ICS 2011
12th INFORMS Computing Society Conference

 | ISBN 978-0-9843378-1-1
DOI 10.1287/ics.2011.0007

Special Case Studies of the Stochastic p-Hub Center Single Allocation Problem with Service Constraints

Li Zhang
Department of Mathematics and Computer Science, The Citadel,
Charleston, South Carolina 29409, li.zhang@citadel.edu

Abstract In time-sensitive distribution systems, one crucial factor to consider is how to keep the delivery time between any origin and destination pair within a time guarantee. With variability in travel or delivery, it is important to maintain the services at a certain level such that the probability of on-time deliveries for all customers is high. In this paper, we address the stochastic p-hub center single allocation problem with service constraints and the assumption that the standard deviation of the travel time between any origin and destination pair is proportional to the mean travel time between them. We propose a mixed integer formulation for the problem and discuss several polynomial solvable cases to the problem.

Keywords hubs; time guarantee; complexity; algorithm

1. Introduction

Hubs are often used in networks such as air passenger travel, postal services, express shipments, and distributed computer systems. These systems usually serve a large number of customers or end users. For example, UPS delivered 3.8 billion packages and documents in 2009 (15.1 million daily) and had a revenue of \$37.9 billion in 2009 (UPS Pressroom [18]). In these networks, direct connections linking all pairs of origin-destination (o-d) nodes can be extremely expensive and impractical. Hubs are special facilities usually used to serve as consolidation, switching and sorting centers, thus direct connections are replaced by fewer, indirect connections. The hub-and-spoke network applications have significantly reduced the start-up, operating and maintenance costs. In UPS' case, six hubs were used to serve 388 airports domestically in 2009 (UPS Pressroom [18]).

In the literature, hub location problems focus on determining the locations of the hubs and allocations of the nonhub nodes to the hub nodes. There are two objectives that have been considered the most: minimizing the sum of the total transportation cost (time, distance, etc.) and minimizing the maximum transportation cost between any o–d pair. If the number of hubs to be located is equal to a predetermined number p, the problem is called the p-hub median problem with the first objective (i.e., find the locations of the p hubs and allocate the nonhub nodes to the p hubs such that the total cost is minimized), and the problem is called the p-hub center problem with the second objective (i.e., find the locations of the p hubs and allocate the nonhub nodes to the p hubs such that the maximum travel cost between any o-d pair is minimized). There are several variants to both problems depending on how the nonhub nodes are allocated to the hub nodes. Single allocation requires a nonhub node to be assigned to one and only one hub node, and multiple allocation requires a nonhub node to be assigned to one or more hub nodes. More details about the hub location problems can be found in the survey papers (Alumur and Kara [1], Bryan and O'Kelly [2], Campbell [5], Klincewicz [10]).

In a hub-and-spoke network application, the p-hub center (single or multiple allocation) problem determines the locations of the p hub nodes and allocations of the nonhub nodes to the p hub nodes such that the maximum cost (distance, time, etc.) between any o-d pair is minimized. This problem is important for a system that delivers time-sensitive items or services, such as express mail services at UPS or FedEx and emergency services at hospitals or fire stations. In these systems, the maximum travel time between any o-d pair represents the best time guarantee that can be offered to all customers, and this value must be kept as low as possible. Like the p-hub median problem, the p-hub center problem in general is NP-hard (both problems remain NP-hard even if hub locations are fixed) Campbell [5], Ernst et al. [7], Kara and Tansel [9], Sohn and Park [14]. Only in special cases such as $p = 2$ or on special graphs, these problems can be solved polynomially (Campbell et al. [4], Sohn and Park [13]).

Since the actual transportation or delivery time from an origin to a destination is often uncertain (instead of a constant as assumed in the p-hub location problems), it would be more applicable to real world situations if we consider this factor in our model. If the transportation time between any o-d pair is a random variable, it is possible that a plane may not arrive on time or a package may not be delivered on time. To reduce the number of failures of on-time delivery and the compensation amount due to these failures, companies may insist on achieving a minimum service level. While the p-hub center problem received some limited research attention in literature, there is only one published article on the stochastic p-hub center problem by Sim et al. [12]. In their paper, the authors present a mixed-integer formulation for the stochastic p-hub center single allocation problem with service constraints and propose three heuristic methods to solve the problem. They also provide the results from their computational experiments.

