The reliable hub-and-spoke design problem: Models and algorithms

Yu An, Yu Zhang*, Bo Zeng

ABSTRACT

Hub-and-spoke structure is widely adopted in industry, especially in transportation and telecommunications applications. Although hub-and-spoke paradigm demonstrates significant advantages in improving network connectivity with less number of routes and saving operating cost, the failure of hubs and reactive disruption management could lead to substantial recovery cost to the operators. Thus, we propose a set of reliable hub-and-spoke network design models, where the selection of backup hubs and alternative routes are taken into consideration to proactively handle hub disruptions. To solve these nonlinear mixed integer formulations for reliable network design problems, Lagrangian relaxation and Branch-and-Bound methods are developed to efficiently obtain optimal solutions. Numerical experiments are conducted with respect to real data to demonstrate algorithm performance and to show that the resulting hub-and-spoke networks are more resilient to hub unavailability.

1. Background and motivation

The hub-and-spoke system has been widely employed in various industrial applications, such as transportation and telecommunications system designs. It is a fully interconnected network with material/information flow between any two nodes being processed at a small number of critical nodes (i.e., hubs) so that the operators can benefit from the economies of scale by consolidating flows from and to spoke nodes and increasing the utilization of equipment and staff at those critical nodes. Clearly, a hub-and-spoke network heavily relies on hubs to make the whole system functional, and therefore it is vulnerable to any disruptions and degradations of hubs. Traditional hub-and-spoke network design solves the problem of hub location and allocations of spoke nodes to hubs, assuming network components work properly. In practice, nevertheless, operators have to face various disruptions and apply disruption management techniques to recover the system. Such an issue is most prominently demonstrated in air transportation where severe weather, labor strikes, terrorism threats, and runway incursions disrupt regular operations and make airports partially or completely unavailable (Palpant et al., 2009; Love and Sörensen, 2001).

To deal with the vulnerability issue of the hub-and-spoke system, several mitigation strategies have been proposed and implemented, such as delaying, canceling, and rerouting in air transportation (Janić, 2005; Ball et al., 2006) and network peering in telecommunications systems (O’Kelly et al., 2006). However, most of mitigation strategies are reactive, which are often costly to implement and inefficient, given that the initial network is designed for perfect conditions. For example,
it is observed in Bratu and Barnhart (2006) that, although the disrupted passengers were only three percent of the total passengers, they suffered 39 percent of the total passenger transportation delays with much lower customer satisfaction. Clearly, the initial network design affects the selections of backup hubs and alternative routes, which affects the cost of mitigation operations. Therefore, to achieve both economic advantage and system reliability, the network design problem should consider both the hub locations and regular route designs as well as the backup hubs and alternative route designs under disruptions in a holistic modeling framework. Therefore, in this paper, we propose a reliable hub-and-spoke network design strategy by explicitly considering the hub unavailability, i.e., backup hub and alternative route decisions will be considered in the design stage and related cost will be included in the objective function of the design problem. With this strategy, we aim to develop a new type of optimization models to minimize the operating cost considering both the normal situation, which is disruption free, and disrupted situations where survived hubs serve as backup hubs for rerouting disrupted flights due to unavailable hubs. As illustrated in Fig. 1, where the solid line denotes a regular route for the flight from Tampa to San Francisco and the dotted line denotes an alternative route using Dallas as a backup hub if the Miami hub is unavailable. This strategy will not only benefit airlines but also other industries who adopted hub-and-spoke distribution paradigm with which they can build and operate their networks with both reliability and economic advantages.

Compared to classical models, the introduction of backup hubs and alternative routes drastically increases the complexity of the network design problem. As the choice of backup hubs and alternative routes depends on the hubs in regular routes, a large number of nonlinear terms are introduced to capture the dependency. As a result, nonlinear mixed integer formulations are constructed. Their structures are further investigated and solution methods developed. To the best of our knowledge, our study is the first analytical work on the reliable hub-and-spoke design with consideration of backup hubs and alternative routes. The developed algorithm is easy to implement and can solve practical instances in a reasonable amount of time. Numerical study demonstrates that our reliable models can serve more passengers under the disruption situations and sensitivity analysis shows that the resulting designs are robust to hub unavailability.

The proposed reliable hub-and-spoke network design also yields a set of useful tools for practitioners, such as airlines, to re-structure their networks or to identify strategic partners to hedge against various disruptions and achieve better performance.

The rest of the paper is organized as follows. In Section 2, literature review on hub-and-spoke design is presented as well as recent research on reliable facility location models. In Section 3, the reliable single allocation hub-and-spoke model is formulated and the solution methods are elaborated. In Section 4, the study is extended to the reliable multiple allocation model. Section 5 demonstrates computational performance of the developed algorithms using the CAB data set from airline operations as the case study and provides comparisons between our reliable hub-and-spoke design models and classical models. In addition, system design and performance with proposed model are analyzed and discussed, including sensitivity
analysis and the demonstration of applying proposed model to a recent airlines merger. Section 6 concludes this paper with some discussions on future research directions.

2. Literature review

The hub-and-spoke design problem is conventionally called hub location problem (HLP), which is concerned with locating hub facilities and allocating spoke nodes to hubs. There are generally two basic structures: single allocation (SA) and multiple allocation (MA). In SA hub-and-spoke model, all outbound/inbound flows of any node must travel directly from/to a specific hub. In MA model, flows of a given node can go directly from/to different hubs. When the number of hubs, denoted by \( p \), is given, the problem is called the \( p \)-hub median problem (HMP). In the remainder of this paper, we use SA-HMP or MA-HMP to denote the corresponding design problem. O’Kelly (1987) proposes the first mathematical formulation for HMP and presented the first quantitative analysis on this type of network structure using the Civil Aeronautics Board (CAB) data set. Since then, as hub-and-spoke structures are of significant theoretical and practical values, a large number of studies have been conducted on developing models with more practical features and on designing efficient algorithms.

We first briefly describe a few important results on formulation and algorithm design. Ernst and Krishnamoorthy (1996, 1998a) formulate SA-HMP and MA-HMP, respectively, based on the idea of “multicommodity flow”. Skorin-Kapov et al. (1996) propose mixed integer formulations for both SA-HMP and MA-HMP that yield tight linear relaxations. As for the customized algorithm development, Branch-and-Bound process and Lagrangian relaxation have been widely used to obtain exact solutions (Ernst and Krishnamoorthy, 1998b and Pirkul and Schilling, 1998). Different from the \( p \)-hub median problem, the hub location problem with fixed costs treats the number of hubs as a decision variable and seeks to minimize the transportation cost and the construction cost where a fixed construction cost is associated with a decision of hub location. O’Kelly (1992) and Campbell (1994) study a few formulations of HLP with fixed costs. There are also extensive literature in search of effective solution algorithms for these problems, see Cunha and Silva (2007), Chen (2007), Cánovas et al. (2007) and Contreras et al. (2011a) for examples. One may refer to Alumur and Kara (2008) and Campbell and O’Kelly (2012) for a comprehensive review of modeling techniques and solution methods of HLP. In the remainder of this paper, unless we explicitly mention, the hub-and-spoke network design problem indicates \( p \)-hub median problem.

Recent studies focused on extending classical SA and MA models by incorporating practical factors, such as hub congestion (Grove and O’Kelly, 1986; Elhedhli and Wu, 2010), hub capacity (Contreras et al., 2012), nonlinear economies of scale (de Camargo et al., 2009), and dynamic/stochastic nature of demand and cost (Contreras et al., 2011b,c).

Nearly all studies on HLP assumed that the chosen hubs would always operate functionally as planned. Nevertheless, in practice, hubs could fail due to different reasons. As the typical cases in air transportation industry, adverse weather often significantly deteriorates the availability of a hub airport and results in huge disruption costs. Similar situations have been observed in facility-and-client based supply chain and logistics systems, where facilities, same as hubs, play the central role and their locations are derived using facility location models. Note that, different from hub-and-spoke design, there is no inter-facility transportation in those systems. To deal with facility disruptions, a facility location model with backup strategy, referred to as the reliable facility location model, was introduced by Snyder and Daskin (2005). Since then, this type of research has received significant attention, including Cui et al. (2010), Li and Ouyang (2010), Lim et al. (2009), Li (2011), An et al. (2014). It is commonly observed that the resulting nonlinear optimization formulations are computationally challenging. Hence, customized algorithms are needed for solving real-size problems, among which Lagrangian relaxation methods and their Branch-and-Bound extensions are the major solution strategies (Snyder and Daskin, 2005; Li and Ouyang, 2010; Cui et al., 2010; Lim et al., 2009; Li, 2011).

In contrast to reliable facility location problems that have attracted the attention of many researchers, up to now, only several recent studies considered reliable hub-and-spoke networks. In Kim and O’Kelly (2009), given that each arc or hub has a reliability (same as availability in this paper), they build SA and MA models to derive an optimal network structure that maximizes the expected network flow, without considering backup hubs and alternative routes. Kim (2008) proposes a \( p \)-hub protection model based on single allocation structure with primary and secondary routes presented. The authors then utilized a heuristic method (tabu search) to solve the real instances with up to 20 nodes. In Zeng et al. (2010), reliable SA and MA models with consideration of hub unavailability and alternative routes have been developed and a heuristic algorithm has been implemented. The authors observe that, different from the reliable facility location model, reliable hub-and-spoke models are much more complicated. Indeed, this type of problems have not been analytically investigated with advanced solution methods to deal with real-size problems. Given that many infrastructure systems, e.g., air transportation and telecommunications systems, adopt hub-and-spoke structures where system reliability is of extremely high importance, we perform an analytical study on reliable hub-and-spoke models and develop efficient algorithms to solve practical instances. A framework of model evaluation under correlated hub disruptions will also be proposed. The study fills in the gap in existing literature and advances the research in reliable hub-and-spoke network design.

3. Reliable single allocation hub-and-spoke model

We are aware that multiple allocation hub-and-spoke model is widely applied instead of single allocation model in air transportation industry. Although we apply aviation case studies later in our study, for the sake of completeness, we study
the formulation and solution algorithms for both reliable single and multiple allocation models in Section 3 and 4. There are two major assumptions in our study. First, we focus on solving the problem with single disruptions. We are aware that under some circumstances, multiple disruptions and even massive disruptions could occur. For example, the volcano ash crisis in Europe in 2010 and 2011 caused the closure of many major airports in that region and Sandy hurricane in 2013 made all three commercial airports in New York area malfunctioned for days. Nevertheless, in airline industry, one carrier often strategically locates its hubs far from each other in its hub-and-spoke network. On one hand, it helps the carrier to fully take advantage of the discounted inter-hub transportation. On the other hand, it can prevent the carrier from being affected by multiple simultaneous hub failures due to the same cause. As an example, during the Grísvötn volcano eruption in Iceland in 2011, two neighboring airports, i.e., Airport Hamburg (HAM) and Airport Bremen (BRE) in Germany (BBC, 2011b; BBC, 2011a), had to be closed. Although HAM and BRE serve as hubs for 19 legacy airlines and low cost carriers, the majority of airlines (17 out of 19) operate just one of the two airports as their hubs (Flylowcostairlines.com, 2012; Mygermancity.com, 2012). Under single disruption and normal condition, our model can provide an optimal routing strategy while multiple disruptions occur, airline companies can take “principle of proximity” and assign disrupted routes to closest functional hubs. We demonstrate, in later section, that the optimized solutions from single disruption model provides better network set-up under multiple disruption scenarios compared to the outcomes from classical model. We will continue tackle the multiple disruption problem in our future study. Second, we assume that for routes going through two hubs, the alternative route is still required to go through the unaffected hub airport. The main reason for doing so is to alleviate the possible impacts of rerouting at tactical operational level to the maintenance scheduling. Certain maintenance requirements have to be followed in airline industry. Type A maintenance check is required every 500–800 flight hours and needs 20–50 man-hours. It can be done overnight at an airport gate or hangar. For other types of checks (B, C, and D), the man-hours needed are much longer and many of them have to be performed at hubs, which usually act the role of maintenance bases. Furthermore, MA structure is adopted in designing alternative routes, regardless of the SA or MA structure used for determining regular routes.

