# Accepted Manuscript


## Multiple Allocation Hub Interdiction and Protection Problems: Model Formulations and Solution Approaches

Prasanna Ramamoorthy, Sachin Jayaswal, Ankur Sinha, Navneet Vidyarthi

PII: S0377-2217(18)30259-5
DOI: 10.1016/j.ejor.2018.03.031
Reference: EOR 15050

To appear in: *European Journal of Operational Research*

Received date: 16 February 2017
Revised date: 23 January 2018
Accepted date: 23 March 2018

Please cite this article as: Prasanna Ramamoorthy, Sachin Jayaswal, Ankur Sinha, Navneet Vidyarthi, Multiple Allocation Hub Interdiction and Protection Problems: Model Formulations and Solution Approaches, *European Journal of Operational Research* (2018), doi: 10.1016/j.ejor.2018.03.031

**Highlights**

- This paper studies a bilevel Hub Interdiction and trilevel Hub Protection Problem.
- We study efficient methods to reduce the bilevel problem to single level.
- We present different closest assignment constraints to enable this reduction.
- We propose a Benders Decomposition method that solves large interdiction problems.
- Hub Protection Problem is solved using an Implicit Enumeration algorithm.

# Multiple Allocation Hub Interdiction and Protection Problems: Model Formulations and Solution Approaches

Prasanna Ramamoorthy[a,*], Sachin Jayaswal[a], Ankur Sinha[a], Navneet Vidyarthi[b]

[a] *Production and Quantitative Methods, Indian Institute of Management, Vastrapur, Ahmedabad, Gujarat, 380 015, India*
[b] *John Molson School of Business and CIRRELT, Concordia University, Montreal, QC, H3G 1M8, Canada.*

## Abstract

In this paper, we present computationally efficient formulations for the multiple allocation hub interdiction and hub protection problems, which are bilevel and trilevel mixed integer linear programs, respectively. In the hub interdiction problem, the aim is to identify a subset of $r$ critical hubs from an existing set of $p$ hubs that when interdicted results in the maximum post-interdiction cost of routing flows. We present two alternate ways of reducing the bilevel hub interdiction model to a single level optimization problem. The first approach uses the dual formulation of the lower level problem. The second approach exploits the structure of the lower level problem to replace it by a set of closest assignment constraints (CACs). We present alternate sets of CACs, study their dominance relationships, and report their computational performances. Further, we propose refinements to CACs that offer computational advantages of an order-of-magnitude compared to the one existing in the literature. Further, our proposed modifications offer structural advantages for Benders decomposition, which lead to substantial computational savings, particularly for large problem instances. Finally, we study and solve large scale instances of the trilevel hub protection problem exactly by utilizing the ideas developed for the hub interdiction problem.

*Keywords:* Location, Hub-and-Spoke network, Interdiction, Protection, Benders Decomposition.

## 1. Introduction

Certain infrastructural assets are critical to the functioning of a nation's economy and societal well being. The United States' Department of Homeland Security[1] identifies sixteen infrastructural sectors as critical, such that their incapacitation or destruction can be debilitating to the national security, economy, and public health (Brown et al., 2006). Three out of these sixteen critical infrastructure sectors, namely transportation systems, communications networks, and energy, employ hub-and-spoke as a dominant network structure because of its operational advantage. Hub-and-spoke networks exploit the economies of scale arising from consolidating the traffic from different origins and/or those destined to different demand points, instead of serving each origin-destination (O-D) pair directly. Flows from the same origin with different destinations in a hub-and-spoke network are consolidated on their route at the hub where they are combined with flows that have different origins but the same destination (Campbell, 1996). In multi-hub networks, traffic concentrated at a hub is directed to a second hub,

*Corresponding Author, Phone: +001-514-848-2424x2990, Fax: +001-514-848-2824
*Email addresses:* prasanna@iima.ac.in (Prasanna Ramamoorthy), sachin@iim.ac.in (Sachin Jayaswal), asinha@iima.ac.in (Ankur Sinha), n.vidyarthi@concordia.ca (Navneet Vidyarthi)

[1] https://www.dhs.gov/what-critical-infrastructure

which distributes it to the final destinations, thereby exploiting the economies of scale on the inter-hub flows. Another advantage of a a hub-and-spoke network is that it results in fewer links, which makes the network construction economical and its maintenance easier, compared to an alternate network with direct connections between all origins and destinations.