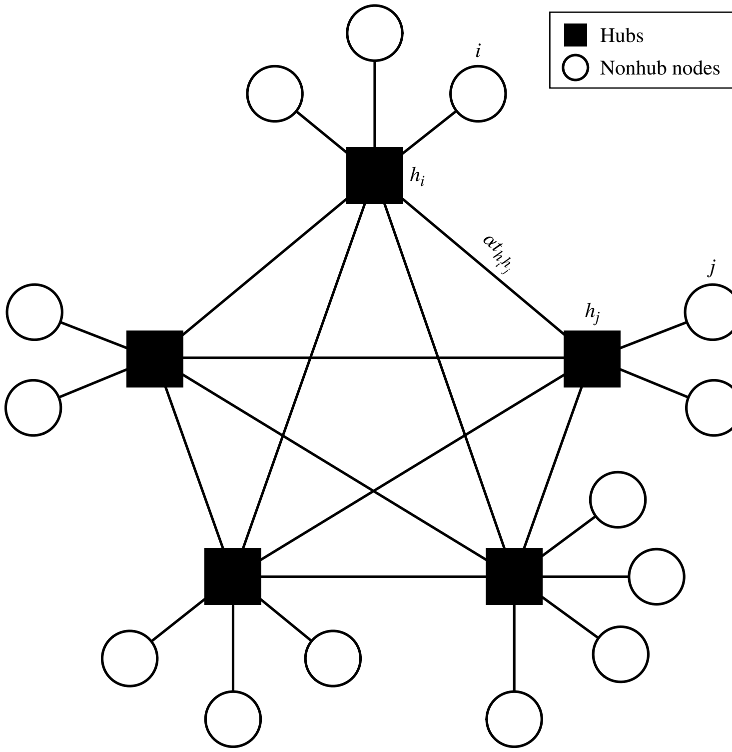
In this paper, we focus on a special case of the stochastic p-hub center single allocation problem with service constraints, where the standard deviation of the travel time between any pair of o-d nodes is assumed to be proportional to the mean travel time between these two nodes. In §2, we provide problem descriptions including assumptions made for the problem and propose a mixed-integer formulation for the model. In §3, we analyze our model and discuss several polynomial solvable cases. In §4, we summarize our results and explore future research directions.

2. Problem Description and Model Formulation

Suppose a given complete network is described by the complete graph $G = (N, E)$ with node set $N = \{1, \ldots, n\}$ and undirected arc set E. The stochastic p-hub center problem with service constraints is to locate the p hubs on G and allocate the nonhub nodes to the p hubs such that the maximum travel time between any o-d pair is minimized for a given service level γ. In practice, γ is usually chosen to be close to 1 such as 0.90 or 0.95. We will only consider the single allocations of nonhub nodes to hub nodes in this paper. Since the flow between every pair of hubs uses a rapid mode of transit as a result of economies of scale, the travel time between every pair of hubs will be less than their corresponding travel time when the rapid mode of transit is not used (i.e., the given travel time on the graph). This can be modeled by a discount factor α (where $0 \le \alpha \le 1$) which is a multiplier that reduces the given travel time between hubs. See Figure 1 for a graphical illustration for the problem, where each nonhub node is assigned to one hub and the mean travel time between a pair of hubs is multiplied by α and the hub network is completely connected.

Similar to the assumptions made in the paper by Sim et al. [12], we assume that the travel time T_{ij} on arc (i, j) from i to j is a normally distributed random variable with mean t_{ij}

FIGURE 1. A graphical illustration of the problem.



and variance δ_{ij}^2, and the T_{ij} values are independent random variables. Assume $t_{ij} = t_{ji}$, $t_{ii} = 0$, and $t_{ij} > 0$ if $i \neq j$, $\forall i, j \in N$. Also, in a solution network (where the nonhub nodes are allocated to the hub nodes), we require that the unique path between any o-d pair with the smallest number of hubs must be followed. Since the hubs are fully connected (i.e., they form a complete network), this requirement implies that the number of hubs on any o-d path is either one (if both of the o-d nodes are assigned to the same hub) or two (if the o-d nodes are assigned to two different hubs). There is no need to make such assumption in the p-hub center or median problems since the travel time values are assumed to satisfy the triangular inequalities and the hubs are completely connected in a solution network, thus it always costs less (or no more) on a path having at most two hubs than the path having three or more hubs. In the stochastic p-hub center single allocation problem with service constraints, we do not assume that the T_{ij} values satisfy the triangular inequalities and in fact the triangular inequalities are often violated due to random travel times being independent. Thus, it is possible that a path having more than two hubs cost less than the path having the same o-d pair and two hubs. See an example about this in Figure 2 in §4.

We use $\widehat{T}_{ij}$ to denote the travel time between an o-d pair i and j on a resulting network G' in a solution. Note that there are no direct connections between any pair of nonhub nodes on G' since each nonhub node has to be assigned to a hub node in a solution. If nodes i and j are nonhub nodes in a solution, and i is assigned to hub k and j is assigned to hub l, then the path from i to j should be the path $i \to k \to l \to j$ (from the requirement that an o-d path should have the smallest number of hubs in a solution network), and $\widehat{T}_{ij} = T_{ik} + \alpha T_{kl} + T_{lj}$. With the assumption that $T_{ij} \sim N(t_{ij}, \delta_{ij}^2)$, $\forall i, j \in N$, and T_{ij} values are independent, we have $\widehat{T}_{ij} = T_{ik} + \alpha T_{kl} + T_{lj} \sim N(t_{ik} + \alpha t_{kl} + t_{lj}, \delta_{ik}^2 + \alpha^2 \delta_{kl}^2 + \delta_{lj}^2)$, i.e., $\widehat{T}_{ij}$ is a random variable having normal distribution with mean $t_{ik} + \alpha t_{kl} + t_{lj}$ and variance $\delta_{ik}^2 + \alpha^2 \delta_{kl}^2 + \delta_{lj}^2$.