### 3.1. Reliable SA model: definition and formulation

In a single allocation problem, every node is assigned to a single hub and all the inbound and outbound flows of this node are routed through that hub. Let \( \mathbf{N} = \{0, 1, \ldots, |N|\} \) be the set of nodes and \( \mathbf{H} \subseteq \mathbf{N} \) be the set of candidate hub locations for this reliable hub-and-spoke design model, i.e., \( R-SAHP\). We assume that \( \mathbf{H} = \mathbf{N} \) throughout this paper. Then, a node \( i \in \mathbf{N} \), is with \( q_i \in [0, 1] \) to represent its disruption probability. We denote a flow by its source \( i \) and destination \( j \) nodes, i.e., an \( i-j \) flow. A route of \( i-j \) flow can be represented by a 4-tuple \( (i,k,m,j) \), where \( k \) and \( m \) represent the first and the second hubs on the route. Unit transportation cost between a pair of nodes \( i \) and \( j \) is \( c_{ij} \), the cost of transporting one unit flow is \( F_{ikmj} = c_{ik} + \alpha c_{km} + c_{mj} \). A discount factor of economies of scale, \( 0 < \alpha < 1 \), is applied to the inter-hub links. So, for \( i-j \) flow taking the route (\( i,k,m,j \)), the cost of transporting one unit flow is \( F_{ikmj} = c_{ik} + \alpha c_{km} + c_{mj} \).

Decision variables in \( R-SAHP \) include hub location and allocation variable \( Y \), route variable \( X \) and backup hub variables \( U \) and \( V \).

\[
Y_{ik} = \begin{cases} 
1, & \text{if } i \text{ is assigned to hub } k, \\
0, & \text{otherwise}; 
\end{cases}
\]

\[
X_{ikmj} = \begin{cases} 
1, & \text{if } i-j \text{ flow is routed through hubs } k \text{ and } m, \\
0, & \text{otherwise}; 
\end{cases}
\]

\[
U_{jn} = \begin{cases} 
1, & \text{hub } n \text{ is the backup hub for the first hub in the route of } i-j \text{ flow}, \\
0, & \text{otherwise}; 
\end{cases}
\]

\[
V_{jn} = \begin{cases} 
1, & \text{hub } n \text{ is the backup hub for the second hub in the route of } i-j \text{ flow}, \\
0, & \text{otherwise}; 
\end{cases}
\]

In our formulation, \( w_{ij} \) is used to denote the traffic volume between nodes \( i \) and \( j \), which is the transportable flow with both \( i \) and \( j \) functional. In other words, if one of these two nodes fails, there will be no traffic between them. We recognize that for air transportation, in practice \( w_{ij} \) might not be constant given that passengers might migrate from a disrupted airport to another airport nearby to complete their critical travel plans. Nevertheless, modeling such migration requires additional information that varies from airport to airport and causes our models intractable in the designing stage. More important, to keep models general for different hub-and-spoke networks where the migration phenomenon may not occur, we assume that \( w_{ij} \) is a constant. We also adopt a convention in many literature (e.g., O’Kelly and Skorin-Kapov, 1996; Pirkul and Schilling, 1998; Sohn and Park, 1998) that \( w_{ij} = w_{ji} \). Given this symmetric structure, in this study we design the network only considering flow from \( i \) to \( j \) with \( j > i \). Note that this assumption also indicates that the first backup hub for route \( (i,k,m,j) \) will be the second one for route \( (j,m,k,i) \). Next, we present \( R-SAHP \) that generalizes and extends the classical SA hub-and-spoke model developed by Skorin-Kapov et al. (1996).
\[ \begin{align*}
\min & \quad \sum_{i \in N} \sum_{k \in H(i)} \sum_{m \in H} \sum_{j \in W(i)} F_{ikmj} w_j (1 - q_k - q_m^o) X_{ikmj} \\
& \quad + \sum_{i \in N} \sum_{j \in W(i)} \left( \sum_{m \in H(j)} \sum_{k \in H} F_{ikmj} w_j (1 - q_m) X_{ikmj} + \sum_{k \in H(j)} F_{ikmj} (1 - q_k^o) X_{ikmj} + F_{ij} w_j X_{ij} \right) \\
& \quad + \rho \left( \sum_{i \in N} \sum_{k \in H} \sum_{m \in H(k)} \sum_{j \in W(j)} \left( F_{ikmj} w_j q_k X_{ikmj} U_{ijn} + F_{ikmj} w_j q_m X_{ikmj} V_{ijn} \right) + \sum_{i \in N} \sum_{k \in H} \sum_{m \in H} \sum_{j \in W(j)} F_{ikmj} w_j q_k X_{ikmj} U_{ijn} \right) \\
\text{subject to} \\
& \quad \sum_{m \in H} X_{ikmj} = Y_{ik} \quad \forall i, j > i, k \\
& \quad \sum_{k \in H} X_{ikmj} = Y_{jm} \quad \forall i, j > i, m \\
& \quad \sum_{k \in H} Y_{ik} = 1 \quad \forall i \\
& \quad \sum_{k \in H} Y_{kk} = p \\
& \quad U_{ij} + \sum_{m \in H} X_{ikmj} \leq Y_{ik} \quad \forall i, j > i, k \\
& \quad \sum_{k \in H} U_{ik} = 1 - \sum_{m \in H} X_{ikmj} - \sum_{m \in H} X_{jimj} \quad \forall i, j > i \\
& \quad V_{jm} + \sum_{k \in H} X_{ikmj} \leq Y_{mm} \quad \forall i, j > i, m \\
& \quad \sum_{m \in H} V_{jm} = 1 - \sum_{k \in H} X_{skij} - \sum_{k \in H} X_{skij} \quad \forall i, j > i \\
& \quad X_{ikmj} \in \{0, 1\} \quad \forall i, j > i, k, m; \ Y_{ik} \in \{0, 1\} \quad \forall i, k; \ U_{ij}, V_{jm} \in \{0, 1\} \quad \forall i, j > i, k
\end{align*} \]

In the R-SAHMP, the objective function is the expected transportation cost considering both the regular and the disrupted situations. Specifically, the first term represents the regular transportation cost for traffic flows with both source and destination at spoke nodes. The probability of regular transportation, given the assumption that in a hub-and-spoke network at most one hub will fail under disruption situation, is computed as \(1 - q_k - q_m^o\), where \(q_m^o\) is 0 if \(k = m\) and \(q_m\) otherwise. By introducing \(q_k^o\), in this way, one formula can capture both cases, i.e., the route is operable with the probability 1 when hubs are 0 if \(k = m\) and \(q_m\) otherwise. The second term represents the regular transportation cost for traffic flows with source or destination at a hub node. The third term in the objective function is the cost of disruption mitigation by diverting flows to alternative routes, which is penalized by a coefficient \(\rho (\rho > 1)\) to represent the impact of disruption to the overall cost in this transportation network (Welman et al., 2010).

Constraints (2)-(5) are classical constraints for the SA \(p\)-median problem (Skorin-Kapov et al., 1996). Constraints (6) and (8) ensure that regular hubs and backup hubs can only be the nodes chosen to be hubs and the regular hubs and the backup hubs must be different. Constraints (7) and (9) are used to ensure that backup routes are selected for all disrupted flows, except the cases where either the source or the destination node of a flow is a hub.

Existing studies have approved that traditional SA hub-and-spoke model is NP-hard. The proposed R-SAHMP problem can be reduced to the traditional one if all nodes are always reliable, so it is also an NP-hard problem. Not only the entire problem is NP-hard, even when all hubs are fixed, the allocation and routing problem in R-SAHMP is still NP-hard (Sohn and Park, 2000). Nevertheless, once all hubs and spoke node allocations are fixed, the design for regular and alternative routes is polynomially solvable. Note that R-SAHMP is an integer quadratic program as its objective function has multiple terms that involve products of two binary variables. Thus, we used the standard linearization method to convert it into a linear model. We also adopted a recent linearization strategy (Chaovalitwongse et al., 2004; Sherali and Smith, 2007; He et al., 2012) to derive a more compact linear reformulation of R-SAHMP. The linearized formulations of the above two methods and computational results are presented in the Appendix A.2 and Section 5, respectively.
3.2. Lagrangian relaxation and Branch-and-Bound

Existing professional mixed integer programming solvers can be applied to seek solutions of the linearized formulas of R-SAHMP. However, due to the large number of variables and constraints in the model, it takes excessive running times (see computational results presented in Section 5). Hence, a Lagrangian relaxation (LR) algorithm is developed after exploring the structure of R-SAHMP. Compared with other algorithms or commercial solvers, the Lagrangian relaxation algorithm often yields high-quality approximate or optimal solutions with much less computational time (Pirkul and Schilling, 1998; Contreras et al., 2011b). Actually, the proposed Lagrangian relaxation technique is able to directly deal with the nonlinear R-SAHMP without linearizing the formulation. Furthermore, variable fixing and Branch-and-Bound methods are implemented to identify an optimal solution if the Lagrangian relaxation algorithm fails to obtain it.