Hub-and-spoke networks have been well studied, starting with the seminal paper by O'Kelly (1986). Campbell (1994) presented integer programming formulations for the $p$-hub median, uncapacitated hub location, $p$-hub center, and hub covering problems. These models are largely inspired by their facility location counterparts, namely $p$-median, fixed charge facility location, $p$-center, and maximal covering problems. Since then, the following variants of hub location problems have been studied in the literature: single allocation (i.e. a non-hub is allocated to only one hub) or multiple allocation (i.e. a non-hub is allocated to one or more than one hub), uncapacitated (no limit on hub capacity) or capacitated (hubs have a fixed capacity). Ernst and Krishnamoorthy (1996), Skorin-Kapov et al. (1996), Ebery et al. (2000), Hamacher et al. (2004) are some of the important works in this area. In recent years, several newer variations of the hub location problems have appeared in the literature. Notable among those are: hub location with congestion (Elhedhli and Hu, 2005); hub location with service level constraints (Jayaswal and Vidyarthi, 2013); cycle hub location problem (Contreras et al., 2016); tree of hubs location problem (Contreras et al., 2010); hub location with flow dependent economies of scale (O'Kelly and Bryan, 1998); hub location with stochastic demands (Contreras et al., 2011b); modular hub location problem (Tanash et al., 2017); and dynamic hub location (Contreras et al., 2011c). Reviews of the hub location literature can be found in: Alumur and Kara (2008), Campbell and O'Kelly (2012) and Farahani et al. (2013).

While hub-and-spoke network structure is attractive due to its cost effectiveness, it is prone to severe disruptions in the event of a failure of any of its hubs due to either random events or deliberate attacks (called interdiction). For example, the recent (January 2017) snowstorm, named Egon, that hit continental Europe caused the international hubs at Heathrow and Frankfurt to close, severely disrupting airline operations[2]. O'Kelly (2015) points out that hub-and-spoke networks exhibit a non-random pattern of node degree, with some hub nodes exhibiting very high connectivity, while many others connected to very few other nodes. An attack at a node chosen at random would most likely do little damage due to the preponderance of low degree nodes, but a deliberate effort to disable (interdict) one of the very high degree nodes (hubs) could be devastating (Albert et al., 2002). A study states that it is possible to disrupt the entire United States' air network by interdicting just 2% of its all airports (Lewis, 2006). Thus, it becomes necessary to identify such critical hubs in advance so that resources may be deployed for their protection to minimize disruptions.

In this paper, we present optimization models to identify vulnerability and build resilience in hub-and-spoke networks to interdiction or extreme incidents. More specifically, the first model studied is a hub interdiction problem (HIP) that identifies critical hubs in hub-and-spoke networks. The second model, called the hub protection problem (HPP), allocates protective resources among critical hubs so that they can be fortified against interdiction. Several classes of HIPs and HPPs may arise depending on the settings of the problem and the underlying assumptions of the model. In this paper, we study the $r$-hub median interdiction problem ($r$-HMIP) and $u$-hub median protection problem ($u$-HMPP) in a multiple allocation hub-and-spoke network with $p$ hubs. $r$-HMIP aims to identify a subset of $r$ critical hubs from an existing set of $p$ hubs