To solve the stochastic p-hub center single allocation problem with service constraints, Sim et al. [12] introduced a mixed-integer formulation having $O(n^4)$ constraints and $O(n^4)$

binary variables, and they claimed that CPLEX was only able to solve problems with $n < 10$. The formulation is apparently not very useful in practice. In this paper, we make a further assumption for the problem and assume δ_{ij} is proportional to the mean t_{ij}, i.e., $\delta_{ij} = \lambda t_{ij}$, $\forall i, j \in N$ and λ is some nonnegative constant. This assumption is reasonable since the standard deviation of the travel time from i to j is often bigger if the mean travel time from i to j is longer (more unknown factors are involved during a longer period of time). Note Sim et al. [12] did use this assumption in their numerical experiment for data testing, but not in their model formulation. Also, some other examples of using the idea that assumes the standard deviation or variance to be proportional to the mean to reduce the size of the formulation can be found in Cai and Zhou [3], Jang and Klein [8], Sheikh et al. [11], Song and Miller [15], Spoerl and Wood [16], Spurrell [17]. Under this assumption, we are able to formulate the problem with $O(n^2)$ constraints and $O(n^2)$ binary variables. Note that if $\delta_{ij} = \lambda t_{ij}$, then $\widehat{T}_{ij} = T_{ik} + \alpha T_{kl} + T_{lj} \sim N(t_{ik} + \alpha t_{kl} + t_{lj}, \delta_{ik}^2 + \alpha^2 \delta_{kl}^2 + \delta_{lj}^2) = N(t_{ik} + \alpha t_{kl} + t_{lj}, \lambda^2(t_{ik}^2 + \alpha^2 t_{kl}^2 + t_{lj}^2))$.

Define X_{ik} to be a binary variable that equals 1 if node i is allocated to hub k and 0 otherwise. If there is a hub at node k, then $X_{kk} = 1$. Let Z represent the maximum travel time between any o-d pair, and our objective is to minimize the Z value for a given service level γ. That is, the probability of any o-d pair having travel time less than or equal to Z should be greater than or equal to γ, i.e., $P(\widehat{T}_{ij} \le Z) \ge \gamma$, $\forall i, j \in N$. Each of these constraints is called a chance constraint for the travel time on the path from origin i to destination j. If $X_{ik} = 1$ and $X_{jl} = 1$, then $\widehat{T}_{ij} = T_{ik} + \alpha T_{kl} + T_{lj} \sim N(t_{ik} + \alpha t_{kl} + t_{lj}, \lambda^2(t_{ik}^2 + \alpha^2 t_{kl}^2 + t_{lj}^2))$. Now $P(\widehat{T}_{ij} \le Z) \ge \gamma$ (the chance constraint on the path $i \to k \to l \to j$) can be expressed as

$$P\left(\frac{\widehat{T}_{ij} - (t_{ik} + \alpha t_{kl} + t_{lj})}{\lambda\sqrt{(t_{ik}^2 + \alpha^2 t_{kl}^2 + t_{lj}^2)}} \le \frac{Z - (t_{ik} + \alpha t_{kl} + t_{lj})}{\lambda\sqrt{(t_{ik}^2 + \alpha^2 t_{kl}^2 + t_{lj}^2)}}\right) \ge \gamma,$$

i.e., $(Z - (t_{ik} + \alpha t_{kl} + t_{lj}))/\lambda\sqrt{(t_{ik}^2 + \alpha^2 t_{kl}^2 + t_{lj}^2)} \ge Z_\gamma$, where Z_γ is the γ-level quantile in the standard normal distribution such that $P(Z_s \le Z_\gamma) = \gamma$ (where Z_s represents the standard normally distributed random variable). Since γ is close to 1, $Z_\gamma \ge 0$. We can rewrite this chance constraint as $Z \ge t_{ik} + \alpha t_{kl} + t_{lj} + Z_\gamma \lambda \sqrt{t_{ik}^2 + \alpha^2 t_{kl}^2 + t_{lj}^2}$. If $X_{ik} = 0$ or $X_{jl} = 0$, then the path $i \to k \to l \to j$ does not exist in the solution network and $T_{ik} + \alpha T_{kl} + T_{lj}$ should be multiplied by 0. Instead of using the four-indexed binary variable Y_{iklj} to represent the existence of the path as in Sim et al. [12] which can increase the complexity of the formulation dramatically, we use radius r_k of a hub k in our formulation. The radius concept was first proposed by Ernst et al. [7] in their study of the p-hub center problems. Let $r_k = \max\{t_{ik} X_{ik}, \forall i \in N\}$, thus r_k represents the maximum mean travel time between hub k and the nonhub nodes that are assigned to it. Every o-d path in a solution network must contain at least one hub, and among all the o-d pairs where all origin nodes are assigned to hub k and all destination nodes are assigned to hub l (k and l are not necessarily different), the maximum value of $t_{ik} + \alpha t_{kl} + t_{lj} + Z_\gamma \lambda \sqrt{t_{ik}^2 + \alpha^2 t_{kl}^2 + t_{lj}^2}$ is $r_k + r_l + \alpha t_{kl} + Z_\gamma \lambda \sqrt{r_k^2 + \alpha^2 t_{kl}^2 + r_l^2}$. Thus, we can express the chance constraints on Z as in (2). Also, let $M = \max\{t_{ij}, \forall i, j \in N\}$. Note that M is a constant for a given network and $r_k \le M$, $\forall k \in N$. The formulation for the stochastic p-hub center single allocation problem with service constraints and assumption that $\delta_{ij} = \lambda t_{ij}$, $\forall i, j \in N$ is as follows:

