Journal of Graph Algorithms and Applications
http://jgaa.info/ vol. 23, no. 1, pp. 93–110 (2019)
DOI: 10.7155/jgaa.00485

Approximation Algorithm for Cycle-Star Hub Network Design Problems and Cycle-Metric Labeling Problems

Yuko Kuroki[1] *Tomomi Matsui*[2]

[1]Department of Computer Science,
Graduate School of Information Science and Technology,
The University of Tokyo, Tokyo, Japan
[2]Department of Industrial Engineering and Economics,
Graduate School of Engineering,
Tokyo Institute of Technology, Tokyo, Japan

Abstract

We consider a single allocation hub-and-spoke network design problem which allocates each non-hub node to exactly one of the given hub nodes in order to minimize the total transportation cost. This paper deals with a cycle-star hub network design problem, in which the hubs are located in a cycle. This problem is essentially equivalent to a cycle-metric labeling problem. It is useful in the design of networks in telecommunications and airline transportation systems. We propose a $2(1-1/h)$-approximation algorithm, where h denotes the number of hub nodes. Our algorithm solves a linear relaxation problem and employs a dependent rounding procedure. We analyze our algorithm by approximating a given cycle-metric matrix using a convex combination of Monge matrices.

Submitted: August 2017	Reviewed: May 2018	Revised: June 2018	Reviewed: July 2018
Revised: August 2018	Accepted: August 2018	Final: August 2018	Published: January 2019
Article type: Regular paper		Communicated by: M.S. Rahman, H.-C. Yen and S.-H. Poon	

Research supported by JSPS KAKENHI Grant Numbers 26285045 and 26242027.

E-mail addresses: ykuroki@ms.k.u-tokyo.ac.jp (Yuko Kuroki) matsui.t.af@m.titech.ac.jp (Tomomi Matsui)

1 Introduction

Hub-and-spoke networks are used in airline transportation systems, delivery systems, and telecommunication systems. Hub networks have an important role when there are many origins and destinations. Hub facilities work as switching points for flows. In order to reduce the transportation costs and set-up costs in a large network, each non-hub node is allocated to exactly one of the hubs instead of assigning every origin-destination pair directly.

Hub location problems (HLPs) involve locating the hubs and designing hub networks so as to minimize the total transportation cost. HLPs were first formulated as quadratic integer programming problems by O'Kelly [26]. Since HLPs were proposed by O'Kelly, many researches on HLPs have been conducted in various applications (see [3, 9, 12, 14, 24, 27] for example).

In this paper, we discuss the situation in which the locations of the hubs are given, and deal with a problem, called a *single allocation hub-and-spoke network design problem*, which involves finding a connection from the non-hubs to the given hubs in order to minimize the total transportation cost. Sohn and Park [29, 30] proposed a polynomial time exact algorithm for a problem with two hubs and proved the NP-completness of the problem even if the number of hubs is equal to three. Thus, for problems with more than three hubs, heuristic algorithms or approximation algorithms have been proposed. An α-approximation algorithm for a minimization problem finds a feasible solution whose objective value is within a ratio α of the optimal value. Iwasa et al. [21] proposed a simple 3-approximation algorithm and a randomized 2-approximation algorithm under the assumptions of triangle inequality. They also proposed a $(5/4)$-approximation algorithm for the special case in which the number of hubs is 3. Ando and Matsui [5] deal with the case in which all the nodes are embedded in a 2-dimensional plane and the transportation cost of an edge per unit flow is proportional to the Euclidean distance between the end nodes of the edge. They proposed a randomized $(1 + 2/\pi)$-approximation algorithm. Saito et al. [28] discussed some facets of polytopes corresponding to the convex hull of feasible solutions of the problem.

Fundamental HLPs assume a full interconnection between hubs. Recently, several researches considered incomplete hub networks, which arise especially in telecommunication systems (see [10, 11, 4, 8] for example). These models are useful when the set-up costs of hub links are considerably large or full interconnection is not required. In some researches , the hub networks were assumed to constitute a particular structure such as a tree [22, 15, 16, 19], star [25, 31, 32], path [18], or cycle [17].

In this paper, we consider a single allocation hub-and-spoke network design problem in which the given hubs are located in a cycle. We call this problem the *cycle-star hub network design problem* (see Figure 1). When the number of hubs is three, the hub network becomes a 3-cycle network and constitutes a complete graph. Therefore, the 4/3-approximation algorithm for a complete 3-hub network proposed in [21] is valid for this special case. In this paper we propose a $2(1 - 1/h)$-approximation algorithm, which is used when a set of h

hubs forms an h-cycle. To the best of our knowledge, our algorithm is the first approximation algorithm for this problem.

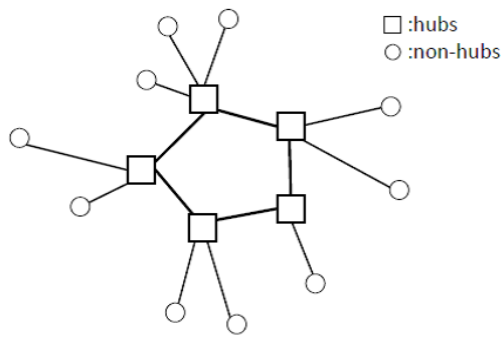

Figure 1: Cycle-star hub network with $h = 5$ hubs.

It is worth pointing out that a single allocation hub-and-spoke network design problem is essentially equivalent to the *metric labeling problem* introduced by Kleinberg and Tardos in [23], which has connections to Markov random field and classification problems that arise in computer vision and related areas. Formally, the metric labeling problem takes as input an n-vertex undirected graph $G(V, E)$ with a nonnegative weight function w on the edges, a set L of labels with metric distance function $d : L \times L \to R$ associated with them, and an assignment cost $c(v, a)$ for each vertex $v \in V$ and label $a \in L$. In the problem, we are asked to find an assignment for every object $v \in V$ to a label $a \in L$, while minimizing the total labeling costs $Q(f)$. The labeling cost $Q(f)$ is based on the contribution of two sets of terms. For each object $v \in V$, *vertex labeling cost* is denoted by $c(v, f(v))$. A vertex labeling cost $c(v, a)$ express an estimate of its likelihood of having each label $a \in L$. These likelihoods are observed from some heuristic preprocessing of the data. For each edge $e = \{u, v\} \in E$, *edge separation cost* is denoted by $w(\{u, v\}) \cdot d(f(u), f(v))$. The weights of the edges express a prior estimate on relationships among objects; if u and v are deemed to be related, then we would like them to be assigned close or identical labels. A distance $d(a, b)$ for $a, b \in L$ represents how similar label a and b are. Thus,

$$Q(f) = \sum_{u \in V} c(u, f(u)) + \sum_{\{u,v\} \in E} w(\{u, v\}) d(f(u), f(v))$$

and the goal is to find a labeling $f : V \to L$ minimizing $Q(f)$. Due to the simple structure and variety of applications, the metric labeling has received much attention since its introduction by Kleinberg and Tardos [23]. From the results in [20] and [23], we have $O(\log h)$-approximation algorithm, where h is the number of labels (hubs). Chuzhoy and Naor [13] showed that there is no polynomial time approximation algorithm with a constant ratio for the problem unless P = NP.