[2] http://www.dw.com/en/storm-egon-brings-chaotic-winter-weather-to-europe/a-37120809

in a hub-and-spoke network, which when interdicted leads to maximal disruption to the system. In $u$-HMPP, the decision maker seeks to fortify/protect a subset of $u$ hubs from an existing set of $p$ hubs against interdiction. We present efficient formulations and solution approaches for the two models. More specifically, we present a bilevel mixed-integer programming (MIP) model for $r$-HMIP, followed by two alternate ways of reducing it to a single-level optimization problem. The first approach uses the dual formulation of the lower level min-cost hub routing problem to combine it with the upper level. The second approach exploits the fact that in hub-and-spoke networks satisfying (i) triangular inequality between every pair of nodes, and (ii) identical economies of scale (represented by a common discount factor) on all inter-hub links, a path between any origin-destination (O-D) pair can have at most two hubs Campbell and O'Kelly (2012). Hence, the total number of paths available in the network is polynomial, which can be enumerated. Therefore, the lower level problem can be replaced by closest assignment constraints (CACs), which ensure that the flow between any O-D pair happens through the least cost path among the polynomial number of available paths post-interdiction. We present alternate sets of CACs and study their dominance relationships. Our proposed refinements to CACs offer computational advantages of order-of-magnitude compared to the one existing in the literature. Further, our CACs offer structural advantages that are explored while applying Benders decomposition to efficiently solve large instances of $r$-HMIP. Benders decomposition offers further computational advantage of orders-of-magnitude over the direct solution of the single-level $r$-HMIP with CACs. Finally, we present a trilevel MIP formulation of $u$-HMPP, and reduce it to bilevel MIP using the proposed CACs. We also present an implicit enumeration algorithm in combination with Benders decomposition for $u$-HMPP. The computational advantage gained for $r$-HMIP by using CACs and Benders decomposition allowed us to further solve large instances of an otherwise intractable $u$-HMPP.

The major contributions of the paper are as follows:

- We present alternate ways of reducing the bilevel $r$-HMIP to single level using different sets of CACs that are more efficient than the one existing in the literature.
- We further present Benders decomposition for the different single-level formulations to efficiently solve large instances of $r$-HMIP.
- Further, using the above contributions, we solve large instances of $u$-HMPP using a combination of implicit enumeration and benders decomposition procedure.

The remainder of the paper is organized as follows. In Section 2, we present a brief review of literature on network interdiction and protection problems. Section 3 describes the hub interdiction problem ($r$-HMIP), and presents its bilevel formulation, followed by two alternate ways of reducing it to single level. Subsection 3.2 presents the single-level reduction by taking dual of the lower level routing problem, whereas Subsection 3.3 presents the single-level reduction using three alternative sets of CACs. The dominance relationships between CACs is described in Subsection 3.4. We further present two reduced formulations of CACs in Subsection 3.5. The computational comparisons of all the single-level reformulations of $r$-HMIP are reported in Section 4. Section 5 describes Benders decomposition for the different reformulations of $r$-HMIP. In Section 6, we present the trilevel model for $u$-HMPP, followed by a solution methodology using implicit enumeration algorithm. Conclusions and some future research directions are outlined in Section 7.

## 2. Literature Review

Interdiction refers to the forbidding or halting of an intelligent adversary's activity through an intentional attack. Interdiction problems involve two players, namely an attacker (also called interdictor) and a defender (also called evader). Such problems have been widely studied with respect to network flows (network interdiction) and facility location-allocation (facility interdiction) problems. The decision maker in an interdiction problem is interested in identifying the set of nodes/arcs (in network interdiction) or facilities (in facility interdiction) that when interdicted causes the maximum disruption/loss to the system. The problem is modeled as a Stackleberg game in which the *attacker* is the leader and the *defender* is the follower. The literature on interdiction can be broadly categorized into three groups: network interdiction, facility interdiction, and hub interdiction.