$$\min \quad Z \tag{1}$$

$$\text{subject to} \quad Z \ge r_k + r_l + \alpha t_{kl} + Z_\gamma \lambda \sqrt{r_k^2 + \alpha^2 t_{kl}^2 + r_l^2}, \quad \forall k \le l \in N, \tag{2}$$

$$r_k \ge t_{ik} X_{ik}, \quad \forall i, k \in N, \tag{3}$$

$$r_k \le M X_{kk}, \quad \forall k \in N, \tag{4}$$

$$\sum_{k\in N} X_{ik} = 1, \quad \forall\, i \in N, \tag{5}$$

$$X_{ik} \le X_{kk}, \quad \forall\, i,k \in N, \tag{6}$$

$$\sum_{k\in N} X_{kk} = p, \tag{7}$$

$$X_{ik} \in \{0,1\}, \quad \forall\, i,k \in N. \tag{8}$$

The objective of minimizing the maximum travel time between any o-d pair satisfying a given service level is expressed in (1) and (2), where constraints in (2) are the chance constraints on Z as explained earlier. Notice that the right-hand side in constraint (2) is for $\forall\, k \le l \in N$ since $t_{kl} = t_{lk}$. Constraints (3) and (4) ensure that r_k is the radius of k if k is a hub and r_k is 0 if k is a nonhub node (Note that since the objective is to minimize Z, constraint (4) may be removed from the formulation to further reduce the size of the model. However, if constraint (4) is removed, it is possible that a non-zero r_k for a nonhub node k in an optimal solution. Thus, constraint (4) can be used to tighten up the formulation). Constraint (5) ensures that each node is assigned to one and only one hub node. Constraint (6) ensures that node i can be assigned to node k if k is a hub node and no nodes can be assigned to k if k is a nonhub node. Constraint (7) ensures that exactly p hubs are located. This formulation has $\frac{5}{2}n^2 + \frac{5}{2}n + 1$ constraints and $n^2 + n + 1$ variables of which n^2 are binary. If the condition $\delta_{ij} = \lambda t_{ij}, \forall\, i,j \in N$ is satisfied, this formulation is able to solve a problem of much larger size than the formulation proposed by Sim et al. [12]. Note that Sim et al. [12] also proposed a reduce-sized model in the case that a triangle inequality holds in a certain stochastic dominance sense.

3. Polynomial Solvable Cases

Since the p-hub center single allocation problem is NP-hard in general (Campbell [5], Ernst et al. [7], Kara and Tansel [9]), the problem we discuss in this paper is also NP-hard in general. In fact, the p-hub center single allocation problem can be considered as a special case of our problem where $t_{ij} = d_{ij}$ (distance between i and j satisfying triangular inequality) and $\delta_{ij} = 0$ (or $\lambda = 0$), $\forall\, i,j \in N$. In this section we discuss several polynomial solvable cases for the problem formulated in §2. Recall that the chance constraint for the travel time on the path $i \to k \to l \to j$ is $Z \ge t_{ik} + \alpha t_{kl} + t_{lj} + Z_\gamma \lambda \sqrt{t_{ik}^2 + \alpha^2 t_{kl}^2 + t_{lj}^2}$. Let $\widetilde{T}_{ij}$ equal the right-hand side of the constraint (i.e., $\widetilde{T}_{ij} = t_{ik} + \alpha t_{kl} + t_{lj} + Z_\gamma \lambda \sqrt{t_{ik}^2 + \alpha^2 t_{kl}^2 + t_{lj}^2}$) for a given allocation of the nonhub nodes to the p hub nodes where $X_{ik} = X_{jl} = 1$. Notice that $\widetilde{T}_{ij} = \widetilde{T}_{ji}$ since $t_{ij} = t_{ji}, \forall\, i,j \in N$. Thus the problem is equivalent to locating the p hubs and allocating the nonhub nodes to the p hubs such that $Z = \max\{\widetilde{T}_{ij}, \forall\, i,j \in N\}$ is minimized.

Theorem 1. *Suppose $t_{ik} \ge t_{i'k}$ and $X_{ik} = X_{i'k} = X_{kk} = 1$, then $\widetilde{T}_{ij} \ge \widetilde{T}_{i'j}$ and $\widetilde{T}_{ji} \ge \widetilde{T}_{ji'}, \forall\, j \in N$.*

Proof. Node j must be assigned to a hub node and suppose j is assigned to a hub node l, i.e., $X_{jl} = X_{ll} = 1$, then $\widetilde{T}_{ij} = t_{ik} + \alpha t_{kl} + t_{lj} + Z_\gamma \lambda \sqrt{t_{ik}^2 + \alpha^2 t_{kl}^2 + t_{lj}^2}$ and $\widetilde{T}_{i'j} = t_{i'k} + \alpha t_{kl} + t_{lj} + Z_\gamma \lambda \sqrt{t_{i'k}^2 + \alpha^2 t_{kl}^2 + t_{lj}^2}$. Since $t_{ik} \ge t_{i'k} \ge 0$, $\sqrt{t_{ik}^2 + \alpha^2 t_{kl}^2 + t_{lj}^2} \ge \sqrt{t_{i'k}^2 + \alpha^2 t_{kl}^2 + t_{lj}^2}$. Also, since $Z_\gamma \ge 0$ and $\lambda \ge 0$, we have $\widetilde{T}_{ij} \ge \widetilde{T}_{i'j}$. Since $\widetilde{T}_{ij} = \widetilde{T}_{ji}$ and $\widetilde{T}_{i'j} = \widetilde{T}_{ji'}$, we also have $\widetilde{T}_{ji} \ge \widetilde{T}_{ji'}$. □