We deal with the *cycle-metric labeling problem* where a given metric matrix is defined by an undirected cycle and non-negative edge length. Therefore, our results form an important class of the metric labeling problem, which has a polynomial time approximation algorithm with a constant approximation ratio.

Our main results are Theorem 1 and Theorem 2. Theorem 1 shows that a cycle-metric matrix can be approximated by a convex combination of Monge matrices. In the case where the given cost matrix is a Monge matrix, the problems can be solvable in polynomial time. Theorem 1 gives a basic idea of our algorithm. Theorem 2 shows that our algorithm is a $2(1-1/h)$-approximation algorithm for the problem. The rest of this paper is structured as follows. In Section 2, we describe the problem formulation. In Section 3, we describe the Monge property and our dependent rounding procedure. In Section 4, we propose the approximation algorithm and analyze the approximation ratio of our algorithm. In Section 5, we discuss the better approximation algorithm under some assumptions of triangle inequalities. In Section 6, we give the conclusion.

2 Problem Formulation

Let $H = \{1, 2, \ldots, h\}$ be a set of hub nodes and N be a set of non-hub nodes where $|H| \geq 3$ and $|N| = n$. This paper deals with a single assignment hub network design problem, which assigns each non-hub node to exactly one hub node. We discuss the case in which the set of hubs forms an undirected cycle, and the corresponding problem is called the *cycle-star hub network design problem*. More precisely, we are given an undirected cycle $\Gamma = (H, T)$ defined by a vertex-set H and an edge-set $T = \{\{1, 2\}, \{2, 3\}, \ldots, \{h-1, h\}, \{h, 1\}\}$. In the rest of this paper, we identify hub i and hub $h+i$ when there is no ambiguity. For each edge $e = \{i, i+1\} \in T$, the corresponding length, which is denoted by c_e or $c_{i\,i+1}$, represents a non-negative cost per unit flow on the edge. For each ordered pair $(p, i) \in N \times H$, c_{pi} also denotes a non-negative cost per unit flow on an undirected edge $\{p, i\}$. We denote a given non-negative amount of flow from a non-hub p to another non-hub q by $w_{pq} (\geq 0)$. Throughout this paper, we assume that $w_{pp} = 0$ $(\forall p \in N)$. We discuss the problem for finding an assignment of non-hubs to hubs, which minimizes the total transportation cost defined below.

When non-hub nodes p and q $(p \neq q)$ are assigned hubs i and j, respectively, an amount of flow w_{pq} is sent along a path $((p, i), \Omega_{ij}, (j, q))$, where Ω_{ij} denotes the shortest path in $\Gamma = (H, T)$ between i and j. For each pair of hub nodes $(i, j) \in H^2$, c_{ij} denotes the length of the shortest path Ω_{ij}. More precisely, cycle Γ contains exactly two paths between i and j, and c_{ij} denotes the minimum of the lengths of these two paths. It is easy to see that $c_{ij} = c_{ji}$. In the rest of this paper, the matrix $C = (c_{ij})$ defined above is called a *cost matrix* and/or a *cycle-metric matrix*. The transportation cost corresponding to a flow from p to q is defined by $w_{pq}(c_{pi} + c_{ij} + c_{qj})$.

Now we describe our problem formally. First, we introduce a 0-1 variable

x_{pi} for each pair $\{p, i\} \in N \times H$ as follows:

$$x_{pi} = \begin{cases} 1 & (p \in N \text{ is assigned to } i \in H), \\ 0 & (\text{otherwise}). \end{cases}$$

We have a constraint $\sum_{i \in H} x_{pi} = 1$ for each $p \in N$, since each non-hub is connected to exactly one hub. Then, the cycle-star hub network design problem can be formulated as follows:

$$\begin{array}{lll} \text{SAP:} & \text{min.} & \displaystyle\sum_{(p,q) \in N^2} w_{pq} \left(\sum_{i \in H} c_{pi} x_{pi} + \sum_{j \in H} c_{jq} x_{qj} + \sum_{(i,j) \in H^2} c_{ij} x_{pi} x_{qj} \right) \\ & \text{s. t.} & \displaystyle\sum_{i \in H} x_{pi} = 1 \qquad\qquad (\forall p \in N), \\ & & x_{pi} \in \{0, 1\} \qquad\qquad (\forall (p, i) \in N \times H). \end{array}$$

The above formulation also appears in [30, 21]. In case $h = 3$, 3-cycle Γ is a complete graph, and therefore the corresponding problem is NP-complete [30].

Next, we describe the integer linear programming problem proposed in [21], which is derived from SAP by employing the linearization technique introduced by Adams and Sherali [1]. We replace $x_{pi}x_{qj}$ with y_{piqj}. We have a new constraint $\sum_{i \in H} y_{piqj} = x_{qj}$ from the equation $\sum_{i \in H} x_{pi} = 1$ by multiplying both sides by x_{qj}. We also obtain a constraint $\sum_{j \in H} y_{piqj} = x_{pi}$ in a similar way. Then, we obtain the following 0-1 integer linear programming problem:

$$\begin{array}{lll} \text{SAPL:} & \text{min.} & \displaystyle\sum_{(p,q) \in N^2} w_{pq} \left(\sum_{i \in H} c_{pi} x_{pi} + \sum_{j \in H} c_{jq} x_{qj} + \sum_{(i,j) \in H^2} c_{ij} y_{piqj} \right) \\ & \text{s. t.} & \displaystyle\sum_{i \in H} x_{pi} = 1 \qquad\qquad (\forall p \in N), \\ & & \displaystyle\sum_{j \in H} y_{piqj} = x_{pi} \qquad\qquad (\forall (p, q) \in N^2, \forall i \in H, p < q), \\ & & \displaystyle\sum_{i \in H} y_{piqj} = x_{qj} \qquad\qquad (\forall (p, q) \in N^2, \forall j \in H, p < q), \\ & & x_{pi} \in \{0, 1\} \qquad\qquad (\forall (p, i) \in N \times H), \\ & & y_{piqj} \in \{0, 1\} \qquad\qquad (\forall (p, q) \in N^2, \forall (i, j) \in H^2). \end{array}$$

By substituting the non-negativity constraints of all the variables for 0-1 constraints in SAPL, we obtain a linear relaxation problem, denoted by LRP. We can solve LRP in polynomial time by employing an interior point algorithm.