3.2.1. Lagrangian lower bound

For R-SAHMP, we dualize the constraints (2), (3), (4), (6), and (8) with \( \delta_{ijk,1}, \delta_{ijk,2}, \beta_{i}, \gamma_{ijk,1}, \) and \( \gamma_{ijk,2} \) as their Lagrangian multipliers, respectively. As a result, we obtain the following relaxation:

\[
\min \left\{ \sum_{i \in N} \sum_{k \in H} C_{ik} Y_{ik} - \sum_{i \in N} \sum_{k \in H} \sum_{j \in H} \sum_{m \in H} (F_{ikm} w_{ij}(1 - q_{im}^k) + \delta_{ijk,1} + \gamma_{ijk,1} + \delta_{ijk,2} + \gamma_{ijk,2}) X_{ikm} + \sum_{i \in N} \sum_{j \in H} \sum_{m \in H} (F_{ijm} w_{ij}(1 - q_{ij}^m) + \delta_{ij,1} + \gamma_{ij,1} + \delta_{ij,2} + \gamma_{ij,2}) X_{ijm} + \sum_{i \in N} \sum_{j \in H} \sum_{k \in H} \sum_{m \in H} \rho F_{ikm} w_{ij} q_{ikm} X_{ikm} + \sum_{i \in N} \sum_{j \in H} \sum_{k \in H} \sum_{m \in H} \rho F_{ijm} w_{ij} q_{ijm} X_{ijm} \right\}
\]

subject to

\[
\begin{align*}
(5), (7), (9), \text{ and } (10) \\
Y_{ik} &\leq Y_{kk} \quad \forall i, k \\
\end{align*}
\]

where

\[
\bar{C}_{ik} = \begin{cases} 
\beta_i - \sum_{j \neq i} \delta_{ijk,1} - \sum_{j \neq i} \delta_{ijk,2}, & \text{if } i \neq k; \\
\beta_k - \sum_{j \neq k} \delta_{ijk,1} - \sum_{j \neq k} \delta_{ijk,2} - \sum_{i \in N} \sum_{j \neq i} \gamma_{ijk,1} + \gamma_{ijk,2}, & \text{otherwise.}
\end{cases}
\]

Note that (12) is implied in R-SAHMP and can be derived from (2) and (6).

Since \( X \) and \( Y \) variables are not linked any more in the relaxed formulation, the problem can be decomposed into two independent subproblems (\( \text{SAsub-1} \) and \( \text{SAsub-2} \)). An optimal solution to the relaxed problem can be obtained by solving the two subproblems and combining their optimal solutions.

\textbf{SAsub-1}

\[
\min \left\{ \sum_{i \in N} \sum_{k \in H} \bar{C}_{ik} Y_{ik} - \sum_{i \in N} \beta_i : \sum_{i \in N} Y_{ik} = p, \ Y_{ik} \leq Y_{kk} \quad \forall i, k, \ Y_{ik} \in \{0, 1\} \quad \forall i, k. \right\}
\]

\( \text{SAsub-1} \) contains only the allocation variable \( Y \) and is solved by a procedure as follows. Note that it can be completed within \( O(|N|^2) \).

Step 1: For \( i, k (i \neq k) \), set \( Y_{ik} = 1 \) if \( \bar{C}_{ik} < 0 \) and \( Y_{ik} = 0 \) otherwise. Compute \( S_i = \sum_{k \in N} \bar{C}_{ik} Y_{ik}, \) for each \( k \).

Step 2: Sort \( S_i \)'s in ascending order, choose \( p \) of the nodes with smaller \( S_i \), and set the corresponding \( Y_{ik} = 1 \) and set the remaining \( Y_{ik} \)'s to 0. Calculate the optimal value of \( \text{SAsub-1} \) by \( \sum_{k \in N} S_k Y_{kk} - \sum_{i \in N} \beta_i \).

Step 3: For \( i, k (i \neq k) \), set \( Y_{ik} \) to 0 if \( Y_{kk} = 0 \).
**SAsub-2**

\[
\begin{aligned}
\min & \quad \sum_{i \in N} \sum_{k \in \mathcal{H}} \sum_{m \in \mathcal{H}} \sum_{j=m}^{N} \left( F_{ikmj}w_{ij}(1 - q_{k} - d_{m}^{i} + \delta_{ijk.1} + \gamma_{ijk.1} + \delta_{ijm.2} + \gamma_{ijm.2})X_{ikmj} \right) \\
& \quad + \sum_{i \in N} \sum_{k \in \mathcal{H}} \sum_{m \in \mathcal{H}} \sum_{j=m}^{N} \left( F_{ikmj}w_{ij}(1 - q_{k} - d_{m}^{i} + \delta_{ijk.1} + \gamma_{ijk.1} + \delta_{ijm.2} + \gamma_{ijm.2})X_{ikmj} \right) \\
& \quad + \sum_{i \in N} \sum_{k \in \mathcal{H}} \sum_{m \in \mathcal{H}} \sum_{j=m}^{N} \left( F_{ikmj}w_{ij}(1 - q_{k} - d_{m}^{i} + \delta_{ijk.1} + \gamma_{ijk.1} + \delta_{ijm.2} + \gamma_{ijm.2})X_{ikmj} \right) \\
& \quad + \sum_{i \in N} \sum_{k \in \mathcal{H}} \sum_{m \in \mathcal{H}} \sum_{j=m}^{N} \left( F_{ikmj}w_{ij}(1 - q_{k} - d_{m}^{i} + \delta_{ijk.1} + \gamma_{ijk.1} + \delta_{ijm.2} + \gamma_{ijm.2})X_{ikmj} \right)
\end{aligned}
\]

subject to

\[
\begin{aligned}
& \quad \sum_{k \in \mathcal{H}} \sum_{m \in \mathcal{H}} X_{ikmj} = 1 \quad \forall i, j > i \\
& \quad U_{ijk} + \sum_{m \in \mathcal{H}} X_{ikmj} \leq 1 \quad \forall i, j > i, k \\
& \quad V_{ijm} + \sum_{k \in \mathcal{H}} X_{ikmj} \leq 1 \quad \forall i, j > i, m \\
& \quad X_{ikmj} \in \{0, 1\} \quad \forall i, j > i, k, m; U_{ijk}, V_{ijm} \in \{0, 1\} \quad \forall i, j > i, k
\end{aligned}
\]

**SAsub-2** includes the regular route variable \( \mathbf{X} \) and the backup hub variables \( \mathbf{U} \) and \( \mathbf{V} \). Constraints (14)–(16) are redundant in the original model. Nevertheless, including them in **SAsub-2** yields solutions that are more likely to be feasible to the original problem and therefore strengthens the lower bound obtained from Lagrangian relaxation. Note that the constraints in (14) require that each \( i - j \) flow has to go through one or two nodes to reach its destination; constraints in (15) and (16) ensure that for each \( i - j \) flow, the first/s node in its backup route must be different from the first/s node of its regular route. Note that, if a regular route is a single-hub route, so is its alternative route. Furthermore, in **SAsub-2**, an optimal solution for one \( i - j \) flow, i.e., a set of \( X_{ikmj} \), \( U_{ijk} \), and \( V_{ijm} \), is independent of those of others. So, it is sufficient to consider each individual \( i - j \) flow with the corresponding cost function from (13) and constraints from **SAsub-2**.

Although the cost function is nonlinear, every feasible solution has a clear combinatorial structure. As shown in Fig. 2(a), if \( i - j \) flow takes \( (i,k,m,j) \) as its regular route satisfying \( i \neq k \) and \( j \neq m \), a cost of \( F_{ikmj}w_{ij}(1 - q_{k} - d_{m}^{i} + \delta_{ijk.1} + \gamma_{ijk.1} + \delta_{ijm.2} + \gamma_{ijm.2}) \) will be incurred; if this flow takes \( n_{i} \) (\( n_{k} \), respectively) as the backup hub for \( k \) (\( m \), respectively), a cost of \( \rho F_{iknj}w_{ij} \gamma_{ijk.1} + \gamma_{ijn.1} \) (\( \rho F_{ikmj}w_{ij} \gamma_{ijm.1} + \gamma_{ijn.2} \), respectively) will be incurred additionally. A similar situation on transportation cost can be observed in Fig. 2(b) when the \( i - j \) flow selects a single-hub regular route. Such observations motivate us to develop the following enumeration procedure to identify an optimal solution to the \( i - j \) flow.

![Fig. 2. Transportation cost of solutions to SAsub-2.](image-url)
Step 1: For one pair of \((k, m)\), i.e., a given regular route, obtain its best alternative route (or best backup hubs) by computing all possible backup hubs that are different from \(k\) and \(m\) and selecting the alternative routes (or a single alternative route if \(k = m\)) with the least transportation cost. Compute the total transportation costs from both the regular route and the alternative routes.

Step 2: Repeat Step 1 for all \((k, m)\) pairs and identify the pair that provides the least total transportation cost. Denote that pair by \((k', m')\) and its corresponding best backup hubs by \(n_k'\) and \(n_m'\).

Step 3: Obtain an optimal solution to \(i - j\) flow by setting \(X_{kmj} = 1\) if \(k = k'\) and \(m = m'\), and otherwise to zero; setting \(U_{1n} = 1\) if \(n = n_k'\), and otherwise to zero; setting \(V_{ijn} = 1\) if \(n = n_m'\), and otherwise to zero.

The computational complexity of this procedure for one \(i - j\) flow is \(O(|N|^4)\) and therefore \(S_{\text{asub-2}}\) can be solved within \(O(|N|^6)\).

3.2.2. Upper bound and multiplier updating

To obtain a feasible solution as well as an upper bound, we apply a procedure similar to the one in Pirkul and Schilling (1998) that exploits the solution of \(S_{\text{asub-1}}\). Specifically, for each node \(i \in N\), its allocation will be retained if (4) is not violated. For the node with allocation infeasible to (4), given that hubs are already fixed after solving \(S_{\text{asub-1}}\), we select the lowest cost allocation. After determining \(Y\) variables, the regular route for each \(i - j\) flow is determined and its alternative routes can also be obtained by evaluating hubs not in the regular route and selecting the best ones.

We apply the classical subgradient algorithm described in Fisher (2004) to iteratively update the Lagrangian multipliers and to search for the best lower bound. Parameters such as step-size multiplier and maximum number of iterations are usually set up while applying the algorithm. The values of such parameters for the experimental study are described in Section 5.1.

3.2.3. Variable fixing

Variable fixing is an approach that uses both primal information from a feasible solution and dual information from Lagrangian multipliers to fix some variables in Lagrangian solution procedure. It has been proven to be effective in reducing search space and computation time (Snyder and Daskin, 2005; Contreras et al., 2011b).

Assume that we have the current best upper bound UB, \((\delta_1, \delta_2, \beta, \gamma_1, \gamma_2)\) are the current Lagrangian multipliers, and \((Y^*, X^*)\) is the corresponding optimal solution to the Lagrangian relaxed problem. Let \(f(\delta_1, \delta_2, \beta, \gamma_1, \gamma_2|C)\) be the optimal objective function value for \((\delta_1, \delta_2, \beta, \gamma_1, \gamma_2)\) under some condition \(C\). Then, we have the following results.

**Proposition 1.** When UB is strictly greater than LB,

1. if \(Y_{kk} = 1\) and \(f(\delta_1, \delta_2, \beta, \gamma_1, \gamma_2|Y_{kk} = 0) > UB\) for some \(k\), we have \(Y_{kk} = 1\) in any optimal solution;
2. if \(Y_{kk} = 0\) and \(f(\delta_1, \delta_2, \beta, \gamma_1, \gamma_2|Y_{kk} = 1) > UB\) for some \(k\), we have \(Y_{kk} = 0\) in any optimal solution.