### *2.1. Network Interdiction*

Network interdiction problems identify critical nodes or arcs in a network. The defender operates on the network to optimize her objective such as: (i) to pass through the network as fast as possible (shortest path network interdiction) (Corley and Sha, 1982; Israeli and Wood, 2002; Cappanera and Scaparra, 2011); (ii) to move through the network without getting caught (most reliable path interdiction) (Morton et al., 2007); (iii) to maximize the amount of flow passing through the network (maximum flow network interdiction) (Wood, 1993; Cormican et al., 1998). The objectives of the attacker in these models are: (i) to intercept or destroy the arc(s)/node(s) so as to maximize the length of the shortest path; (ii) to maximize the probability of detection in the network; (iii) to minimize the maximum flow in the network. These models find applications in disrupting enemy flows (McMasters and Mustin, 1970), infectious disease control (Assimakopoulos, 1987), counter-terrorism (Farley, 2003), interception of nuclear material (Pan and Morton, 2008; Gutfraind et al., 2009), and contraband smuggling (Washburn and Wood, 1995). A review of network interdiction models with applications can be found in Collado and Papp (2012).

In one of the early papers in this area, Wood (1993) presented a network interdiction problem in which an enemy attempts to maximize flow through a capacitated network, while an interdictor tries to minimize this maximum flow by interdicting arcs using limited resources. He presented integer programming formulations for the discrete interdiction case, and presented valid inequalities and derived a reformulation to tighten the LP relaxation of some of these models. Extension of the model to allow for continuous interdiction, multiple origins and destinations, undirected networks, multiple interdiction resources, and multiple commodities are also described. Smith et al. (2007) presented a three-stage model for designing a survivable network under several interdiction scenarios. In particular, they examine the case in which an enemy, subject to some interdiction budget, can destroy any portion of any arc that a designer constructs on the network. The problem is modeled using a two-player game, in which the designer acts first to construct a network and transmit an initial set of flows through the network. The enemy acts next to destroy a set of constructed arcs in the designer's network followed by the designer who acts last to transmit a final set of flows in the network. They present solution approaches for three different profiles of enemy action: (i) based on arc capacities, (ii) based on initial flows, and (iii) interdiction to minimize the network designer's maximum profit obtained from transmitting flows. Lim and Smith (2007) studied the multicommodity flow network interdiction problem, in which an attacker disables, subject to an interdiction budget, a set of network arcs in order to minimize the maximum profit that can be obtained from shipping commodities in the network. The authors examine the problem under: (i) discrete interdiction (i.e., an interdicted arc is completely disabled), and (ii) continuous interdiction (an interdicted arc is not completely destroyed, but operates with reduced capacity).

### *2.2. Facility Interdiction and Protection*

Facility interdiction problems identify critical facilities in a supply network, which when destroyed causes maximum disruption. Church et al. (2004) proposed $r$-interdiction median problem (r-IMP) and $r$-interdiction covering problem (r-ICP) to study interdiction of facilities under different location scenarios. The r-IMP identifies a subset of $r$ facilities to remove from an existing set of $p$ facilities so as to maximize the overall transportation cost of serving customers from the remaining facilities post-interdiction. On the other hand, r-ICP identifies a subset of $r$ facilities from an exisiting set of $p$ facilities that when removed minimizes the total demand that can be covered within a specific distance or time. Different variants of r-IMP have been studied in the literature. Church and Scaparra (2007a) studied an extension of the problem where the success of the attack is uncertain. The authors assumed that the attacks are successful with a given probability. Losada et al. (2012) studied another type of uncertainty in r-IMP, which is the uncertainty in the degree of impact from an attack. This problem identifies the disruption scenario that result in the maximum overall transportation distance for serving all customers in the system. A key assumption here is that the degree of interdiction impact on a facility is proportional to the amount of reorigins employed. The problems described above assume no restrictions on the capacity of the facilities. Aksen et al. (2014) studied a partial interdiction of capacitated $r$-IMP, wherein facilities operate with a reduced capacity post-interdiction. Though various versions of $r$-IMP are studied (capacitated and uncapacitated, partial and full interdiction), their $r$-ICP counterparts have received limited attention in the literature.