3 Monge Property and Dependent Rounding Procedure

First, we give the definition of a Monge matrix. A comprehensive research on the Monge property appears in a recent survey [7].

Definition 1 *An $m \times n$ matrix C' is a Monge matrix if and only if C' satisfies the so-called Monge property*

$$c'_{ij} + c'_{i'j'} \le c'_{ij'} + c'_{i'j} \qquad \text{for all} \quad 1 \le i < i' \le m, 1 \le j < j' \le n.$$

Although the Monge property depends on the orders of the rows and columns, in this paper, we say that a matrix is a Monge matrix when there exist permutations of rows and columns that yield the Monge property.

For each edge $e \in T$, we define a path $\Gamma^e = (H, T \setminus \{e\})$ obtained from cycle Γ by deleting the edge e. Let $C^e = (c^e_{ij})$ be a cost matrix, where c^e_{ij} denotes the length of the unique subpath of Γ^e connecting i and j.

Lemma 1 *For any edge $e = \{\ell, \ell+1\} \in T$, a Monge matrix is obtained from C^e by permuting the rows and columns simultaneously in the ordering $(\ell+1, \ell+2, \ldots, h, 1, 2, \ldots, \ell)$.*

Proof is omitted (see Appendix A or [7] for example).

Next, we approximate a given cost matrix (cycle-metric matrix) C by a convex combination of h Monge matrices $\{C^e \mid e \in T\}$. Alon et al. [2] considered approximating a cycle-metric matrix by a probability distribution over path-metric matrices, and showed a simple distribution such that the expected length of each edge is no more than twice its original length. The following theorem improves upon their result, especially when the size of the cycle (number of hubs) is small.

Theorem 1 *Let C be a cost matrix obtained from a cycle $\Gamma = (H, T)$ and non-negative edge lengths $(c_e \mid e \in T)$. Then, there exists a vector of coefficients $(\theta_e \mid e \in T)$ satisfying*

$$\theta_e \ge 0 \ (\forall e \in T), \ \sum_{e \in T} \theta_e = 1, \ \ \text{and} \ \ C \le \sum_{e \in T} \theta_e C^e \le 2\left(1 - \frac{1}{h}\right) C.$$

Proof: When there exists an edge $e^\circ \in T$ satisfying $c_{e^\circ} \ge (1/2) \sum_{f \in T} c_f$, it is easy to see that for every pair $(i, j) \in H^2$, there exists a shortest path Ω_{ij} on cycle $\Gamma = (H, T)$ between i and j excluding edge e°. Thus, a given cost matrix C is equivalent to the Monge matrix C^{e°. In this case, the desired result is trivial.

We assume that $2c_e < L = \sum_{f \in T} c_f$ $(\forall e \in T)$ and introduce a positive coefficient θ_e for each $e \in T$ defined by

$$\theta_e = \frac{c_e}{K} \prod_{f \in T \setminus \{e\}} (L - 2c_f)$$

where K is a normalizing constant, which yields the equality $\sum_{e\in T}\theta_e = 1$. Let $\Omega_{ij} \subseteq T$ be a set of edges in the shortest path in Γ between i and j. The definition of the coefficients $(\theta_e \mid e \in T)$ directly implies that for each pair $(i,j) \in H^2$,

$$\begin{aligned}\sum_{e\in T}\theta_e c^e_{ij} &= \sum_{e\notin\Omega_{ij}}\theta_e c_{ij} + \sum_{e\in\Omega_{ij}}\theta_e(L-c_{ij}) = \sum_{e\in T}\theta_e c_{ij} + \sum_{e\in\Omega_{ij}}\theta_e(L-2c_{ij})\\ &\le c_{ij}\sum_{e\in T}\theta_e + \sum_{e\in\Omega_{ij}}\theta_e(L-2c_e)\\ &= c_{ij} + \sum_{e\in\Omega_{ij}}\left((L-2c_e)\frac{c_e}{K}\prod_{f\in T\setminus\{e\}}(L-2c_f)\right)\\ &= c_{ij} + \frac{\prod_{f\in T}(L-2c_f)}{K}\sum_{e\in\Omega_{ij}}c_e = c_{ij} + \frac{\prod_{f\in T}(L-2c_f)}{K}c_{ij}.\end{aligned}$$

From the assumption, the last term appearing above is positive. Then, we have

$$\frac{K}{\prod_{f\in T}(L-2c_f)} = \frac{\sum_{e\in T}\left(c_e\prod_{f\in T\setminus\{e\}}(L-2c_f)\right)}{\prod_{f\in T}(L-2c_f)} = \sum_{e\in T}\frac{c_e}{L-2c_e}.$$

Now we introduce a function $f(z_1,\ldots,z_h) = \sum_{\ell=1}^{h} z_\ell/(L-2z_\ell)$ defined on a domain $\{\boldsymbol{z} \in [0,L/2)^h \mid z_1+\cdots+z_h = L\}$. From the convexity and symmetry of variables of f, the minimum of f is attained at $z_1 = z_2 = \cdots = z_h = L/h$, and $f(L/h,\ldots,L/h) = 1/(1-2/h)$, which gives the following inequality

$$\sum_{e\in T}\theta_e c^e_{ij} \le c_{ij} + \frac{\prod_{f\in T}(L-2c_f)}{K}c_{ij} \le c_{ij} + \left(1-\frac{2}{h}\right)c_{ij} = 2\left(1-\frac{1}{h}\right)c_{ij}.$$

Since $C \le C^e$ $(\forall e \in T)$, it is obvious that $C \le \sum_{e\in T}\theta_e C^e$. □

Next, we describe the rounding technique proposed in [21]. We will describe the connection between the Monge matrix and the rounding technique later.