**Proof.** We provide the proof for (1). Results in (ii) can be proven using similar arguments.

Note that \(f(\delta_1, \delta_2, \beta, \gamma_1, \gamma_2|Y_{kk} = 0)\) is a lower bound to \(R-\text{SAHMP}\) with a spoke node located in \(k\) for the given Lagrangian multipliers \((\delta_1, \delta_2, \beta, \gamma_1, \gamma_2)\). So, if

\[
f(\delta_1, \delta_2, \beta, \gamma_1, \gamma_2|Y_{kk} = 0) > UB,
\]

any solution to \(R-\text{SAHMP}\) with a spoke node in \(k\) will generate more cost than the current best feasible solution. Therefore, we have \(Y_{kk} = 1\) in any optimal solution to \(R-\text{SAHMP}\).

We mention that although more variable fixing rules can be developed, such as rules for the case of \(Y_{kk} = 0\), they will either be time-consuming to implement or have less impact on the Lagrangian relaxation. So, we only perform variable fixing procedure on \(Y_{kk}\) variable for each \(k\) with the best multipliers ever found once the Lagrangian procedure is terminated.

3.2.4. Branch-and-Bound strategies

If the subgradient method reaches the maximum number of iterations and the gap is still larger than the tolerance, the Lagrangian relaxation algorithm discussed in the previous section will be embedded in a Branch-and-Bound framework to further reduce the gap.

The Branch-and-Bound technique with Lagrangian relaxation has been implemented in reliable facility location models (Cui et al., 2010; Snyder and Daskin, 2005; Li and Ouyang, 2011). The results imply that branching on facility location variables is sufficient for determining an optimal network structure (Cui et al., 2010; Snyder and Daskin, 2005; Li and Ouyang, 2011). However, this is not the case for \(R-\text{SAHMP}\). Note that for a classical single allocation hub-and-spoke model, given fixed hubs, the remaining problem on spoke node allocation has been proven to be NP-hard. Thus, a more sophisticated two-stage Branch-and-Bound framework is developed and implemented in a width-first manner.
The first stage Branch-and-Bound is similar to that used for reliable facility location models in Snyder and Daskin (2005), where branching is made for $Y_{ik}$ (hub location) variables. At each Branch-and-Bound node, the hub location variable $Y_{ik}$ selected for branching is the unfixed open hub with the greatest assigned flow (without considering alternative routes), i.e.,

$$k^* = \arg \max_{k \in N} \left\{ \sum_{i \in N} \sum_{j \neq i} \sum_{m \in H_{(i)}} \sum_{p \in H_{(m)}} F_{ikmp} w_{ij} X_{ikm} + \sum_{i \in N} \sum_{j \neq i} \sum_{m \in H_{(k)}} w_{ij} X_{ikm} \right\}.$$

$Y_{ik}$ is forced to be 0 and then 1. The first stage Branch-and-Bound process will be terminated either with an optimal (including $\varepsilon$-optimal) solution or with $p$ hubs forced to open (or equivalently, $|N| - p$ hubs forced to close).

In the latter case, the second stage Branch-and-Bound method is applied to close the gap. Branching is made for $Y_{ik}$ (allocation) variables for spoke node $i$. First, the level of violation, $\nu^i$, for spoke node $i$ is computed. Given the current solution to the relaxed problem, the total number of violations to constraints in (2)–(4) for each $i$ are then calculated. The spoke node with the largest violation level $\nu^i$, say $i^*$, is selected for branching. Then, we partition the hub set $H$ (note that hub locations are already determined) into two sets $H_1$ and $H_2$ and create two nodes. Correspondingly, constraint $\sum_{k \in H_1} Y_{ik} = 1$ is added to the left-hand node and constraint $\sum_{k \in H_2} Y_{ik} = 1$ to the right-hand node. Once hub and spoke node allocation decisions are made, the remaining problem, including regular route and alternative route decisions, is polynomially solvable, which implies that no further branching is necessary.

During the whole Branch-and-Bound procedure, the set of Lagrangian multipliers that yields the smallest gap at a given node is passed to its child nodes as initial multipliers.

### 4. Reliable multiple allocation hub-and-spoke model

In this section, we consider the reliable MA-HLP model ($R$-MAHMP). Compared with the single allocation model, the multiple allocation model does not restrict flows from one source (or to one destination) to route through the same hub. As a result, we do not need to introduce spoke-hub allocation variables but simply introduce binary variables to define hubs.

#### 4.1. Reliable MA model: definition and formulation

The formulation for $R$-MAHMP is given below, most constraints reflect the requirements similar to those in $R$-SAHMP, $R-MAHMP$

$$\begin{align*}
\min & \sum_{i \in N} \sum_{k \in H} \sum_{m \in H_{(i)}} \sum_{p \in H_{(m)}} F_{ikmp} w_{ij} (1 - q_k - q_m^j) X_{ikm} \\
& + \sum_{i \in N} \sum_{j \in N} \left( \sum_{m \in H_{(j)}} F_{imj} w_{ij} (1 - q_m^j) X_{imj} + \sum_{k \in H_{(i)}} F_{ikj} w_{ij} (1 - q_k^j) X_{ikj} + F_{ij} w_{ij} X_{ij} \right) \\
& + \rho \left( \sum_{i \in N} \sum_{k \in H} \sum_{m \in H_{(k)}} \sum_{p \in H_{(m)}} \left( F_{ikmj} w_{ij} q_k^j X_{ikm} U_{ijn} + F_{ikmj} w_{ij} q_m^j X_{ikm} V_{ijn} \right) + \sum_{i \in N} \sum_{k \in H} \sum_{j \in N} \sum_{m \in H} F_{ikmj} w_{ij} q_k^j X_{ikmj} U_{ijn} \right) 
\end{align*}$$

subject to

$$\begin{align*}
\sum_{k \in H} X_{ikj} &= Y_j & \forall i, j > i \\
\sum_{m \in H_{(i)}} X_{imj} &= Y_i & \forall i, j > i \\
\sum_{i \in N} Y_i &= p \\
\sum_{k \in H} \sum_{m \in H_{(m)}} X_{ikm} &= 1 & \forall i, j > i \\
U_{ik} + \sum_{m \in H} X_{ikm} &\leq Y_k & \forall i, j > i, k \\
\sum_{k \in H} U_{ik} - \sum_{m \in H} X_{ikm} - \sum_{m \in H} X_{ikm} &= 1 & \forall i, j > i \\
V_{ijn} + \sum_{k \in H} X_{ikm} &\leq Y_m & \forall i, j > i, m \\
\sum_{m \in H} V_{ijn} - \sum_{k \in H} X_{ikj} - \sum_{k \in H} X_{ikj} &= 1 & \forall i, j > i \\
X_{ikm} &\in \{0, 1\} & \forall i, j > i, k, \forall m; & Y_k \in \{0, 1\} & \forall k; & U_{ik}, V_{ijn} \in \{0, 1\} & \forall i, j > i, k
\end{align*}$$
We use a binary variable $Y_k$ to indicate whether $k$ is a hub. Constraints (19) and (20) imply that if $i$ (or $j$) is a hub, it must be the first (or the second) hub in the routes of all flows from $i$ (or $j$). Constraints (22) require that each $i-j$ flow must have a route through hub(s).

Compared to the R-SAHMP, R-MAHMP is much simpler. First, Campell (1994) states that, for the classical MA-HLP, since there is no capacity restriction on links, each $i-j$ flow should be routed through the least-cost hub pair. So one optimal solution would always force the $X$ variables to be 1 or 0 and therefore there is no need to restrict $X$ variables to be binary. Second, MA-HLP is polynomial solvable if $p$ is fixed. In fact, these two observations still hold in R-MAHMP. For a given $p$, the R-MAHMP problem is polynomially solvable, and there exists one optimal solution such that all the flow variables $X_{ikm}$ take either 0 or 1 for all $i,j > i,k$ and $m$.

4.2. Solution methods for R-MAHMP

Note that the two linearization approaches described in Appendix A.2 could be applied to R-MAHMP with little modification. So, we only describe the development of a Lagrangian relaxation algorithm for R-MAHMP. We dualize constraints linking the route variables $X$ and the hub variables $Y$ and solve two resulting subproblems separately. By dualizing constraints in (19), (20), (23), and (25) with Lagrangian multipliers $\delta_{ij1}, \delta_{ij2}, \gamma_{jk1} \geq 0$, and $\gamma_{jm2} \geq 0$, we obtain subproblems MA-sub-1 and MA-sub-2 as follows.

MA-sub-1

$$\min \{ \sum_{k \in H} \sum_{i \in N} Y_k : \sum_{k \in H} Y_k = p, \ Y_k \in \{0,1\} \ \forall k \}$$

where $\bar{c}_k = -\sum_{i \in N, i \neq k} \delta_{ik1} - \sum_{i \in N, i \neq k} \delta_{ik2} - \sum_{i \in N, i \neq k} \gamma_{ik1}$. Clearly, MA-sub-1 can be solved by sorting variables according to their coefficients and selecting $p$ of them with smaller coefficients.

MA-sub-2

$$\min \sum_{i \in N} \sum_{k \in H} \sum_{m \in M} \sum_{j \in N} (F_{ikm}w_i(1 - q_k - q_m^{d_i}) + \gamma_{ij1} + \gamma_{ij2})X_{ikm}$$

$$+ \sum_{i \in N} \sum_{j \in N} \sum_{m \in M} F_{ijm}w_j(1 - q_m^{d_j}) + \delta_{ij1} + \gamma_{ij2})X_{ijm}$$

$$+ \sum_{i \in N} \sum_{j \in N} \sum_{m \in M} F_{ijm}w_j(1 - q_j^{d_j}) + \delta_{ij1} + \gamma_{ij2})X_{ijm}$$

$$+ \sum_{i \in N} \sum_{j \in N} \sum_{m \in M} \rho F_{ikm}w_jq_kX_{ikm}U_{ik} + \sum_{i \in N} \sum_{j \in N} \gamma_{ij1}U_{ij}$$

$$+ \sum_{i \in N} \sum_{j \in N} \sum_{m \in M} \rho F_{ikm}w_jq_mX_{ikm}V_{ijm} + \sum_{i \in N} \sum_{j \in N} \gamma_{jm2}V_{ijm}$$

subject to

$$\sum_{k \in H} \sum_{j \in N} X_{ikm} \leq 1 \ \forall i,j > i,k$$

$$\sum_{k \in H} \sum_{m \in M} X_{ikm} \leq 1 \ \forall i,j > i,m$$

Similar to SAsub-2, constraints (29) and (30) are supplied to get a tighter lower bound. Again, MA-sub-2 can be solved by using the combinatorial structure of each single $i-j$ flow. To obtain a feasible solution, as well as an upper bound, we take advantage of the result from MA-sub-1 to fix hubs. Then, an optimal solution for those given hubs can be determined by deriving an optimal route for each individual $i-j$ flow.

