortest-path problem. *Journal of Algorithms*, 21(2):267–305, 1996.
- Martin Skutella. Approximating the single source unsplittable min-cost flow problem. *Mathematical Programming*, 91(3):493–514, 2002.
- S. Srivastava, G. Agrawal, M. Pioro, and D. Medhi. Determining link weight system under various objectives for OSPF networks using a Lagrangian relaxation-based approach. *IEEE Transactions on Network and Service Management*, 2(1):9–18, 2005.
- A.S. Tanenbaum and D. Wetherall. *Computer Networks*. Pearson Prentice Hall, Boston, MA, 2011.
- Thomas M. Thomas. *OSPF Network Design Solutions*. Cisco Press, Indianapolis, IN, 2003.
- Hao Wang, Haiyong Xie, Lili Qiu, Yang Richard Yang, Yin Zhang, and Albert Greenberg. COPE: Traffic engineering in dynamic networks. *SIGCOMM Computer Communication Review*, 36(4):99–110, 2006.
- Yufei Wang, Zheng Wang, and Leah Zhang. Internet traffic engineering without full mesh overlaying. In *INFOCOM 2001. Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, volume 1, pages 565–571, 2001.
- Shaoji Xu and Jianzhong Zhang. An inverse problem of the weighted shortest path problem. *Japan Journal of Industrial and Applied Mathematics*, 12(1):47–59, 1995.
- F. Benjamin Zhan and Charles E. Noon. Shortest path algorithms: An evaluation using real road networks. *Transportation Science*, 32(1):65–73, 1998.
- Changyong Zhang. Comparison on objective functions of the unique shortest path routing problem. In *Proceedings of the Eighth INFORMS Telecommunications Conference*, Dallas, Texas, March 30-April 2, 2006.
- Changyong Zhang and Robert Rodošek. Modelling and constraint hardness characterisation of the unique-path OSPF weight setting problem. In *Lecture Notes in Computer Science*, volume 3514, pages 804–811. Springer Berlin Heidelberg, 2005a.
- Changyong Zhang and Robert Rodošek. Two mathematically equivalent models of the unique-path OSPF weight setting problem. In *Lecture Notes in Computer Science*, volume 3421, pages 318–326. Springer Berlin Heidelberg, 2005b.
- Jianzhong Zhang and Zhenhong Liu. Calculating some inverse linear programming problems. *Journal of Computational and Applied Mathematics*, 72(2):261 – 273, 1996.