`Dependent Rounding` $(\boldsymbol{x},\boldsymbol{y};\pi)$

Input: A feasible solution $(\boldsymbol{x},\boldsymbol{y})$ of LRP and a total order π of the hubs.

Step 1: Generate a random variable U which that a uniform distribution defined on $[0,1)$.

Step 2: Assign each non-hub node $p \in N$ to a hub $\pi(i)$, where $i \in \{1,2,\ldots h\}$ is the minimum number that satisfies $U < x_{p\pi(1)} + \cdots + x_{p\pi(i)}$.

The above procedure can be explained roughly as follows (see Figure 2). For each non-hub $p \in N$, we subdivide a rectangle of height 1 with horizontal segments into smaller rectangles whose heights are equal to the values of the given feasible solution $x_{p\pi(1)}, x_{p\pi(2)}, \ldots, x_{p\pi(h)}$. Here, we note that $\sum_{i \in H} x_{p\pi(i)} = 1$ $(\forall p \in N)$. We assume that the smaller rectangles are heaped in the order π. We generate a horizontal line whose height is equal to the random variable U and round a variable x_{pi} to 1 if and only if the corresponding rectangle intersects the horizontal line.

Given a feasible solution $(\boldsymbol{x}, \boldsymbol{y})$ of LRP and a total order π of H, the vector of random variables X^π, which is indexed by $N \times H$, denotes the solution obtained by `Dependent Rounding` $(\boldsymbol{x}, \boldsymbol{y}; \pi)$. In the following, we discuss the probability $\Pr[X^\pi_{pi} X^\pi_{qj} = 1]$.

Lemma 2 [21] *Let* $(\boldsymbol{x}, \boldsymbol{y})$ *be a feasible solution of LRP and* π *be the total order of* H*. The vector of random variables* $\boldsymbol{X}^\pi$ *obtained by* `Dependent Rounding` $(\boldsymbol{x}, \boldsymbol{y}; \pi)$ *satisfies*

(1) $\mathrm{E}[X^\pi_{pi}] = x_{pi}$ $\quad (\forall (p, i) \in N \times H)$,

(2) $\mathrm{E}[X^\pi_{pi} X^\pi_{qj}] = y^\pi_{piqj}$ $\quad (\forall (p, q) \in N^2, \forall (i, j) \in H^2)$,

where $\boldsymbol{y}^\pi$ *is a unique solution of the following system of equalities*

$$\sum_{i=1}^{i'} \sum_{j=1}^{j'} y^\pi_{p\pi(i)q\pi(j)} = \min \left\{ \sum_{i=1}^{i'} x_{p\pi(i)}, \sum_{j=1}^{j'} x_{q\pi(j)} \right\} \left(\forall (p, q) \in N^2, \forall (i', j') \in H^2 \right). \tag{1}$$

Proof is omitted (see Appendix C or [21]).

In the rest of this paper, the pair of vectors $(\boldsymbol{x}, \boldsymbol{y}^\pi)$ defined by (1) is called a *north-west corner rule solution* with respect to $(\boldsymbol{x}, \boldsymbol{y}; \pi)$. When $\boldsymbol{x}$ is non-negative and $\sum_{i \in H} x_{pi} = \sum_{i \in H} x_{qi}$ holds, the unique solution of (1) gives the so-called *north-west corner rule solution* for a Hitchcock transportation problem (Figure 3 shows an example of a north-west corner rule solution, whose details are given in Appendix B or [21]). Here we note that the above definition is different from the ordinary definition of the north-west corner rule solution,

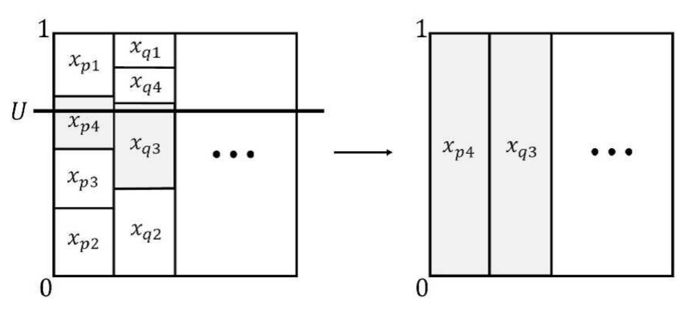

Figure 2: Dependent Rounding π_ℓ where $(\pi(1), \pi(2), \pi(3), \pi(4)) = (2, 3, 4, 1)$.

which is the result of the procedure for finding a feasible solution of a Hitchcock transportation problem. In the rest of this section, we describe the Hitchcock transportation (sub)problems contained in LRP.

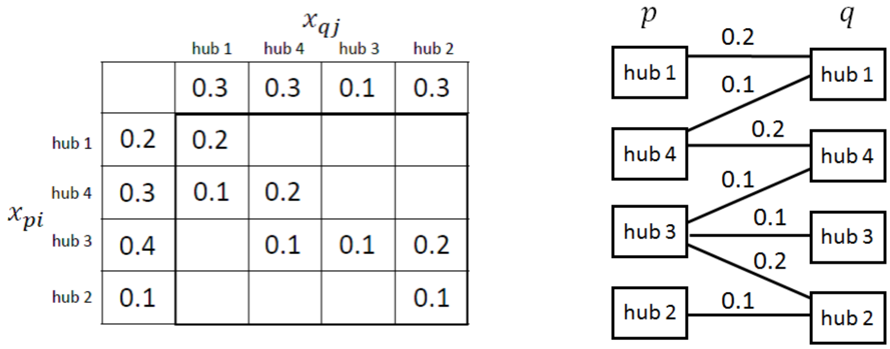

Figure 3: North-west corner rule solution where $(\pi(1), \pi(2), \pi(3), \pi(4)) = (2, 3, 4, 1)$. In this case, $\mathrm{E}[X^{\pi}_{p1} X^{\pi}_{q1}] = 0.2$, $\mathrm{E}[X^{\pi}_{p4} X^{\pi}_{q1}] = 0.1$, $\mathrm{E}[X^{\pi}_{p4} X^{\pi}_{q4}] = 0.2$, $\mathrm{E}[X^{\pi}_{p3} X^{\pi}_{q4}] = 0.1$, $\mathrm{E}[X^{\pi}_{p3} X^{\pi}_{q3}] = 0.1$, $\mathrm{E}[X^{\pi}_{p3} X^{\pi}_{q2}] = 0.2$, and $\mathrm{E}[X^{\pi}_{p2} X^{\pi}_{q2}] = 0.1$.