Lagrangian multipliers are updated iteratively by applying the classical subgradient algorithm. Also, a variable fixing strategy and a Branch-and-Bound technique which consider only hub location variables, are developed and implemented. Given that optimal routing decisions can be obtained in polynomial time if all hubs are fixed, this Branch-and-Bound procedure is guaranteed to be completed by branching on hub location variables only, which has a similar complexity to that of reliable facility location models in Snyder and Daskin (2005), Li and Ouyang (2011) and Cui et al. (2010).
5. Computational experiments

5.1. Data and design of experiments

We test our algorithms on the widely-used CAB data set (O’Kelly, 1987), which contains the distance between two nodes (interpreted as the transportation cost $c$) and origin-destination traffic flow $w$. We set the disruption rate $q_i$ to a random number within $[0.01, 0.05]$ for $i \in N$. We consider 36 combinations structured from setting the number of nodes $|N| = 10, 15, 20, 25$, the number of hubs $p = 3, 5, 7$, and inter-hub transportation cost discount factor $\alpha = 0.3, 0.5, 0.7$. Because rerouting flows will cause more operations and much longer waiting times, we set $\rho$ to 2 to represent this effect (Welman et al., 2010).

The aforementioned instances provide a test bed for both $R$-SAHMP and $R$-MAHMP models. We set the optimality tolerance, $\epsilon$, to 0.1% for all solution methods, including the off-the-shelf MIP solver CPLEX 12.1 that is adopted for benchmark. For the Lagrangian relaxation/Branch-and-Bound algorithm, the initial values of all multipliers are set to zero. The step-size multiplier, $\Delta$, is set to 6; the maximum number of iterations allowed to obtain an improvement of the lower bound is set to 50, i.e., when 50 consecutive iterations fail to improve the lower bound, $\Delta$ will be halved and the Lagrangian multipliers will be reset to the values used to get the best lower bound. The maximum number of iterations at the root node in the Branch-and-Bound tree is set to 3000 and at a child node it is set to 200. In the implementation of subgradient method, we terminate the Lagrangian procedure if one of the following conditions is met: (i) all Lagrangian multipliers are zero, which implies the current solution is proven to be optimal; (ii) the difference between the upper and lower bounds is below a threshold value $\epsilon$, i.e., an $\epsilon$-optimal solution is found; and (iii) the maximum number of iterations, 3000, is reached. If (iii) happens, the variable fixing procedure starts, then if applying variable fixing fails to reduce the gap to less than $\epsilon$, Branch-and-Bound is embedded into the Lagrangian relaxation algorithm. The maximum computation time is set as 3600 s. The problem is reported as unsolvable if no optimal solution is obtained within 3600 s.

All algorithms are implemented in C++, and all instances are tested on a Dell Optiplex 760 desktop computer (Intel Core 2 Duo CPU, 3.0 GHz, 3.25 GB of RAM) in Windows XP environment.

5.2. Performance of Lagrangian relaxation and Branch-and-Bound

Table 1 summarizes the computational results of our Lagrangian relaxation and Branch-and-Bound methods for instances of $R$-SAHMP and $R$-MAHMP. The column marked $\text{Iter}$. indicates the total number of Lagrangian iterations in all Branch-and-Bound nodes; the column marked $\text{Gap(\%)}$ provides the smallest relative gap we have achieved within the time limit. The column $BB_{\text{Nodes}}$ shows the total number of nodes evaluated in the procedure of Branch-and-Bound (excluding the root node); the column marked $\text{Time(s)}$ presents the total computational time in seconds for obtaining optimal solution, if some instances cannot be solved due to time limit or memory issue, we use $T$ or $M$, respectively, to represent the reason.

Similarly, Table 2 presents computational results of CPLEX 12.1 used to solve two types of linearized formulations, i.e., those obtained by the standard and a compact linearization methods, for $R$-SAHMP and $R$-MAHMP. Detailed derivations and concrete linear formulations are presented in the appendix. Results of instances with $|N| > 15$ are omitted because CPLEX fails to deal with larger instances within 3600 s.

The outcomes of the computational experiments show that: (i) The commercial solver CPLEX is of a very limited capability to solve practical instances with more than 10 nodes. With compact linearization formulation, the solver can provide feasible solutions; (ii) the Lagrangian relaxation algorithm with variable fixing and Branch-and-Bound is efficient in solving reliable models. All 72 instances can be solved to optimality within 1000 s; (iii) the Branch-and-Bound technique is necessary to derive optimal solutions for quite a few instances. This observation clearly shows that reliable models are more challenging than the classical ones for which study presented in Pirkul and Schilling (1998) shows that Lagrangian relaxation method itself is sufficient to solve CAB instances; and (iv) comparing reliable SA and MA models, the former often involves more Branch-and-bound nodes and longer computation times, which also confirms that the former one is of a higher complexity level than the latter one.

5.3. Analysis and discussion on system design and performance

In this section, we discuss the impact of reliable design paradigm on the system configurations and performance. The network configurations are compared with those determined by the classical hub-and-spoke models, which actually are special cases of the proposed $R$-SAHMP and $R$-MAHMP with the disruption probability $q = 0$.

5.3.1. Impact of hub unavailability on system design

Hub locations and spoke node allocations of reliable models could be different from those of classical models. Fig. 3 demonstrates a case with $|N| = 25$ (their associated disruption probabilities are presented in Table A2 in the appendix), $p = 5$, and $\alpha = 0.7$. Note that with classical hub-and-spoke network design, Philadelphia is selected as one of the hubs. Nevertheless, when the reliability issue is considered in the design, this hub is replaced by New York, and the spoke nodes in the service region of Philadelphia are re-allocated to New York as well.
Expected numbers of served passengers are calculated as the performance metrics and compared for different network configurations. It is a better measurement of airlines service quality for this study because the objective functions of reliable hub-and-spoke network models include the costs under both normal and disruption conditions which make them incomparable with the objective functions from classical models that only count the costs under normal condition. The following formulas are used to calculate the expected numbers of served passengers of classical ($P_{\text{sgc}}$) and reliable hub-and-spoke networks ($P_{\text{sgr}}$) respectively.

\[
P_{\text{sgc}} = \sum_{i \in N} \sum_{k \in H(i)} \sum_{m \in H} \sum_{j \in \{m\}} w_{ij} (1 - q_k - q_m^k) X_{ikmj} \sum_{i \in N} \sum_{k \in H(i)} \sum_{m \in H} \sum_{j \in \{m\}} w_{ij} (1 - q_m^k) X_{ikmj} + \sum_{k \in H(i)} \sum_{m \in H} \sum_{j \in \{m\}} w_{ij} (1 - q_m^k) X_{ikmj} + w_{ij} X_{ij}
\]

and

\[
P_{\text{sgr}} = P_{\text{sgc}} + \sum_{i \in N} \sum_{k \in H(i)} \sum_{m \in H} \sum_{j \in \{m\}} w_{ij} q_k X_{ikmj} U_{ijn} + \sum_{i \in N} \sum_{k \in H(i)} \sum_{m \in H} \sum_{j \in \{m\}} w_{ij} q_m X_{ikmj} U_{ijn} + \sum_{k \in H(i)} \sum_{m \in H} \sum_{j \in \{m\}} w_{ij} q_k X_{ikmj} U_{ijn}
\]

Given the disruption probabilities presented in Table A2, for the particular case discussed in this subsection, the classical network configuration is expected to transport 4,126,900 passengers and the reliable one 4,270,000 passengers (by both regular and alternative routes) with a 3.47% improvement. In fact, we want to highlight that, even without considering backup hubs and alternative routes, the derived reliable network system can transport more passengers (4,127,250) just by its regular routes than the classical network configuration. Such an observation indicates that it is necessary to consider the availability issue of network components when we design the network system for better performance.

### Table 1

<table>
<thead>
<tr>
<th>$N$</th>
<th>$p$</th>
<th>$\alpha$</th>
<th>R-SAHMP</th>
<th>R-MAHMP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Iter.</td>
<td>BB_Nodes</td>
<td>Gap (%)</td>
<td>Time (s)</td>
</tr>
<tr>
<td>10</td>
<td>3</td>
<td>0.3</td>
<td>250</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>5</td>
<td>0.3</td>
<td>565</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>0.3</td>
<td>184</td>
<td>0</td>
<td>0.098</td>
</tr>
<tr>
<td>5</td>
<td>0.5</td>
<td>257</td>
<td>0</td>
<td>0.098</td>
</tr>
<tr>
<td>5</td>
<td>0.5</td>
<td>1902</td>
<td>6</td>
<td>0.095</td>
</tr>
<tr>
<td>7</td>
<td>0.5</td>
<td>184</td>
<td>0</td>
<td>0.099</td>
</tr>
<tr>
<td>3</td>
<td>0.7</td>
<td>182</td>
<td>2</td>
<td>0.099</td>
</tr>
<tr>
<td>5</td>
<td>0.7</td>
<td>1515</td>
<td>4</td>
<td>0.096</td>
</tr>
<tr>
<td>7</td>
<td>0.7</td>
<td>323</td>
<td>0</td>
<td>0.098</td>
</tr>
<tr>
<td>15</td>
<td>3</td>
<td>0.3</td>
<td>1015</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>0.3</td>
<td>1353</td>
<td>4</td>
<td>0.099</td>
</tr>
<tr>
<td>7</td>
<td>0.3</td>
<td>1722</td>
<td>6</td>
<td>0.099</td>
</tr>
<tr>
<td>3</td>
<td>0.5</td>
<td>1362</td>
<td>4</td>
<td>0.095</td>
</tr>
<tr>
<td>5</td>
<td>0.5</td>
<td>1701</td>
<td>6</td>
<td>0.080</td>
</tr>
<tr>
<td>7</td>
<td>0.5</td>
<td>1313</td>
<td>6</td>
<td>0.090</td>
</tr>
<tr>
<td>3</td>
<td>0.7</td>
<td>980</td>
<td>2</td>
<td>0.099</td>
</tr>
<tr>
<td>5</td>
<td>0.7</td>
<td>1540</td>
<td>4</td>
<td>0.099</td>
</tr>
<tr>
<td>7</td>
<td>0.7</td>
<td>512</td>
<td>0</td>
<td>0.099</td>
</tr>
<tr>
<td>20</td>
<td>3</td>
<td>0.3</td>
<td>482</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>0.3</td>
<td>553</td>
<td>0</td>
<td>0.099</td>
</tr>
<tr>
<td>3</td>
<td>0.3</td>
<td>118</td>
<td>0</td>
<td>0.100</td>
</tr>
<tr>
<td>5</td>
<td>0.5</td>
<td>1762</td>
<td>6</td>
<td>0.098</td>
</tr>
<tr>
<td>7</td>
<td>0.5</td>
<td>1584</td>
<td>4</td>
<td>0.099</td>
</tr>
<tr>
<td>3</td>
<td>0.7</td>
<td>589</td>
<td>0</td>
<td>0.099</td>
</tr>
<tr>
<td>3</td>
<td>0.7</td>
<td>3925</td>
<td>16</td>
<td>0.097</td>
</tr>
<tr>
<td>5</td>
<td>0.7</td>
<td>3871</td>
<td>14</td>
<td>0.099</td>
</tr>
<tr>
<td>7</td>
<td>0.7</td>
<td>2095</td>
<td>8</td>
<td>0.097</td>
</tr>
<tr>
<td>25</td>
<td>3</td>
<td>0.3</td>
<td>2020</td>
<td>6</td>
</tr>
<tr>
<td>5</td>
<td>0.3</td>
<td>965</td>
<td>2</td>
<td>0.100</td>
</tr>
<tr>
<td>3</td>
<td>0.3</td>
<td>812</td>
<td>2</td>
<td>0.100</td>
</tr>
<tr>
<td>5</td>
<td>0.5</td>
<td>1745</td>
<td>6</td>
<td>0.097</td>
</tr>
<tr>
<td>7</td>
<td>0.5</td>
<td>2774</td>
<td>10</td>
<td>0.099</td>
</tr>
<tr>
<td>5</td>
<td>0.7</td>
<td>201</td>
<td>0</td>
<td>0.082</td>
</tr>
<tr>
<td>3</td>
<td>0.7</td>
<td>765</td>
<td>4</td>
<td>0.076</td>
</tr>
<tr>
<td>5</td>
<td>0.7</td>
<td>7318</td>
<td>34</td>
<td>0.096</td>
</tr>
<tr>
<td>7</td>
<td>0.7</td>
<td>879</td>
<td>2</td>
<td>0.100</td>
</tr>
</tbody>
</table>
5.3.2. Performance of reliable hub-and-spoke networks

The expected numbers of served passengers of reliable models and those of the classical models are further compared for more scenarios. Results are listed in Table 3. In the table, the performance measures ($P_{\text{sgc}}$ and $P_{\text{sgr}}$) for the classical and reliable model are presented with numerical values and the relative improvements (denoted by $\text{Improvement}$) achieved by the reliable model are shown in percentages. In all experiments, the inter-hub transportation cost discount factor $\alpha$ is set to 0.7.