Church and Scaparra (2007b) studied an extension of $r$-IMP, known as $r$-interdiction median problem with fortification ($r$-IMF), which identifies optimal fortification/protection strategies against interdiction. The model assumes that a protected/fortified facility becomes completely immune to attacks. Scaparra and Church (2008a) formulated $r$-IMF as a bilevel MIP, which is solved using an implicit enumeration algorithm. Scaparra and Church (2008b) proposed an alternate method for $r$-IMF based on its reformulation as a single level maximal covering problem with precedence constraints. The authors devise an approximate heuristic to identify the upper and lower bound to the problem, which are used to reduce the size of the original problem. This reduced problem is then solved to optimality using a commercial MIP solver. Aksen et al. (2010), Losada et al. (2010), Scaparra and Church (2012), Liberatore et al. (2012), Aksen and Aras (2012), and Aksen et al. (2013) are other related works in this area.

### *2.3. Hub Interdiction and Protection*

Interdiction of hubs in a hub-and-spoke network has received scarce attention in the literature, despite its many useful applications, as discussed in Section 1. However, there have been a few studies in closely related areas. An et al. (2015) and Azizi et al. (2016), for example, studied the reliable hub-and-spoke network design problem, which includes the possibility of re-routing flows through backup hubs when the active hubs are disrupted. However, the objective in both these papers is to minimize the weighted sum of pre-disruption and the expected value (over all disruption scenarios) of post-disruption transportation cost. Chaharsooghi et al. (2017) studied the reliable uncapacitated multiple allocation hub location problem under hub disruptions. Similar to Azizi et al. (2016), they assume that customers originally assigned to a disrupted hubs are either reassigned to other surviving hubs or they do not receive service, in which case a penalty should be paid. The problem is modeled as a two-stage stochastic program, and a metaheuristic algorithm based on the adaptive large neighborhood search is proposed. HIPs by contrast, study the worst-case loss to the defender.

In terms of the problem studied in this paper, Lei (2013), Parvaresh et al. (2014), Ghaffarinasab and Motallebzadeh (2017), and Ghaffarinasab and Atayi (2017) are the closest to our work. Lei (2013) presented a bilevel and a trilevel formulation for $r$-HMIP and $u$-HMPP, respectively. However, due to the complexity of the problem, computational results are presented only for small instances of $r$-HMIP, and none for $u$-HMPP. Parvaresh et al. (2014) presented two multi-objective metaheuristics based on simulated annealing and tabu search to solve the problem. More recently, Ghaffarinasab and Motallebzadeh (2017) studied $r$-HMIP, along with $r$-hub interdiction maximal covering and $r$-hub interdiction center problems, and solved them using simulated annealing techniques. Ghaffarinasab and Atayi (2017) presented an implicit enumeration algorithm for $r$-HMIP and a two level implicit enumeration algorithm for $u$-HMPP (one level of implicit enumeration for interdiction and another level for protection). Our paper differs from these papers essentially in terms of the solution method to efficently solve large instances of $r$-HMIP and $u$-HMPP. In contrast to Ghaffarinasab and Atayi (2017), which used implicit enumeration to solve the bilevel $r$-HMIP, we present alternate ways to reduce the problem to single level, and to solve them efficiently using Benders decomposition. Ghaffarinasab and Motallebzadeh (2017), similar to us, used alternate sets of CACs to reduce the bilevel $r$-HMIP to a single level. However, we further study in detail the dominance relations among the various CACs, and also present their reduced versions, which make the resulting single level formulation computationally very efficient. We further exploit the structure of the resulting alternate single level formulations of $r$-HMIP in Benders decomposition to further reduce the computational times. Benders decomposition for $r$-HMIP also allows us to further improve the efficiency of the implicit enumeration used for solving $u$-HMPP.