Let $(\boldsymbol{x}^{\circ},\ \boldsymbol{y}^{\circ})$ be a feasible solution of LRP. For any $p \in N$, $\boldsymbol{x}^{\circ}_p$ denotes a subvector of $\boldsymbol{x}^{\circ}$ defined by $(x^{\circ}_{p1}, x^{\circ}_{p2}, \ldots, x^{\circ}_{ph})$. When we fix the variables $\boldsymbol{x}$ in LRP to $\boldsymbol{x}^{\circ}$, we can decompose the obtained problem into n^2 Hitchcock transportation problems $\{\mathrm{HTP}(\boldsymbol{x}^{\circ}_p, \boldsymbol{x}^{\circ}_q, C) \mid (p, q) \in N^2\}$, where

$$
\begin{array}{llll}
\mathrm{HTP}(\boldsymbol{x}^{\circ}_p, \boldsymbol{x}^{\circ}_q, C): & \text{min.} & \displaystyle\sum_{i \in H} \sum_{j \in H} c_{ij} y_{piqj} & \\
& \text{s. t.} & \displaystyle\sum_{j \in H} y_{piqj} = x^{\circ}_{pi} & (\forall i \in H), \\
& & \displaystyle\sum_{i \in H} y_{piqj} = x^{\circ}_{qj} & (\forall j \in H), \\
& & y_{piqj} \geq 0 & (\forall (i, j) \in H^2).
\end{array}
$$

Next, we describe a well-known relation between the north-west corner rule solution of a Hitchcock transportation problem and the Monge property.

Lemma 3 *If a given cost matrix $C = (c_{ij})$ is a Monge matrix with respect to the total order π of hubs, then the north-west corner rule solution $\boldsymbol{y}^{\pi}$ defined by* (1) *gives optimal solutions of all the Hitchcock transportation problems $\{HTP(\boldsymbol{x}^{\circ}_p, \boldsymbol{x}^{\circ}_q, C) \mid (p, q) \in N^2\}$.*

Proof is omitted here (see [6, 7] for example).

4 Approximation Algorithm

In this section, we propose an algorithm and discuss its approximation ratio. First, we describe our algorithm. and then we discuss the approximation ratio.

Alogorithm 4

Step 1: Solve the linear relaxation problem LRP and obtain an optimal solution $(\boldsymbol{x}^*, \boldsymbol{y}^*)$.

Step 2: For each edge $e \in T$, execute `Dependent Rounding` $(\boldsymbol{x}^*, \boldsymbol{y}^*; \pi^e)$, where π^e denotes the total order $(\pi^e(1), \pi^e(2), \ldots, \pi^e(h)) = (\ell+1, \ell+2, \ldots, h, 1, 2, \ldots, \ell-1, \ell)$.

Step 3: Output the best solution obtained in Step 2.

For simplicity, we denote the objective function by

$$W_1^* = \sum_{(p,q)\in N^2} w_{pq} \left(\sum_{i\in H} c_{pi} x_{pi}^* + \sum_{j\in H} c_{jq} x_{qj}^* \right) \text{ and}$$

$$W_2^* = \sum_{(p,q)\in N^2} w_{pq} \left(\sum_{(i,j)\in H^2} c_{ij} y_{piqj}^* \right)$$

where $(\boldsymbol{x}^*, \boldsymbol{y}^*)$ is an optimal solution of LRP.

Theorem 2 *Algorithm 4 is a* $2(1-1/h)$*-approximation algorithm for cycle-star hub network design problems and cycle-metric labeling problems with* h *hubs(labels).*

Proof: Let z^{**} be the optimal value of the original problem SAP and $(\theta_e \mid e \in T)$ be a vector of coefficients defined in Theorem 1. For each $e \in T$, $(X^{\pi(e)})$ denotes a solution obtained by `Dependent Rounding`$(\boldsymbol{x}^*, \boldsymbol{y}^*; \pi^e)$, and $\boldsymbol{y}^{\pi(e)}$ is the north-west corner rule solution defined by (1) (where π is set to π^e). Then, we have that