Note that since our model can handle any single hub disruption, the number of served passengers is exactly the total transportable flow $\sum_{i=1}^{N} \sum_{j=2}^{N} w_{ij}$, which is constant for each fixed $|N|$. It is observed that the reliable network always transports more passengers compared to classical model, with the magnitude increasing with the growth of the network scale $|N|$. Therefore, in terms of the expected number of served passengers, the reliable models clearly outperform the classical ones.

5.3.3. Verification with correlated multiple disruptions

One assumption we made in developing reliable models is that no more than one hub will fail at any time. In some extreme cases, such an assumption may not valid and multiple failures could occur simultaneously. So, in this section, we perform numerical experiments to evaluate the influence of the single disruption (SD) assumption. We study the optimal network configurations obtained from our models in an environment that correlated multiple disruption (MD) may occur.

Table 2
Solver performance for R-SAHMP and R-MAHMP.

| $|N|$ | $p$ | $\alpha$ | R-SAHMP | R-MAHMP |
|-----|-----|--------|---------|---------|
|     |     |        | StdLinear | CptLinear | Time (s) | Gap (%) | Time (s) | Gap (%) | Time (s) | Gap (%) |
| 10  | 3   | 0.3    | 33.7 0.032 T | 0.514 | 641.1 0.100 | 1456.7 0.100 |
| 5   | 0.3 | 24.5 0.047 T | 1.827 | 3516.3 0.100 | T 1.896 |
| 7   | 0.3 | 5.2 0.000 2.2 0.000 | 138.4 0.100 | 4.4 0.100 |
| 3   | 0.5 | 40.4 0.000 T | 2.069 | 343.5 0.099 T | 0.321 |
| 5   | 0.5 | 35.3 0.000 T | 3.414 | T 0.164 | 2041.3 0.000 |
| 7   | 0.5 | 7.1 0.006 4.5 0.094 | 520.9 0.100 | 76.5 0.100 |
| 3   | 0.7 | 50.1 0.000 T | 2.007 | 407.6 0.100 | 1335.4 0.100 |
| 5   | 0.7 | 39.2 0.010 T | 1.660 | M 0.760 T | 1.951 |
| 7   | 0.7 | 7.6 0.000 19.7 0.099 | M 0.330 M | 0.740 |
| 15  | 3   | 0.3 | M NA T | 4.030 | M 16.360 T | 4.440 |
| 5   | 0.3 | M NA M | 5.070 | M 14.480 T | 5.441 |
| 7   | 0.3 | M NA T | 4.789 | M 18.660 T | 3.669 |
| 3   | 0.5 | M NA T | 3.729 | M 11.650 T | 4.531 |
| 5   | 0.5 | M NA M | 5.340 | M 14.960 T | 4.620 |
| 7   | 0.5 | M NA T | 4.020 | M 13.560 T | 3.117 |
| 3   | 0.7 | M NA T | 4.907 | M 10.110 T | 3.949 |
| 5   | 0.7 | M NA M | 4.560 | M 9.770 T | 3.723 |
| 7   | 0.7 | M NA T | 3.480 | M 9.600 T | 2.662 |

(a) Configuration from classical model (b) Configuration from reliable model

Fig. 3. Optimal system configurations in different SA models.
Letting the random variable \(D_k\) be the status for any hub \(k\), i.e., \(D_k = 1\) when hub \(k\) is down and 0 otherwise, we use the following equations to recalculate the expected number of passengers to be served with possible multiple hub disruptions in the real situation:

\[
P_{sg_0} = \sum_{i \in H} \sum_{j \in H} \sum_{k \in H} \sum_{m \in H} \sum_{n \in H} \left( \sum_{k \in H} \sum_{m \in H} \sum_{n \in H} \sum_{j \in H} \sum_{i \in H} \left( \sum_{i \in H} \sum_{j \in H} \sum_{k \in H} \sum_{m \in H} \sum_{n \in H} \right) \right)
\]

and

\[
P_{sg_r} = \sum_{i \in H} \sum_{j \in H} \sum_{k \in H} \sum_{m \in H} \sum_{n \in H} \left( \sum_{i \in H} \sum_{j \in H} \sum_{k \in H} \sum_{m \in H} \sum_{n \in H} \right) \]

Given that \(P(D_k = 1) = q_k\) for any hub \(k\), by setting a correlation \(corr(D_k, D_m)\) of any pair of random variables \((D_k, D_m)\) and assuming a relationship between \(P(D_m = 0, D_n = 0|D_k = 1)\) and \(P(D_m = 0, D_n = 0)\) we can obtain the probabilities in (33) and (34). Specifically, we want the correlation between given nodes \(k\) and \(m\) decreases as the distance \(c_{km}\) grows, so we choose \(corr(D_k, D_m) = e^{-\Gamma_1 c_{km}}\) where \(\Gamma_1\) is a positive constant. In order to avoid the situation in which the correlation decreases too fast, \(\Gamma_1\) is set to \(1 \times 100\) (see Fig. 4). Note that under this correlation assumption, the geographically close nodes can have high correlations. For instance, \(corr(D_5, D_8) = 0.624\) (Cleveland and Detroit). Based on the correlation function, we can derive the

![Fig. 4. Curve of Correlation between Dk and Dm.](image-url)
required \( P(D_k = 0, D_m = 0) \) and \( P(D_k = 1, D_n = 0) \). For the probabilities involving three nodes like \( P(D_k = 1, D_m = 0, D_n = 0) \), we further assume that \( P(D_m = 0, D_n = 0|D_k = 1) = P(D_m = 0, D_n = 0) \left( 1 - \frac{c_{kn} \cdot \text{C}}{c_{km} \cdot \text{C}} \right) \), i.e., \( P(D_m = 0, D_n = 0|D_k = 1) \) is related to but smaller than \( P(D_m = 0, D_n = 0) \) and also determined by the average distance \( \frac{c_{kn} \cdot \text{C}}{c_{km} \cdot \text{C}} \); then \( P(D_k = 1, D_m = 0, D_n = 0) \) can be easily calculated. See Appendix A.4 for details. We mention that by changing the form of the correlation function, we can even model negative correlation. Therefore, (33) and (34) provide us a useful tool to evaluate a hub-and-spoke network in the real practice in which correlated multiple node failures may occur.

First, the relative decrease of expected served passengers with respect to that under the single disruption assumption is listed in the column “Change(%)” of Table 4. It is easy to observe that in terms of expected served passengers, the influence of the multiple hub disruptions to the system performance is small (all less than 0.5%). Next, expected served passengers of the reliable model and the classical model under multiple disruptions are computed and listed in Table 5. According to the results, proposed reliable models outperform the classical ones under correlated multiple disruptions as well.

Finally, a sensitivity analysis of failure rates on system configurations is conducted both for classical and reliable models. Assuming that all nodes have the same hub disruption probability, we investigate the impact of small variation in failure rate \( q \) on the aforementioned performance measures, \( P_{sg} \) and \( P_{sg}' \). Both low \((q = 0.009)\) and high probability scenarios \((q = 0.04)\) are considered for multiple disruption scenarios. In Table 6, numerical results for \( |N| = 25, p = 3.5, 7, \) and \( z = 0.7 \) are presented, with the columns \( P_{sg} \) and \( P_{sg}' \) representing the expected number of passengers of the corresponding network configuration with initial hub failure rates and the column \( \text{Change(\%)} \) representing the percentage change from \( P_{sg}' \) to \( P_{sg} \) when \( q \) is increased by 0.001 while keeping the network configuration fixed.

A clear observation is that the reliable model is much less sensitive than the classical model to the variations of hub availability. The reliable networks have a higher survivability and are more robust to disruptions. Such observations again demonstrate the importance of taking into account hub unavailabilities in designing robust hub-and-spoke networks.

### Table 4

Relative change of expected served passengers with different assumptions.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>SA model</th>
<th></th>
<th>MA model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>SD</td>
<td>MD</td>
<td>Change (%)</td>
</tr>
<tr>
<td>10</td>
<td>3</td>
<td>499,513</td>
<td>497,023</td>
<td>-0.498</td>
</tr>
<tr>
<td>5</td>
<td>499,513</td>
<td>497,444</td>
<td>-0.414</td>
<td>499,513</td>
</tr>
<tr>
<td>7</td>
<td>499,513</td>
<td>498,563</td>
<td>-0.190</td>
<td>499,513</td>
</tr>
<tr>
<td>15</td>
<td>3</td>
<td>1,182,470</td>
<td>1,181,630</td>
<td>-0.071</td>
</tr>
<tr>
<td>5</td>
<td>1,182,470</td>
<td>1,179,170</td>
<td>-0.279</td>
<td>1,182,470</td>
</tr>
<tr>
<td>7</td>
<td>1,182,470</td>
<td>1,179,450</td>
<td>-0.255</td>
<td>1,182,470</td>
</tr>
<tr>
<td>20</td>
<td>3</td>
<td>2,877,300</td>
<td>2,871,970</td>
<td>-0.185</td>
</tr>
<tr>
<td>5</td>
<td>2,877,300</td>
<td>2,873,540</td>
<td>-0.131</td>
<td>2,877,300</td>
</tr>
<tr>
<td>7</td>
<td>2,877,300</td>
<td>2,873,220</td>
<td>-0.244</td>
<td>2,877,300</td>
</tr>
<tr>
<td>25</td>
<td>3</td>
<td>4,270,000</td>
<td>4,263,040</td>
<td>-0.143</td>
</tr>
<tr>
<td>5</td>
<td>4,270,000</td>
<td>4,262,770</td>
<td>-0.170</td>
<td>4,270,000</td>
</tr>
</tbody>
</table>