Corollary 1. *Suppose node k is a hub and set $K = \{\forall\, i \in N \mid X_{ik} = 1\}$. Then $\max\{\widetilde{T}_{ij}, \forall\, i,j \in K\} = \widetilde{T}_{i^*i^*}$ for some $i^* \in K$ where $t_{i^*k} \ge t_{ik}, \forall\, i \in K$. Furthermore, $\widetilde{T}_{i^*i^*} = \widetilde{T}_{i^*j^*}$ for some $j^* \in K$ iff $t_{j^*k} = t_{i^*k}$.*

Proof. Since $t_{i^*k} \geq t_{ik}, \forall i \in K$, $\widetilde{T}_{i^*i^*} \geq \widetilde{T}_{ii^*} \geq \widetilde{T}_{ij}, \forall j \in K$ by Theorem 1. Thus, $\max\{\widetilde{T}_{ij}, \forall i,j \in K\} = \widetilde{T}_{i^*i^*}$.

Assume $t_{j^*k} = t_{i^*k}$. We have $\widetilde{T}_{i^*j^*} = t_{i^*k} + t_{j^*k} + Z_\gamma\lambda\sqrt{t^2_{i^*k} + t^2_{j^*k}} = t_{i^*k} + t_{i^*k} + Z_\gamma\lambda\sqrt{t^2_{i^*k} + t^2_{i^*k}} = \widetilde{T}_{i^*i^*}$.

Assume $\widetilde{T}_{i^*i^*} = \widetilde{T}_{i^*j^*}$. If $i^* = j^*$, then $t_{j^*k} = t_{i^*k}$. Suppose $i^* \neq j^*$. If $t_{i^*k} > t_{j^*k}$, then $t^2_{i^*k} > t^2_{j^*k}$ since $t_{i^*k} > t_{j^*k} \geq 0$. Also, $\widetilde{T}_{i^*i^*} = t_{i^*k} + t_{i^*k} + Z_\gamma\lambda\sqrt{t^2_{i^*k} + t^2_{i^*k}} > t_{i^*k} + t_{j^*k} + Z_\gamma\lambda\sqrt{t^2_{i^*k} + t^2_{i^*k}} > t_{i^*k} + t_{j^*k} + Z_\gamma\lambda\sqrt{t^2_{i^*k} + t^2_{j^*k}} = \widetilde{T}_{i^*j^*}$, contrary to the assumption that $\widetilde{T}_{i^*i^*} = \widetilde{T}_{i^*j^*}$. Similarly, $t_{i^*k} \not< t_{j^*k}$. Thus, $t_{j^*k} = t_{i^*k}$. □

An allocation is said to be a *feasible* allocation for a given value T if $\widetilde{T}_{ij} \leq T$, $\forall i,j \in N$.

Corollary 2. *If there exists a feasible allocation for $T = \widetilde{T}_{i^*j^*}$ where both i^* and j^* are assigned to the same hub k, then $t_{i^*k} = t_{j^*k}$ and $T = \widetilde{T}_{i^*i^*} = \widetilde{T}_{j^*j^*}$.*

Proof. Let set $K = \{\forall i \in N \mid X_{ik} = 1\}$. Since $\widetilde{T}_{ij} \leq T, \forall i,j \in N$ and $T = \widetilde{T}_{i^*j^*}$ for some i^* and j^* in K, $T = \max\{\widetilde{T}_{ij}, \forall i,j \in K\}$. Also, $t_{i^*k} \geq t_{ik}, \forall i \in K$ (otherwise, if there is an $i \in K$ such that $t_{ik} > t_{i^*k}$, then $\widetilde{T}_{ij^*} > \widetilde{T}_{i^*j^*} = T$, contrary to the assumption that $\widetilde{T}_{ij} \leq T$, $\forall i,j \in N$). By Corollary 1, $T = \widetilde{T}_{i^*i^*}$, and $T = \widetilde{T}_{i^*j^*}$ implies that $t_{i^*k} = t_{j^*k}$, therefore $T = \widetilde{T}_{i^*i^*} = \widetilde{T}_{j^*j^*}$. □

Corollary 2 implies that if $T = \widetilde{T}_{i^*j^*}$ for some i^* and j^* in K and $t_{i^*k} \neq t_{j^*k}$, then there does not exist a feasible allocation for T.

3.1. Case $p = 1$

If k is the one hub node in a solution and $t_{i^*k} \geq t_{ik}$, $\forall i \in N$, then by Corollary 1, $Z = \widetilde{T}_{i^*i^*} = 2t_{i^*k} + \sqrt{2}\lambda Z_\gamma t_{i^*k} = t_{i^*k}(2 + \sqrt{2}\lambda Z_\gamma)$ for that solution. To minimize Z is equivalent to minimize t_{i^*k}. Thus, the problem becomes minimizing maximum t_{ik}, $\forall i,k \in N$ and it is similar to the vertex 1-center problem Daskin [6]. This can be done in at most $O(n^2)$ time as shown in Algorithm 1.