$$
\begin{aligned}
&2\left(1-\frac{1}{h}\right)z^{**} \geq 2\left(1-\frac{1}{h}\right)(\text{optimal value of LRP}) = 2\left(1-\frac{1}{h}\right)(W_1^* + W_2^*) \\
&\geq W_1^* + 2\left(1-\frac{1}{h}\right)W_2^* = W_1^* + \sum_{(p,q)\in N^2} w_{pq}\left(\sum_{(i,j)\in H^2} 2\left(1-\frac{1}{h}\right)c_{ij}y^*_{piqj}\right) \\
&\geq \sum_{e\in T}\theta_e W_1^* + \sum_{(p,q)\in N^2} w_{pq}\left(\sum_{(i,j)\in H^2}\sum_{e\in T}\theta_e c^e_{ij}y^*_{piqj}\right) \quad \text{(Theorem 1)} \\
&= \sum_{e\in T}\theta_e W_1^* + \sum_{e\in T}\theta_e\left(\sum_{(p,q)\in N^2} w_{pq}\left(\sum_{(i,j)\in H^2} c^e_{ij}y^*_{piqj}\right)\right) \\
&\geq \sum_{e\in T}\theta_e W_1^* + \sum_{e\in T}\theta_e\left(\sum_{(p,q)\in N^2} w_{pq}\left(\text{optimal value of HTP}(\boldsymbol{x}^*_p, \boldsymbol{x}^*_q, C^e)\right)\right) \\
&= \sum_{e\in T}\theta_e W_1^* + \sum_{e\in T}\theta_e\left(\sum_{(p,q)\in N^2} w_{pq}\left(\sum_{(i,j)\in H^2} c^e_{ij}y^{\pi(e)}_{piqj}\right)\right) \quad \text{(Lemma 3)} \\
&= \sum_{e\in T}\theta_e\left(\sum_{(p,q)\in N^2} w_{pq}\left(\sum_{i\in H} c_{pi}x^*_{pi} + \sum_{j\in H} c_{jq}x^*_{qj} + \sum_{(i,j)\in H^2} c_{ij}y^{\pi(e)}_{piqj}\right)\right) \\
&= \sum_{e\in T}\theta_e\left(\sum_{(p,q)\in N^2} w_{pq}\left(\begin{array}{c}\sum_{i\in H} c_{pi}\mathrm{E}[X^{\pi(e)}_{pi}] + \sum_{j\in H} c_{jq}\mathrm{E}[X^{\pi(e)}_{qj}] \\ +\sum_{(i,j)\in H^2} c_{ij}\mathrm{E}[X^{\pi(e)}_{pi}X^{\pi(e)}_{qj}]\end{array}\right)\right) \\
&= \mathrm{E}\left[\sum_{e\in T}\theta_e\left(\sum_{(p,q)\in N^2} w_{pq}\left(\begin{array}{c}\sum_{i\in H} c_{pi}X^{\pi(e)}_{pi} + \sum_{j\in H} c_{jq}X^{\pi(e)}_{qj} \\ +\sum_{(i,j)\in H^2} c_{ij}X^{\pi(e)}_{pi}X^{\pi(e)}_{qj}\end{array}\right)\right)\right] \\
&\geq \mathrm{E}\left[\min_{e\in T}\left\{\sum_{(p,q)\in N^2} w_{pq}\left(\begin{array}{c}\sum_{i\in H} c_{pi}X^{\pi(e)}_{pi} + \sum_{j\in H} c_{jq}X^{\pi(e)}_{qj} \\ +\sum_{(i,j)\in H^2} c_{ij}X^{\pi(e)}_{pi}X^{\pi(e)}_{qj}\end{array}\right)\right\}\right] \\
&= \mathrm{E}[Z]
\end{aligned}
$$

where Z denotes the objective value of the solution obtained by Algorithm 4. The last inequality in the above transformation is obtained from the equality $\sum_{e\in T}\theta_e = 1$ and the non-negativity of coefficients $(\theta_e \mid e \in T)$. □

5 A Better Analysis with Some Assumption

Lastly, we discuss a simple independent rounding technique which independently connects each non-hub node $p \in N$ to a hub node $i \in H$ with probability x^*_{pi}, where $(\boldsymbol{x}^*, \boldsymbol{y}^*)$ is an optimal solution of LRP. Iwasa et al. [21] discussed a case where the set of hubs forms a complete graph. They showed that the independent rounding technique gives a 2-approximation algorithm under the following assumption.

Assumption 1 *A given symmetric non-negative cost matrix C satisfies $c_{ij} \leq c_{ik} + c_{kj}$ $(\forall(i,j,k) \in H^3)$ and $c_{ij} \leq c_{pi} + c_{pj}$ $(\forall(i,j,p) \in H^2 \times N)$.*

The following theorem implies that if we add the assumption that C is a cycle-metric matrix to Assumption 1, we can achieve a better approximation ratio by combining the independent rounding technique and our proposed method.

Theorem 3 *Under the assumptions that C is a cycle-metric matrix and $c_{ij} \leq c_{pi} + c_{pj}$ $(\forall(i,j,p) \in H^2 \times N)$, a $\left(\frac{3}{2} - \frac{1}{2(h-1)}\right)$-approximation algorithm (for a cycle-star hub network design problem) is obtained by choosing the better of the two solutions given by Algorithm 4 and the independent rounding technique.*

Proof: Let a random variable Z_2 be an objective function value with respect to a solution obtained by the independent rounding technique. Iwasa et al. [21] showed that $\mathrm{E}[Z_2] \leq 2W_1^* + W_2^*$. In the proof of Theorem 2, we have shown that $\mathrm{E}[Z] \leq W_1^* + 2(1 - 1/h)W_2^*$. By combining these results, we obtain that

$$\begin{aligned}
&\mathrm{E}[\min\{Z_2, Z\}] \leq \mathrm{E}\left[\frac{h-2}{2(h-1)} Z_2 + \frac{h}{2(h-1)} \mathrm{E}[Z]\right] \\
&= \frac{h-2}{2(h-1)} \mathrm{E}[Z_2] + \frac{h}{2(h-1)} \mathrm{E}[Z] \\
&\leq \frac{h-2}{2(h-1)} (2W_1^* + W_2^*) + \frac{h}{2(h-1)} \left(W_1^* + 2\left(1 - \frac{1}{h}\right) W_2^*\right) \\
&= \left(\frac{3}{2} - \frac{1}{2(h-1)}\right)(W_1^* + W_2^*) \leq \left(\frac{3}{2} - \frac{1}{2(h-1)}\right) (\text{opt. val. of SAP}).
\end{aligned}$$

□

6 Conclusions

In this paper, we proposed a polynomial time $2(1 - 1/h)$-approximation algorithm for a cycle-star hub network design problem with h hubs. Our algorithm solves the linear relaxation problem and employs a dependent rounding procedure. The attained approximation ratio is based on an approximation of a cycle-metric matrix by a convex combination of Monge matrices. Moreover, for the case where the given metric satisfies the triangle inequalities, we proposed the $\frac{3}{2} - \frac{1}{2(h-1)}$-approximation algorithm by combining the existing independent rounding based algorithm and our rounding based algorithm.

Appendix A Proof of Lemma 1

Here, we briefly provide the a proof of Lemma 1.

Proof: Let C' be a matrix obtained from C^e by permuting the rows and columns simultaneously in the ordering $(\ell+1, \ell+2, \ldots, h, 1, 2, \ldots, \ell)$. We consider each quadruple (i, i', j, j') of indices satisfying $1 \leq i < i' \leq m, 1 \leq j < j' \leq n$. Since C^e is symmetric, it suffices to show the following three cases:

(i) if $i \leq i' \leq j \leq j'$, then $c'_{ij} + c'_{i'j'} = c'_{ii'} + 2c'_{i'j} + c'_{jj'} = (c'_{ij'} + c'_{i'j})$,

(ii) if $i \leq j \leq i' \leq j'$, then $c'_{ij} + c'_{i'j'} \leq c'_{ij} + c'_{i'j'} + c'_{ji'} + c'_{i'j} = (c'_{ij'} + c'_{i'j})$,

(iii) if $i \leq j \leq j' \leq i'$, then $c'_{ij} + c'_{i'j'} \leq c'_{ij} + c'_{j'i'} + 2c'_{jj'} = c'_{ij'} + c'_{ji'} = (c'_{ij'} + c'_{i'j})$.