### Table 5

Performance of reliable models under the multiple disruption assumption.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>SA model</th>
<th></th>
<th></th>
<th>MA model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Classical</td>
<td>( P_{sg} )</td>
<td>( P_{sg}' )</td>
<td>Reliable Improvement (%)</td>
</tr>
<tr>
<td>10</td>
<td>3</td>
<td>488,612</td>
<td>497,023</td>
<td>1.721</td>
<td>490,613</td>
</tr>
<tr>
<td>5</td>
<td>491,128</td>
<td>497,444</td>
<td>1.286</td>
<td>494,187</td>
<td>498,744</td>
</tr>
<tr>
<td>7</td>
<td>494,733</td>
<td>498,563</td>
<td>0.774</td>
<td>496,130</td>
<td>498,853</td>
</tr>
<tr>
<td>15</td>
<td>3</td>
<td>1,155,140</td>
<td>1,181,630</td>
<td>2.293</td>
<td>1,163,250</td>
</tr>
<tr>
<td>5</td>
<td>1,157,080</td>
<td>1,179,170</td>
<td>1.909</td>
<td>1,164,630</td>
<td>1,179,760</td>
</tr>
<tr>
<td>7</td>
<td>1,162,000</td>
<td>1,179,450</td>
<td>1.502</td>
<td>1,171,440</td>
<td>1,180,440</td>
</tr>
<tr>
<td>20</td>
<td>3</td>
<td>2,783,080</td>
<td>2,871,970</td>
<td>3.194</td>
<td>2,828,600</td>
</tr>
<tr>
<td>5</td>
<td>2,802,660</td>
<td>2,873,540</td>
<td>2.529</td>
<td>2,833,540</td>
<td>2,874,820</td>
</tr>
<tr>
<td>7</td>
<td>2,805,300</td>
<td>2,870,280</td>
<td>2.316</td>
<td>2,846,600</td>
<td>2,873,950</td>
</tr>
<tr>
<td>25</td>
<td>3</td>
<td>4,137,710</td>
<td>4,263,040</td>
<td>3.022</td>
<td>4,164,820</td>
</tr>
<tr>
<td>5</td>
<td>4,129,250</td>
<td>4,263,040</td>
<td>3.240</td>
<td>4,182,190</td>
<td>4,265,340</td>
</tr>
<tr>
<td>7</td>
<td>4,183,440</td>
<td>4,262,290</td>
<td>1.885</td>
<td>4,217,820</td>
<td>4,264,660</td>
</tr>
</tbody>
</table>
5.3.4. Application of proposed reliable models to a recent airlines merger

The recent merger between United and Continental Airlines brings the new United Airlines (UA) eight domestic hubs. The hub at Cleveland Hopkins Airport shares a great functional similarity with the hub at Chicago O’Hare and is expected to be closed to save cost by industrial experts (Grossman, 2010). In this section, we apply the proposed reliable models to UA network and evaluate different network configurations in a quantitative way. Our analysis uses the proposed reliable MA network with CAB data set under the correlated multiple disruption assumption with current eight hubs in UA. Parameter $q$ is shown in Table A2, $j N$ is set as 25 and $\alpha = 0.7$. We evaluate two performance measurements, i.e., the expected number of served passengers and the expected transportation cost, under different single hub closing options. We point out that our study is simply for demonstration as UA’s coverage and traffic flows may be very different from those from CAB data set.

We first compute the impact of closing Cleveland and obtain corresponding results: the expected number of served passengers is 4.25603 × 10^8 and the expected transportation cost is 3.54629 × 10^9. Then, we compute results of closing any of other hubs and calculate the differences compared with the result of closing Cleveland. The outcomes are presented in Fig. 5.

For example, closing the hub in New York will result in 4.255 × 10^8 served passengers and a transportation cost of 3.71088 × 10^9, which are 1030 less passengers and 1.6459 × 10^8 more cost compared to the performance of closing the Cleveland hub (as shown in Fig. 5(a)).

It is observed that the disruption probability of Cleveland ($q_{5}$) is relatively high in Table A2 (0.047 in the range of 0.012–0.050 for all 25 nodes). A different scenario with $q_{5}$ equal to 0.025 is evaluated and the corresponding results are presented in Fig. 5(b). We observe that, from the perspective of transportation cost, the hub in Cleveland is always the optimal choice to be closed. This quantitative analysis endorses the opinion from the industrial expert. Nevertheless, if the number of served passengers is of a higher priority, closing the hub at Washington DC becomes a better option. Although no current information of UA but the CAB data set is used, this quantitative analysis demonstrates that the proposed reliable models and algorithms can be used to provide decision support to the management of airlines to re-structure their networks. Similarly, they can be used by airlines for identifying strategic partners/alliance to hedge against disruptions and achieve their desired operational goals.

6. Conclusions

In this study, we construct reliable single and multiple allocation hub-and-spoke models that generalize their classical counterparts. Our models seek to build hub-and-spoke systems with backup hubs and alternative routes to better hedge...
against various disrupted situations in practice. Due to the complexity of the reliable models, we develop a set of easy-to-implement Lagrangian relaxation/Branch-and-Bound algorithms that can compute optimal solutions efficiently. Computational study demonstrates the effectiveness of these algorithms, as well as the superiority of the proposed models to classical models in terms of serving passengers and being robust subject to the variations of hub failure rates.

To the best of our knowledge, our work is the first analytical study on reliable hub-and-spoke network design problem. It theoretically extends the existing literature on reliable network design and also has a clear practical impact on transportation and telecommunications systems. The proposed models can be slightly modified to deal with different situations, such as just allowing a subset of nodes being chosen to be hubs and allowing a subset of flows to be rerouted. Therefore, they are powerful decision support tools for system designers to derive optimal system configuration with a desired trade-off among performance measures.

Nevertheless, the proposed models have significant caveats that need to be addressed in future research. Although it is demonstrated that the resulted network settings from proposed models outperform those from classical models under correlated multiple disruption scenarios, explicitly including multiple disruption into mathematical modeling is a desire and should be considered in future research. Furthermore, more complicated issues in practice, such as congestion effect, should also be taken into account.

Acknowledgement

We thank Dr. Anthony Chen for providing valuable comments and several useful references to our research.

Appendix A

A.1. Nomenclature

See Table A1.

<table>
<thead>
<tr>
<th>Type</th>
<th>Notation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set</td>
<td>N</td>
<td>Set of nodes</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>Set of candidate hubs</td>
</tr>
<tr>
<td>Index</td>
<td>i, k, m, j, n</td>
<td>Node indices, range from 0 to</td>
</tr>
<tr>
<td>Parameter</td>
<td>c_{ij}</td>
<td>Unit transportation cost between i and j</td>
</tr>
<tr>
<td></td>
<td>w_{ij}</td>
<td>Transportable flow between i and j</td>
</tr>
<tr>
<td></td>
<td>q_k</td>
<td>Failure probability of node k</td>
</tr>
<tr>
<td></td>
<td>q_{mk}</td>
<td>Failure probability of node m if m ≠ k; 0, otherwise</td>
</tr>
<tr>
<td></td>
<td>z</td>
<td>Discount factor of inter-hub links</td>
</tr>
<tr>
<td></td>
<td>F_{ikmj}</td>
<td>Unit transportation cost of the route (i, k, m, j)</td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>Penalty factor of using alternative routes</td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>Total number of hubs</td>
</tr>
<tr>
<td></td>
<td>a_{ik}</td>
<td>Coefficient of the variable Y_k in SAsub_1</td>
</tr>
<tr>
<td></td>
<td>s_i</td>
<td>Summation of T_{ij}Y_k over i</td>
</tr>
<tr>
<td></td>
<td>Δ</td>
<td>Step-size multiplier of the subgradient algorithm</td>
</tr>
<tr>
<td></td>
<td>ɛ</td>
<td>Violation level of node i</td>
</tr>
<tr>
<td></td>
<td>c_1, c_2</td>
<td>Lagrangian multipliers</td>
</tr>
<tr>
<td></td>
<td>b</td>
<td>Lagrangian multiplier</td>
</tr>
<tr>
<td>Variable</td>
<td>Y_k</td>
<td>1 if node k is chosen as a hub; 0, otherwise</td>
</tr>
<tr>
<td></td>
<td>Y_i</td>
<td>1 if node i is assigned to k; 0, otherwise</td>
</tr>
<tr>
<td></td>
<td>X_{ikmj}</td>
<td>1 if the flow between node i and j is routed through k and m; 0, otherwise</td>
</tr>
<tr>
<td></td>
<td>U_{ij}</td>
<td>1 if n is the backup hub of the first hub in the route between i and j; 0, otherwise</td>
</tr>
<tr>
<td></td>
<td>V_{ij}</td>
<td>1 if n is the backup hub of the second hub in the route between i and j; 0, otherwise</td>
</tr>
<tr>
<td></td>
<td>s_1, s_2</td>
<td>Lagrangian multipliers</td>
</tr>
<tr>
<td></td>
<td>β</td>
<td>Lagrangian multiplier</td>
</tr>
<tr>
<td></td>
<td>Y_1, Y_2</td>
<td>Lagrangian multipliers</td>
</tr>
</tbody>
</table>
A.2. Linearization techniques and CPLEX performance

A.2.1. Standard linearization of R-SAHP

A linear reformulation for R-SAHP is obtained by using standard linearization techniques (Nemhauser and Wolsey, 1988), which is denoted by StdLinear.

\[
\text{StdLinear} \\
\min \sum_i \sum_k \sum_m \sum_{j:j'\neq j} F_{ikmj} w_{ij} (1 - q_k - q_{m}^j) X_{ikmj} + \sum_i \sum_{j:j'\neq j} \left( \sum_{m\neq k} F_{ikmj} w_{ij} (1 - q_k) X_{ikmj} + \sum_{k\neq k} F_{ikjj} w_{ij} (1 - q_k) X_{ikjj} + F_{ijjj} w_{ij} X_{ijjj} \right) \\
+ \sum_i \sum_k \sum_m \sum_{j:j'\neq j} \rho (F_{ikmj} w_{ij} q_k Z_{ikmj} + F_{ikjj} w_{ij} q_k Z_{ikjj}^2) + \sum_i \sum_k \sum_{j:j'\neq j} \sum_{n} \rho F_{iknj} w_{ij} q_k Z_{iknj}^2 \\
\text{subject to} \\
(2) - (10) \\
Z_{ikmj}^1 \leq X_{ikmj}, Z_{ikmj}^1 \leq U_{ijn}, Z_{ikmj}^2 \geq X_{ikmj} + U_{ijn} - 1 \quad \forall i, k, m, j > i, n \\
Z_{ikmn}^1 \leq X_{ikmn}, Z_{ikmn}^2 \leq V_{ijn}, Z_{ikmn}^2 \leq X_{ikmn} + V_{ijn} - 1 \quad \forall i, k, m, j > i, n \\
Z_{ikmj}^1, Z_{ikmj}^2 \geq 0 \quad \forall i, k, m, j > i, n
\]

Compared with the quadratic form in (1)–(10), \(X_{ikmj} U_{ijn}\) is replaced by \(Z_{ikmj}^1\) and \(X_{ikmj} V_{ijn}\) is replaced by \(Z_{ikmj}^2\). Also, a few sets of constraints are added to enforce that \(Z_{ikmj}^1 = X_{ikmj} U_{ijn}\) and \(Z_{ikmj}^2 = X_{ikmj} V_{ijn}\). Note that this mixed integer linear reformulation has to deal with a huge number of additional variables and constraints.