Algorithm 1 (Case $p = 1$).

Let $Z = \infty$ and hub $= 0$;
for $(k = 1;\ k \leq n;\ k{+}{+})\{$
  let $r_k = 0$;
    for $(i = 1;\ i \leq n;\ i{+}{+})\{$
      if $(t_{ik} > r_k)$
        let $r_k = t_{ik};\}$
  if $(r_k < Z)$
    let $Z = r_k$ and hub $= k;\}$
for $(i = 1;\ i \leq n;\ i{+}{+})\{$
  assign i to hub;}
return $Z = Z(2 + \sqrt{2}\lambda Z_\gamma)$.

3.2. Case $p = 2$

The basic idea here is to find the optimal allocation of the nonhub nodes to a pair of given hub nodes and the corresponding Z value (i.e., solving an allocation problem), then compare this Z value to the Z value of which optimal allocation is found for a different pair of given hub nodes. Since there are $\binom{n}{2} = n(n-1)/2$ possible pairs of hub nodes, the optimal solution to the problem can be found among these $n(n-1)/2$ allocation problems as shown in Algorithm 2. Next we discuss how to solve a specific allocation problem where the locations of the two hubs are given.

Suppose the two hubs are fixed at nodes k and l in an allocation problem. For any pair of o-d nodes i and j in N, there are at most four possible paths from i to j: $i \to k \to l \to j$, $i \to k \to j$, $i \to l \to j$, or $i \to l \to k \to j$. Each path corresponds to a $\widetilde{T}_{ij}$ value for which a feasible allocation may or may not exist. Corollary 1 and Corollary 2 imply that we should only consider the $\widetilde{T}_{ij}$ values for the following paths: $i \to k \to l \to j$ (where $i < j$), $i \to k \to i$, $i \to l \to i$, or $i \to l \to k \to j$ (where $i < j$), for $\forall\, i, j \in N$. There are $O(n^2)$ such $\widetilde{T}_{ij}$ values. The optimal Z value for this allocation problem is among these values. Thus, we first compute these $\widetilde{T}_{ij}$ values and then perform a binary search on them to find Z. For a given $\widetilde{T}_{i^*j^*}$ value, we retain information about the o-d path and assignments of o-d nodes i^* and j^* to the two hub nodes. Also, we check if there exists an allocation for $\forall\, i \in N$ as shown in Algorithm 2 (step 4 to step 18 in the procedure Allocation (k,l)). We will show in Theorem 2 that this allocation is a feasible allocation for the given $\widetilde{T}_{i^*j^*}$ value (i.e., $\widetilde{T}_{ij} \le \widetilde{T}_{i^*j^*}$, $\forall\, i, j \in N$). If such allocation exists, then we search for a smaller value, otherwise we search for a larger one until we find the optimal Z for this allocation problem.

Algorithm 2 (Case $p = 2$).

```
1. let Z = ∞, hub1 = 0 and hub2 = 0;
2.    for (k = 1; k ≤ n; k++){
3.       for (l = k + 1; l ≤ n; l++){
4.          if Allocation (k, l) < Z
5.             let Z = Allocation (k, l), hub1 = k and hub2 = l; }}
6. return Z.
```

Allocation (k,l)

```
1. Compute all candidate O(n^2) T̃_ij values and store them in set S.
2. Sort the values in S in an descending order.
3. Perform a binary search on the values in S {for a given T = T̃_{i*j*}:
4.   if i* is assigned to hub k and j* and i* are the same node
5.      let r_k = t_{i*k};
6.      for (i = 1; i ≤ n; i++){
7.         if i can be assigned to k such that t_ik ≤ r_k, then assign i to hub k;
8.         else if i can be assigned to l such that T̃_{ii*} ≤ T and T̃_ii ≤ T, then assign i to hub l;
9.         else break and go to step 3 for the next (bigger) value in the binary search};
10.     if i* is assigned to hub l and j* and i* are the same node, repeat step 5 to step 9
           and interchange k and l;
11.     if i* is assigned to hub k and j* is assigned to hub l
12.        let r_k = t_{i*k} and r_l = t_{j*l};
13.        if T̃_{i*i*} > T or T̃_{j*j*} > T, go to step 3 for the next (bigger) value in the binary search;
14.        for (i = 1; i ≤ n; i++){
15.           if i can be assigned to k such that t_ik ≤ r_k, then assign i to hub k;
16.           else if i can be assigned to l such that t_il ≤ r_l, then assign i to hub l;
17.           else break and go to step 3 for the next (bigger) value in the binary search};
18.  if i* is assigned to hub l and j* is assigned to hub k, repeat step 12 to step 17
        and interchange k and l;
19.  let Z = T and go to step 3 for the next (smaller) value in the binary search};
20. return the optimal Z value and the corresponding allocation of nonhub nodes
      to the two hub nodes.
```

Theorem 2. *For a given $T = \widetilde{T}_{i^*j^*}$ value, if for $\forall\, i \in N$, i can be assigned to a hub node (k or l) as described in the procedure Allocation (k,l) (step 4 to step 18), then the allocation is a feasible allocation for T, i.e., $\widetilde{T}_{ij} \le T$, $\forall\, i, j \in N$.*

Proof. There are four possible cases regarding the allocation of i^* and j^* to the hub nodes k and l: $X_{i^*k}=1$ (step 4), $X_{i^*l}=1$ (step 10), $X_{i^*k}=X_{j^*l}=1$ (step 11), or $X_{i^*l}=X_{j^*k}=1$ (step 18).