## 3. Problem Description and Model Formulation for r-HMIP

Consider a multiple allocation hub-and-spoke network with a set $\mathrm{H} \subseteq \mathrm{N}$ of $p$ hubs. Suppose that the follower (defender) has a set of flows $(W_{ij})$ between every origin node $i \in N$ and destination node $j \in N$, which is routed through one or at most two of the hubs from the set $H$. Let $d_{ijkm}$ represent the cost per unit flow from the origin $i$ to destination $j$, through hubs $k$ and $m$, in that order. Then, $d_{ijkm} = \alpha c_{ik} + \delta c_{km} + \gamma c_{mj}$, where $\alpha$, $\delta$, and $\gamma$ are the discount factors on collection, transhipment, and distribution links, respectively and $c_{ik}$, $c_{km}$, and $c_{mj}$ represent the cost of traversing from node $i$ to $k$, $k$ to $m$, and $m$ to $j$, respectively. Typically, $\delta < \alpha$ and $\delta < \gamma$ due to economies of scale arising from consolidation of flows on inter-hub links.

We model $r$-HMIP as a Stackleberg game in which the leader (attacker) makes the first move by interdicting a subset of $r$ hubs from the existing set $H$ of $p$ hubs with the objective to maximize the follower's (defender's) optimal routing/transportation cost through the $p - r$ surviving hubs in the network post-interdiction. We assume $r < p$ since the attacker usually has limited resources to interdict the hubs. We also assume that an interdicted hub is completely disabled, i.e., partial flows through an interdicted hub is not permitted. We formulate this game as a bilevel MIP. The hierarchical structure of the problem is shown in Figure 1.

### *3.1. Bilevel Programming Formulation*

In this subsection, we provide a mathematical formulation for $r$-HMIP. To begin with, we introduce the following notation.