Thus, we have the desired result.

□

Appendix B North-West Corner Rule

A Hitchcock transportation problem is defined on a complete bipartite graph consisting of a set of supply points $A = \{1, 2, \ldots, I\}$ and a set of demand points $B = \{1, 2, \ldots, J\}$. Given a pair of non-negative vectors $(\boldsymbol{a}, \boldsymbol{b}) \in \mathrm{R}^I \times \mathrm{R}^J$ satisfying $\sum_{i=1}^{I} a_i = \sum_{j=1}^{J} b_j$ and an $I \times J$ cost matrix $C' = (c'_{ij})$, a Hitchcock transportation problem is formulated as follows:

$$
\begin{array}{lll}
\mathrm{HTP}(\boldsymbol{a}, \boldsymbol{b}, C'): & \text{min.} & \displaystyle\sum_{i=1}^{I}\sum_{j=1}^{J} c'_{ij} y_{ij} \\
& \text{s. t.} & \displaystyle\sum_{j=1}^{J} y_{ij} = a_i \qquad (i \in \{1, 2, \ldots, I\}), \\
& & \displaystyle\sum_{i=1}^{I} y_{ij} = b_j \qquad (j \in \{1, 2, \ldots, J\}), \\
& & y_{ij} \geq 0 \qquad (\forall (i, j) \in \{1, 2, \ldots, I\} \times \{1, 2, \ldots, J\}),
\end{array}
$$

where y_{ij} denotes the amount of flow from a supply point $i \in A$ to a demand point $j \in B$.

We describe the north-west corner rule in Algorithm NWCR, which finds a feasible solution of the Hitchcock transportation problem $\mathrm{HTP}(\boldsymbol{a}, \boldsymbol{b}, C')$. It is easy to see that the north-west corner rule solution $Y = (y_{ij})$ satisfies the following equalities:

$$
\sum_{i=1}^{i'}\sum_{j=1}^{j'} y_{ij} = \min\left\{\sum_{i=1}^{i'} \alpha_i \ , \ \sum_{j=1}^{j'} \beta_j\right\} \quad (\forall (i', j') \in \{1, 2, \ldots, I\} \times \{1, 2, \ldots, J\}).
$$

Algorithm NWCR

Step 1: Set all the elements of matrix Y to 0 and set the target element y_{ij} to y_{11} (top-left corner).

Step 2: Allocate the maximum possible amount of transshipment to the target element without making the row or column total of the matrix Y exceed the supply or demand respectively.

Step 3: If the target element is y_{IJ} (the south-east corner element), then stop.

Step 4: Denote the target element by y_{ij}. If the sum total of the jth column of Y is equal to b_j, set the target element to y_{ij+1}. Else (the sum total of Y of ith row is equal to a_i), set the target element to y_{i+1j}. Go to Step 2.

Since the coefficient matrix of the above equality system is nonsingular, the north-west corner rule solution is a unique solution of the above equality system. Thus, the above system of equalities has a unique solution which is feasible to HTP$(\boldsymbol{a}, \boldsymbol{b}, C')$.

Appendix C Proof of Lemma 2

Here, we provide the proof of Lemma 2 briefly. See [21] for detail.

Proof: (1) When we introduce an index i' satisfying $\pi(i') = i$, it is easy to see that

$$\mathrm{E}[X^{\pi}_{pi}] = \Pr[X^{\pi}_{pi} = 1] = \Pr\left[\sum_{j=1}^{i'-1} x_{p\pi(j)} \leq U < \sum_{j=1}^{i'} x_{p\pi(j)}\right] = x_{p\pi(i')} = x_{pi}.$$

(2) We denote $\mathrm{E}[X^{\pi}_{pi}X^{\pi}_{qj}] = \Pr[X^{\pi}_{pi}X^{\pi}_{qj} = 1]$ by y'_{piqj} for simplicity. Then, for any pairs $(p, q) \in N^2$ and $(i', j') \in H^2$, the vector $\boldsymbol{y}'$ satisfies

$$\begin{aligned}
&\sum_{i=1}^{i'}\sum_{j=1}^{j'} y'_{p\pi(i)q\pi(j)} = \Pr\left[\left[\sum_{i=1}^{i'} X_{p\pi(i)} = 1\right] \wedge \left[\sum_{j=1}^{j'} X_{q\pi(j)} = 1\right]\right] \\
&= \Pr\left[\left[U < \sum_{i=1}^{i'} x_{p\pi(i)}\right] \wedge \left[U < \sum_{j=1}^{j'} x_{q\pi(j)}\right]\right] \\
&= \Pr\left[U < \min\left\{\sum_{i=1}^{i'} x_{p\pi(i)}, \sum_{j=1}^{j'} x_{q\pi(j)}\right\}\right] \\
&= \min\left\{\sum_{i=1}^{i'} x_{p\pi(i)}, \sum_{j=1}^{j'} x_{q\pi(j)}\right\}.
\end{aligned}$$

From the above, $(\boldsymbol{x}, \boldsymbol{y}')$ satisfies system (1) defined by $(\boldsymbol{x}, \boldsymbol{y}; \pi)$. The non-singularity of (1) implies $\boldsymbol{y}' = \boldsymbol{y}^{\pi}$. □

Acknowledgements

This work was done while the first author was at Tokyo Institute of Technology. This work was supported by JSPS KAKENHI Grant Numbers 26285045 and 26242027.

References

[1] M. P. Adams and H. D. Sherali. A tight linearization and an algorithm for zero-one quadratic programming problems. *Management Science*, 32:1274–1290, 1986. doi:10.1287/mnsc.32.10.1274?journalCode=mnsc.

[2] N. Alon, R. M. Karp, D. Peleg, and D. West. A graph-theoretic game and its application to the k-server problem. *SIAM Journal on Computing*, 24:78–100, 1995. doi:10.1137/S0097539792224474.

[3] S. A. Alumur and B. Y. Kara. Network hub location problems: The state of the art. *European Journal of Operational Research*, 190:1–21, 2008. doi:10.1016/j.ejor.2007.06.008.