Case 1. $X_{i^*k}=1$: We have $r_k=t_{i^*k}$. For $\forall i\in N$, suppose i can be assigned to k such that $t_{ik}\le r_k$ (step 7) or i can be assigned to l such that $\widetilde{T}_{ii^*}\le T$ and $\widetilde{T}_{ii}\le T$ (step 8). Thus, for $\forall i,j\in N$,

(1) if both i and j are assigned to k, then by Corollary 1, $\widetilde{T}_{ij}\le\widetilde{T}_{i^*i^*}=T$;
(2) if i is assigned to k and j is assigned to l, then by Theorem 1, $\widetilde{T}_{ij}=\widetilde{T}_{ji}\le\widetilde{T}_{ji^*}\le T$;
(3) if i is assigned to l and j is assigned to k, similar arguments as in (2);
(4) if i is assigned to l and j is assigned to l, then $\widetilde{T}_{ij}\le\max\{\widetilde{T}_{ii},\widetilde{T}_{jj}\}\le T$.

Case 2. $X_{i^*l}=1$: use similar arguments as in Case 1.

Case 3. $X_{i^*k}=X_{j^*l}=1$: We have $r_k=t_{i^*k}$, $r_l=t_{j^*l}$, $\widetilde{T}_{i^*i^*}\le T$ and $\widetilde{T}_{j^*j^*}\le T$. For $\forall i\in N$, suppose i can be assigned to k such that $t_{ik}\le r_k$ (step 15) or i can be assigned to l such that $t_{il}\le r_l$ (step 16). Thus, for $\forall i,j\in N$,

(1) if both i and j are assigned to k, then by Corollary 1, $\widetilde{T}_{ij}\le\widetilde{T}_{i^*i^*}\le T$;
(2) if i is assigned to k and j is assigned to l, then by Theorem 1, $\widetilde{T}_{ij}\le\widetilde{T}_{ij^*}\le\widetilde{T}_{i^*j^*}=T$;
(3) if i is assigned to l and j is assigned to k, similar arguments as in (2);
(4) if i is assigned to l and j is assigned to l, then $\widetilde{T}_{ij}\le\widetilde{T}_{j^*j^*}\le T$.

Case 4. $X_{i^*l}=X_{j^*k}=1$: use similar arguments as in Case 3. □

In the procedure Allocation (k,l), step one takes $O(n^2)$ time and step two takes $O(n^2\log n)$ time to sort the $O(n^2)$ values found in step 1. Steps 3 to 19 require $O(n\log n)$ time, thus the total time requirement for procedure Allocation (k,l) is $O(n^2\log n)$, and the total time requirement for Algorithm 2 is $O(n^4\log n)$.

3.3. Case $\alpha=0$ and p is a constant

If p is a constant, then there are $\binom{n}{p}$ possible hub locations. We solve an allocation problem for each given locations of the p hubs.

Theorem 3. *Assume $\alpha=0$ in an allocation problem where the locations of the p hubs are given. The optimal solution can be obtained by assigning all the nonhub nodes to their nearest hubs.*

Proof. Let H be the set consisting of the p given hubs. If $\alpha=0$, then $\widetilde{T}_{ij}=t_{ik}+t_{jl}+Z_\gamma\lambda\sqrt{t_{ik}^2+t_{jl}^2}$ on the o-d path $i\to k\to l\to j$ for $\forall i,j\in N$, $\forall k,l\in H$. If for some k^* and l^* in H such that $t_{ik^*}\le t_{ik}$ and $t_{jl*}\le t_{jl}$, $\forall i,j\in N$, $\forall k,l\in H$, then $t_{ik^*}+t_{jl^*}+Z_\gamma\lambda\sqrt{t_{ik^*}^2+t_{jl^*}^2}\le t_{ik}+t_{jl}+Z_\gamma\lambda\sqrt{t_{ik}^2+t_{jl}^2}$. Thus, $Z=\max\{\widetilde{T}_{ij},\,\forall i,j\in N\}$ is minimized if every nonhub node is assigned to its nearest hub. □

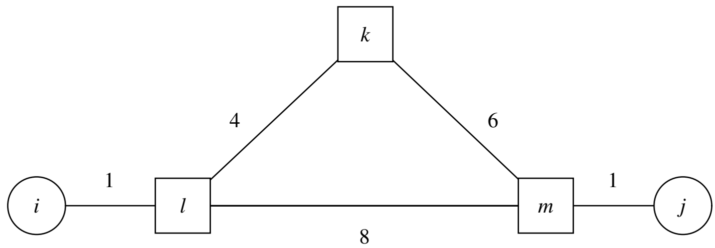
Since enumerating the locations of the p hubs takes $O(n^p)$ time and assigning the nonhub nodes to their nearest hubs takes $O(n)$ time, the problem can be solved in $O(n^{p+1})$ time.