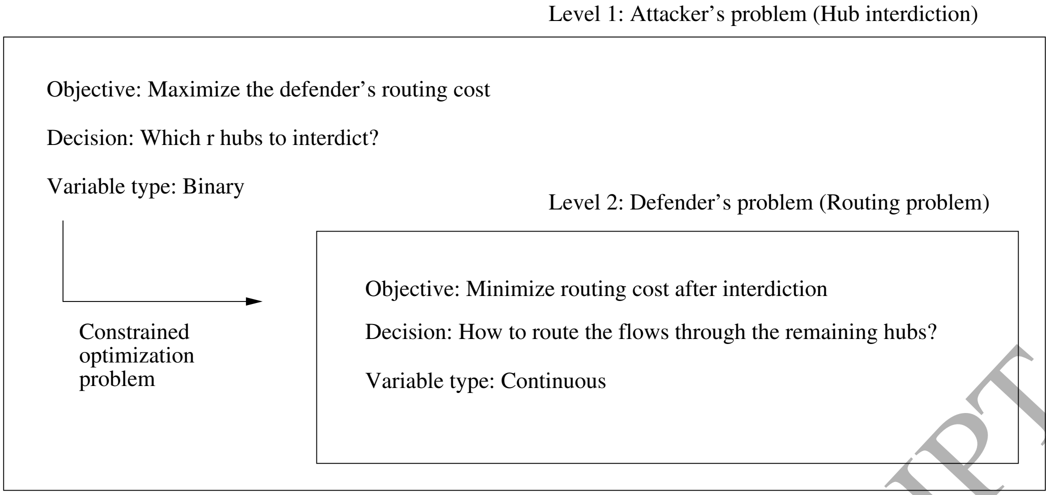


Figure 1: Hub interdiction problem as a bilevel MIP

*Indices and Parameters:*

| | | |
|---|---|---|
| $N$ | : | Set of all nodes |
| $H$ | : | Set of all hubs, $H \subseteq N$ |
| $i$ | : | Index for origin nodes, $i \in N$ |
| $j$ | : | Index for destination nodes, $j \in N$ |
| $k$ | : | Index for hub which is connected to origin nodes, $k \in H$ |
| $m$ | : | Index for hub which is connected to destination nodes, $m \in H$ |
| $\alpha$ | ; | Discount factor for collection (origin to hub), $(i \to k)$ |
| $\delta$ | ; | Discount factor for transhipment (hub to hub), $(k \to m)$ |
| $\gamma$ | ; | Discount factor for distribution (hub to destination), $(m \to j)$ |
| $W_{ij}$ | : | Demand (of flow) from origin $i$ to destination $j$ |
| $c_{ij}$ | : | Cost of traversing from node $i$ to $j$ |
| $d_{ijkm}$ | : | Cost of traversing from the origin $i$ to destination $j$, through hubs $k$ and $m$; $d_{ijkm} = \alpha c_{ik} + \delta c_{km} + \gamma c_{mj}$ |
| $p$ | : | No. of open hubs in the system |
| $r$ | : | No. of hubs to interdict. |

*Decision Variables:*

| | | |
|---|---|---|
| $X_{ijkm}$ | : | Fraction of flows from origin $i$ to destination $j$ through hubs $k$ and $m$ post-interdiction |
| $z_k$ | : | 1 if hub $k$ survives interdiction (is not interdicted), 0 otherwise |

With the above notation, the bilevel formulation of the multiple allocation $r$-HMIP can be mathematically stated as follows:

$$[r\text{-HMIP}_{2L}] : \max_{\mathbf{z}} \; T \tag{1}$$

$$\text{s.t.} \sum_{k \in H} z_k = p - r \tag{2}$$

$$z_k \in \{0, 1\} \qquad \forall k \in H \tag{3}$$

$$T = \min_{\mathbf{X}} \sum_{i \in N} \sum_{j \in N} \sum_{k \in H} \sum_{m \in H} W_{ij} d_{ijkm} X_{ijkm} \tag{4}$$

$$\text{s.t.} \sum_{k \in H} \sum_{m \in H} X_{ijkm} = 1 \qquad \forall i, j \in N \tag{5}$$

$$\sum_{m \in H} X_{ijkm} + \sum_{m \in H \setminus \{k\}} X_{ijmk} \leqslant z_k \qquad \forall i, j \in N; k \in H \tag{6}$$

$$X_{ijkm} \geqslant 0 \qquad \forall i, j \in N; k, m \in H \tag{7}$$

The leader's (attacker's) objective function (1) maximizes the defender's optimal total transportation cost post-interdiction, which the follower wants to minimize in its objective function (4). Constraint (2) ensures that $p - r$ hubs remain open post-interdiction. (4) to (7) form the follower's problem at the lower level. Constraint set (5) ensures that the demand between every O-D pair $(i, j)$ is satisfied using paths containing at most two hubs, while constraint set (6) ensures that this demand is routed only via surviving hubs post-interdiction. Constraint set (6) can be alternatively represented by the following two sets of constraints, as done by Lei (2013).

$$\sum_{k \in H} X_{ijkm} \leqslant z_m \qquad \forall\, i, j \in N; m \in H;$$

$$\sum_{m \in H} X_{ijkm} \leqslant z_k \qquad \forall\, i, j \in N; k \in H.$$

However, the constraint set of the form (6) has been proven to be facet defining (Hamacher et al., 2004). Hence, constraint set (6) provides a tighter linear programming (LP) relaxation, which is effective in solving large instances of $r$-HMIP. Note that the lower level problem in the above formulation is a linear program (LP) as a result of multiple allocation of non-hub nodes to hubs.

Bilevel optimization problems, even with linear programs at both levels, are known to be NP-hard (Frangioni, 1995; Audet et al., 1997). As such, they are traditionally solved by reducing the problem to single-level using various reduction techniques (Sinha et al., 2017). In this paper, we present two alternate ways of reducing the bilevel $r$-HMIP to a single-level MIP such that it is tractable. The first approach is based on the dual of the lower level LP, while the second approach exploits the structure of the solution to the lower level problem so that it can be replaced by CACs.