[4] S. A. Alumur and B. Y. Kara. Network hub location problems: The state of the art. *European Journal of Operational Research*, 190:1–21, 2008. doi:10.1016/j.ejor.2007.06.008.

[5] R. Ando and T. Matsui. Algorithm for single allocation problem on hub-and-spoke networks in 2-dimensional plane. *Lecture Notes in Computer Science*, 7074:474–483, 2011. doi:10.1007/978-3-642-25591-5_49.

[6] W. Bein, P. Brucker, K. Park, and K. Pathak. A monge property for the d-dimensional transportation problem. *Discrete Applied Mathematics*, 58:97–109, 1995. doi:10.1016/0166-218X(93)E0121-E.

[7] R. E. Burkard, B. Klinz, and R. Rudolf. Perspectives of monge properties in optimization. *Discrete Applied Mathematics*, 70:95–161, 1996. doi:10.1016/0166-218X(95)00103-X.

[8] H. Calik, S. A. Alumur, B. Y. Kara, and O. E. Karasan. A tabu-search based heuristic for the hub covering problem over incomplete hubnetworks. *Computers and Operations Research*, 36:3088–3096, 2009. doi:10.1016/j.cor.2008.11.023.

[9] J. F. Campbell. A survey of network hub location. *Studies in Locational Analysis*, 6:31–47, 1994. doi:10.1016/j.cie.2013.01.012.

[10] J. F. Campbell, A. T. Ernst, and M. Krishnamoorthy. Hub arc location problems: Part i: introduction and results. *Management Science*, 51:1540–1555, 2005. doi:10.1287/mnsc.1050.0406.

[11] J. F. Campbell, A. T. Ernst, and M. Krishnamoorthy. Hub arc location problems: Part ii: formulations and optimal algorithms. *Management Science*, 51:1556–1571, 2005. doi:10.1287/mnsc.1050.0407.

[12] J. F. Campbell and M. E. O'Kelly. Twenty-five years of hub location research. *Transportation Science*, 46:153–169, 2012. doi:10.1287/trsc.1120.0410.

[13] J. Chuzhoy and J. Naor. The hardness of metric labeling. *SIAM Journal on Computing*, 36:1376–1386, 2007. doi:10.1137/S0097539703422479.

[14] I. Contreras. Hub location problems. *in: Location Science, Chapter 12, G. Laporte, S. Nickel, and F. S. da Gama (Eds), Springer*, pages 311–344, 2015.

[15] I. Contreras, E. Fernández, and A. Marín. Tight bounds from a path based formulation for the tree of hub location problem. *Computers and Operations Research*, 36:3117–3127, 2009. doi:10.1016/j.cor.2008.12.009.

[16] I. Contreras, E. Fernández, and A. Marín. The tree of hubs location problem. *European Journal of Operational Research*, 202:390–400, 2010. doi:10.1016/j.ejor.2009.05.044.

[17] I. Contreras, M. Tanash, and N. Vidyarthi. Exact and heuristic approaches for the cycle hub location problem. *Annals of Operations Research*, pages 1–23, 2016. doi:10.1007/s10479-015-2091-2.

[18] E. M. de Sá, I. Contreras, J. F. Cordeau, R. de Camargo, and G. de Miranda. The hub line location problem. *Transportation Science*, 49:500–518, 2015. doi:10.1287/trsc.2014.0576.

[19] E. M. de Sá, R. S. de Camargo, and G. de Miranda. An improved benders decomposition algorithm for the tree of hubs location problem. *European Journal of Operational Research*, 226:185–202, 2013. doi:10.1016/j.ejor.2012.10.051.

[20] J. Fakcharoenphol, S. Rao, and K. Talwar. A tight bound on approximating arbitrary metrics by tree metrics. *Journal of Computer and System Sciences*, 60:485–497, 2004. doi:10.1016/j.jcss.2004.04.011.

[21] M. Iwasa, H. Saito, and T. Matsui. Approximation algorithms for the single allocation problem in hub-and-spoke networks and related metric labeling problems. *Discrete Applied Mathematics*, 157:2078–2088, 2009. doi:10.1016/j.dam.2008.11.016.

[22] J. G. Kim and D. W. Tcha. Optimal design of a two-level hierarchical network with tree-star configuration. *Computers and Industrial Engineering*, 22:273–281, 1992. doi:10.1016/0360-8352(92)90005-5.

[23] J. Kleinberg and E. Tardos. Approximation algorithms for classification problems with pairwise relationships. *Journal of ACM*, 49:616–639, 2002. doi:10.1145/585265.585268.

[24] J. G. Klincewicz. Hub location in backbone/tributary network design: a review. *Location Science*, 6:307–335, 1998. doi:10.1016/S0966-8349(98)00042-4.

[25] M. Labbé and H. Yaman. Solving the hub location problem in a star-star network. *Networks*, 51:19–33, 2008. doi:10.1002/net.20193.

[26] M. E. O'Kelly. A quadratic integer program for the location of interacting hub facilities. *European Journal of Operational Research*, 32:393–404, 1987. doi:10.1016/S0377-2217(87)80007-3.

[27] M. E. O'Kelly and H. J. Miller. The hub network design problem: A review and synthesis. *Journal of Transport Geography*, 2:31–40, 1994. doi:10.1016/0966-6923(94)90032-9.

[28] H. Saito, T. Fujie, T. Matsui, and S. Matuura. A study of the quadratic semi-assignment polytope. *Discrete Optimization*, 6:37–50, 2009. doi:10.1016/j.disopt.2008.08.003.

[29] J. Sohn and S. Park. A linear program for the two-hub location problem. *European Journal of Operational Research*, 100:617–622, 1997. doi:10.1016/S0377-2217(96)00233-0.

[30] J. Sohn and S. Park. The single allocation problem in the interacting three-hub network. *Networks*, 35:17–25, 2000. doi:10.1002/(SICI)1097-0037(200001)35:1<17::AID-NET2>3.0.CO;2-N.

[31] H. Yaman. Star p-hub median problem with modular arc capacities. *Computers and Operations Research*, 35:3009–3019, 2008. doi:10.1016/j.cor.2007.01.014.

[32] H. Yaman and S. Elloumi. Star p-hub center problem and star p-hub median problem with bounded path lengths. *Computers and Operations Research*, 39:2725–2732, 2012. doi:10.1016/j.cor.2012.02.005.