4. Summary

In this paper we discussed the stochastic p-hub center single allocation problem with service constraints and assumption that $\delta_{ij}=\lambda t_{ij}$, $\forall i,j\in N$. We proposed a mixed integer formulation with $\frac{5}{2}n^2+\frac{5}{2}n+1$ constraints and n^2+n+1 variables. If the condition $\delta_{ij}=\lambda t_{ij}$, $\forall i,j\in N$ is satisfied, this formulation can solve a much larger size problem compared to the formulation proposed by Sim et al. [12]. Also, we studied some properties related to the problem and discussed three polynomial solvable cases in §3.

One of the assumptions we made for the problem is that the unique path between any o-d pair with the smallest number of hubs (either one or two) must be followed in a solution

FIGURE 2. An example of the shortest path between an o-d pair.



network. However, the unique path between a pair of o-d nodes may not always cost less than a path that has the same o-d pair and is allowed to have more than two hubs in a solution network. In the example shown in Figure 2, suppose i and j are nonhub nodes that are assigned to hub nodes l and m, respectively. The number on top of each arc is the mean travel time between a pair of nodes. Suppose $\alpha = 0.5$, $Z_\gamma = 2$, and $\lambda = 2$. Clearly the shortest path from i to j is $i \to l \to m \to j$ in the p-hub center single allocation problem. However, in the stochastic p-hub center single allocation problem with service constraints, $\widetilde{T}_{ij} = t_{il} + \alpha t_{lm} + t_{mj} + Z_\gamma \lambda \sqrt{t_{il}^2 + \alpha^2 t_{lm}^2 + t_{mj}^2} = 1 + 0.5 \times 8 + 1 + 2 \times 2 \times \sqrt{1^2 + 0.5^2 \times 8^2 + 1^2} = 22.97$ on the path $i \to l \to m \to j$, and $\widetilde{T}_{ij} = t_{il} + \alpha t_{lk} + \alpha t_{km} + t_{mj} + Z_\gamma \lambda \sqrt{t_{il}^2 + \alpha^2 t_{lk}^2 + \alpha^2 t_{km}^2 + t_{mj}^2} = 1 + 0.5 \times 4 + 0.5 \times 6 + 1 + 2 \times 2 \times \sqrt{1^2 + 0.5^2 \times 4^2 + 0.5^2 \times 6^2 + 1^2} = 22.49$ on the path $i \to l \to k \to m \to j$, which is less than the other one. Thus, we may take this into consideration in our future study of the problem. Other directions in our future study include numerical experiment on the model formulation and heuristic method development.

Acknowledgments

We wish to thank two anonymous referees for their helpful suggestions and the Citadel Foundation for a grant for travel.

References

[1] S. Alumur and B. Kara. Network hub location problems: The state of the art. *European Journal of Operational Research* 190:1–21, 2008.

[2] D. Bryan and M. O'Kelly. Hub-and-spoke networks in air transportation: An analytical review. *Journal of Regional Science* 39:275–295, 1999.

[3] X. Cai and S. Zhou. Scheduling stochastic jobs with asymmetric earliness and tardiness penalties. *Naval Research Logistics* 44:531–557, 1997.

[4] A. Campbell, T. Lowe, and L. Zhang. The p-hub center allocation problem. *European Journal of Operational Research* 176:819–835, 2007.

[5] J. F. Campbell. Integer programming formulations of discrete hub location problems. *European Journal of Operational Research*, 72:387–405, 1994.

[6] M. S. Daskin. *Network and Discrete Location: Models, Algorithms and Applications*. Wiley, New York, 1995.

[7] A. Ernst, H. Hamacher, H. Jiang, M. Krishnamoorthy, and G. Woeginger. Uncapacitated single and multiple allocation p-hub center problems. *Computers & Operations Research* 36(7):2230–2241, 2009.

[8] W. Jang and C. M. Klein. Minimizing the expected number of tardy jobs when processing times are normally distributed. *Operations Research Letters* 30:100–106, 2002.

[9] B. Y. Kara and B. C. Tansel. On the single-assignment p-hub center problem. *European Journal of Operational Research*, 125(3):648–655, 2000.

[10] J. G. Klincewicz. Hub location in backbone/tributary network design: A review. *Location Science* 6:307–335, 1998.

[11] A. K. Sheikh, L. A. Kendall, and S. M. Pandit. Probabilistic optimization of multitool machining operations. *Journal of Engineering for Industry-Transactions of the ASME*, 102(3):239–246, 1980.

[12] T. Sim, T. Lowe, and B. Thomas. The stochastic p-hub center problem with service constraints. *Computers & Operations Research* 36(7):3166–3177, 2009.

[13] J. Sohn and S. Park. A linear program for the two-hub location problem. *European Journal of Operational Research* 100(3):617–622, 1997.

[14] J. Sohn and S. Park. The single allocation problem in the interacting three-hub network. *Networks* 35(1):17–25, 2000.

[15] Y. Song and G. Y. Miller. Effects of marketing decisions on net present value of pork production for independent and allied swine producers effects of marketing decisions on net present value of pork production for independent and allied swine producers. *Review of Agricultural Economics* 24(1):181–195, 2002.

[16] D. R. Spoerl and R. K. Wood. A stochastic generalized assignment problem. Working paper, Naval Postgraduate School, Monterey, CA, 2004.

[17] D. J. Spurrell. A simple method of production planning in multi-item situations. *Operations Research* 18(2):149–159, 1967.

[18] UPS Pressroom. Technical report. http://www.pressroom.ups.com/Fact+Sheets/UPS+Fact +Sheet, Accessed August 18,2010.