A.2.2. Compact linear reformulation of R-SAHP

A recent linearization approach and its variants are developed for quadratic 0–1 programs to obtain a compact linear reformulation (see Chaovalitwongse et al. (2004); Sherali and Smith, 2007; and He et al., 2012). While the standard one introduces a quadratic number of extra variables and constraints, this type of linearization method introduces only a linear number of extra variables and constraints. Therefore, we adopt and extend this linearization technique to reformulate our quadratic R-SAHP model.

First, we point out that \(\sum_k \sum_{m\neq k} X_{ikmj} \in \{0, 1\}\) for all \(i, j > i\). Because this expression appears in the objective function of R-SAHP, we can treat it simply as a binary variable as a whole and perform linearization with respect to \(\sum_k \sum_{m\neq k} X_{ikmj}\) and \(U_{ijn}\) (\(\sum_{k,m} X_{ikmj}\) and \(V_{ijn}\) can be linearized similarly). We obtain the compact CptLinear formulation as follows.

\[
\text{CptLinear} \\
\min \sum_i \sum_k \sum_m \sum_{j:j'\neq j} F_{ikmj} w_{ij} (1 - q_k - q_{m}^j) X_{ikmj} + \sum_i \sum_{j:j'\neq j} \left( \sum_{m\neq k} F_{ikmj} w_{ij} (1 - q_k) X_{ikmj} + \sum_{k\neq k} F_{ikjj} w_{ij} (1 - q_k) X_{ikjj} + F_{ijjj} w_{ij} X_{ijjj} \right) \\
+ \sum_i \sum_k \sum_m \sum_{j:j'\neq j} \sum_{n} (\Omega_{ijn} - \sigma_j U_{ijn}) + \sum_i \sum_{j:j'\neq j} \sum_{n} (\Theta_{ijn} - \sigma_j V_{ijn}) + \sum_i \sum_{j:j'\neq j} \sum_{n} (\Gamma_{ijn} - \sigma_j U_{ijn}) \\
\text{subject to} \\
(2) - (10) \\
\sum_{k,m\neq k} \rho w_{ij} q_k F_{ikmj} X_{ikmj} - s_{ijn} + \sigma_j = \Omega_{ijn} \quad \forall i, j > i, n \quad (A1) \\
\sum_{k,m\neq k} \rho w_{ij} q_k F_{ikmj} X_{ikmj} - t_{ijn} + \sigma_j = \Theta_{ijn} \quad \forall i, j > i, n \quad (A2) \\
\sum_{k,m\neq k} \rho w_{ij} q_k F_{ikmj} X_{ikmj} - r_{ijn} + \sigma_j = \Gamma_{ijn} \quad \forall i, j > i, n \quad (A3) \\
\sum_{k,m\neq k} \rho w_{ij} q_k F_{ikmj} X_{ikmj} - u_{ijn} + \sigma_j = \Omega_{ijn} \quad \forall i, j > i, n \quad (A4) \\
\sum_{k,m\neq k} \rho w_{ij} q_k F_{ikmj} X_{ikmj} - v_{ijn} + \sigma_j = \Theta_{ijn} \quad \forall i, j > i, n \quad (A5) \\
\sum_{k,m\neq k} \rho w_{ij} q_k F_{ikmj} X_{ikmj} - w_{ijn} + \sigma_j = \Gamma_{ijn} \quad \forall i, j > i, n \quad (A6) \\
\Omega_{ijn}, s_{ijn}, \Theta_{ijn}, t_{ijn}, \Gamma_{ijn}, r_{ijn} \geq 0 \quad \forall i, j > i, n \quad (A7)
\]

where \(\mu_j = \rho w_{ij} \max_{k,m} (F_{ikmj}) \max_k (q_k)\), and \(\sigma_j \geq 0\) is a predetermined coefficient for \(i, j > i\). In our numerical study, we set \(\sigma_j = 0\) for all \(i\) and \(j\). The linearization for the quadratic term \(\sum_i \sum_{k,m\neq k} \sum_{j:j'\neq j} \sum_{n} \rho F_{ikmj} w_{ij} q_k X_{ikmj} U_{ijn}\) in R-SAHP is completed by new variables \(\Omega_{ijn}\), \(s_{ijn}\) and constraints (A1) and (A2). Similarly, \(\Theta_{ijn}, t_{ijn}, (A3)\) and (A4) are used for linearization of
The linear CptLinear formulation is equivalent to the quadratic R-SAHMP model. An optimal solution to CptLinear yields an optimal solution to R-SAHMP. □

Assume that \(\text{corr}(D_k, D_m) = f(c_{km})\) for any \(k\) and \(m\). We let \(f(x)\) takes the form \(e^{-x^2}\) so that the correlation becomes smaller as the distance between two nodes increases.

Since \(D_k\) is a binary random variable, \(E(D_k) = 0 \ast P(D_k = 0) + 1 \ast P(D_k = 1)\), we have

\[
E(D_k) = q_k.
\]

So,

\[
\frac{E(D_k - q_k)(D_m - q_m)}{\sigma_{D_k} \sigma_{D_m}} = \text{corr}(D_k, D_m),
\]

\[
\frac{E(D_kD_m) - q_kq_m}{\sigma_{D_k} \sigma_{D_m}} = \text{corr}(D_k, D_m).
\]

Noting that \(D_k^2\) and \(D_kD_m\) are both binary random variable,

\[
E(D_kD_m) = P(D_k = 1, D_m = 1),
\]

\[
\sigma_{D_k}^2 = E(D_k^2) - E^2(D_k) = q_k - q_k^2.
\]

Hence we can obtain \(P(D_k = 1, D_m = 1)\) by the following equation:

\[
P(D_k = 1, D_m = 1) = \text{corr}(D_k, D_m) \sqrt{q_k - q_k^2} \sqrt{q_m - q_m^2} + q_kq_m.
\]

Then,

\[
P(D_k = 1, D_m = 0) = P(D_k = 1) - P(D_k = 1, D_m = 1),
\]

\[
P(D_k = 0, D_m = 1)\) can be obtained similarly.

We also need \(P(D_m = 0, D_k = 0)\):

\[
P(D_m = 0, D_k = 0) = 1 - P(D_m = 1, D_k = 1) - P(D_m = 0, D_k = 1) - P(D_m = 1, D_k = 0)
\]

### A.3. Sample disruption probabilities for CAB Data Set

See Table A2.

### A.4. Disruption probabilities in (33) and (34)

#### A.4.1. The probabilities involving two nodes \(P(D_k = 0, D_m = 0)\) and \(P(D_k = 1, D_m = 0)\)

Since \(D_k\) is a binary random variable, \(E(D_k) = 0 \ast P(D_k = 0) + 1 \ast P(D_k = 1)\), we have

\[
E(D_k) = q_k.
\]

So,

\[
\frac{E(D_k - q_k)(D_m - q_m)}{\sigma_{D_k} \sigma_{D_m}} = \text{corr}(D_k, D_m),
\]

\[
\frac{E(D_kD_m) - q_kq_m}{\sigma_{D_k} \sigma_{D_m}} = \text{corr}(D_k, D_m).
\]

Noting that \(D_k^2\) and \(D_kD_m\) are both binary random variable,

\[
E(D_kD_m) = P(D_k = 1, D_m = 1),
\]

\[
\sigma_{D_k}^2 = E(D_k^2) - E^2(D_k) = q_k - q_k^2.
\]

Hence we can obtain \(P(D_k = 1, D_m = 1)\) by the following equation:

\[
P(D_k = 1, D_m = 1) = \text{corr}(D_k, D_m) \sqrt{q_k - q_k^2} \sqrt{q_m - q_m^2} + q_kq_m.
\]

Then,

\[
P(D_k = 1, D_m = 0) = P(D_k = 1) - P(D_k = 1, D_m = 1),
\]

\[
P(D_k = 0, D_m = 1)\) can be obtained similarly.

We also need \(P(D_m = 0, D_k = 0)\):

\[
P(D_m = 0, D_k = 0) = 1 - P(D_m = 1, D_k = 1) - P(D_m = 0, D_k = 1) - P(D_m = 1, D_k = 0)
\]

### Table A2

<table>
<thead>
<tr>
<th>No.</th>
<th>City</th>
<th>Value</th>
<th>No.</th>
<th>City</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Atlanta</td>
<td>0.023</td>
<td>9</td>
<td>Houston</td>
<td>0.026</td>
</tr>
<tr>
<td>1</td>
<td>Baltimore</td>
<td>0.017</td>
<td>10</td>
<td>Kansas City</td>
<td>0.018</td>
</tr>
<tr>
<td>2</td>
<td>Boston</td>
<td>0.047</td>
<td>11</td>
<td>Los Angeles</td>
<td>0.049</td>
</tr>
<tr>
<td>3</td>
<td>Chicago</td>
<td>0.041</td>
<td>12</td>
<td>Memphis</td>
<td>0.024</td>
</tr>
<tr>
<td>4</td>
<td>Cincinnati</td>
<td>0.026</td>
<td>13</td>
<td>Miami</td>
<td>0.027</td>
</tr>
<tr>
<td>5</td>
<td>Cleveland</td>
<td>0.047</td>
<td>14</td>
<td>Minneapolis</td>
<td>0.013</td>
</tr>
<tr>
<td>6</td>
<td>Dallas-Fort Worth</td>
<td>0.012</td>
<td>15</td>
<td>New Orleans</td>
<td>0.019</td>
</tr>
<tr>
<td>7</td>
<td>Denver</td>
<td>0.015</td>
<td>16</td>
<td>New York</td>
<td>0.050</td>
</tr>
<tr>
<td>8</td>
<td>Detroit</td>
<td>0.035</td>
<td>17</td>
<td>Philadelphia</td>
<td>0.024</td>
</tr>
</tbody>
</table>

Disruption probabilities of potential hubs in reliable model.
A.4.2. The probabilities involving three nodes $P(D_{m} = 1, D_{n} = 0, D_{k} = 0)$ and $P(D_{k} = 0, D_{m} = 1, D_{n} = 0)$

For any different nodes $k, m,$ and $n$. We need to make further assumptions: fix $D_{k} = 1$, assume a new probability

$$
P(D_{m} = 0, D_{n} = 0| D_{k} = 1) = \frac{\exp(-\lambda_{mn})}{10}.
$$

Then, $P(D_{m} = 0, D_{n} = 0, D_{k} = 1) = P(D_{m} = 0, D_{n} = 0| D_{k} = 1)P(D_{k} = 1)$. $P(D_{k} = 0, D_{m} = 1, D_{n} = 0)$ can be obtained.

References


Flylowlcoastairlines.com, 2012. Airport Hamburg. HAM.