### *3.2. Single-level Reduction using Dual Formulation*

In the bilevel formulation $r$-HMIP$_{2L}$, the lower level problem given by (4)-(7), is an LP. Further, since the objective functions at the two levels are the same, while their objectives are exactly opposite (maximization for the upper level and minimization for the lower level), $r$-HMIP$_{2L}$ can be reduced to a single-level MIP by taking the dual of the lower level LP. For a fixed upper level variable $\bar{z}_k$, associating dual variables $\phi_{ij}$ and $\delta_{ijk}$ with constraint sets (5) and (6), respectively we get the following dual LP of the lower level problem:

$$\max_{\boldsymbol{\phi}, \boldsymbol{\delta}} \sum_{i \in N} \sum_{j \in N} \phi_{ij} - \sum_{i \in N} \sum_{j \in N} \sum_{k \in H} \delta_{ijk} \bar{z}_k$$

$$\text{s.t.} \quad \phi_{ij} - \delta_{ijk} \leqslant W_{ij} d_{ijkm} \qquad \forall i, j \in N; \forall k, m \in H, k = m$$

$$\phi_{ij} - \delta_{ijk} - \delta_{ijm} \leqslant W_{ij} d_{ijkm} \qquad \forall i, j \in N; \forall k, m \in H, k \neq m$$

$$\phi_{ij} \text{ unbounded}; \quad \delta_{ijk} \geqslant 0 \qquad \forall i, j \in N; \forall k \in H$$

Replacing the lower level problem with the above dual gives the following single level nonlinear MIP reformulation for $r$-HMIP:

$$\max_{\mathbf{z},\boldsymbol{\phi},\boldsymbol{\delta}} \sum_i \sum_j \phi_{ij} - \sum_i \sum_j \sum_k \delta_{ijk} z_k \tag{8}$$

$$\text{s.t.} \sum_{k \in H} z_k = p - r \tag{9}$$

$$\phi_{ij} - \delta_{ijk} \leqslant W_{ij} d_{ijkm} \qquad \forall i, j \in N; \forall k, m \in H, k = m \tag{10}$$

$$\phi_{ij} - \delta_{ijk} - \delta_{ijm} \leqslant W_{ij} d_{ijkm} \qquad \forall i, j \in N; \forall k, m \in H, k \neq m \tag{11}$$

$$\phi_{ij} \text{ unbounded}; \quad \delta_{ijk} \geqslant 0; \quad z_k \in \{0, 1\} \qquad \forall i, j \in N; \forall k \in H \tag{12}$$

Model (8)-(12) has bilinear terms $\sum_i \sum_j \delta_{ijk} z_k$ in its objective function (8), which can be linearized by introducing a new variable $V_k$ and the following two constraint sets:

$$V_k \leqslant M_k z_k \qquad \forall k \in H \tag{13}$$

$$V_k \geqslant \sum_{i \in N} \sum_{j \in N} \delta_{ijk} - M_k(1 - z_k) \qquad \forall k \in H \tag{14}$$

where $M_k$ is a sufficiently large number. The linearized single-level dual formulation of $r$-HMIP is given by:

$$[r\text{-HMIP}_{DF}] : \max_{\mathbf{z},\boldsymbol{\phi},\boldsymbol{\delta},\mathbf{V}} \sum_{i \in N} \sum_{j \in N} \phi_{ij} - \sum_{k \in H'} V_k \tag{15}$$

$$\text{s.t. } (9) - (12)$$

$$(13), (14)$$

$$V_k \geqslant 0 \qquad \forall k \in H \tag{16}$$

The computational efficiency of the above formulation depends on the specific value of $M_k$ chosen for the problem. A good value of $M_k$ provides a tighter LP relaxation of the model. In the following proposition, we present a valid value of $M_k$.

**Proposition 1.** *For a given O-D pair $(i, j)$, let $d_{ijk_1m_1} = max_{k,m}\{d_{ijkm}\}$ and $d_{ijkm_2} = min_m\{d_{ijkm}\}$. Then, $\overline{M}_k = \sum_{i \in N} \sum_{j \in N} W_{ij} \left(d_{ijk_1m_1} - d_{ijkm_2}\right)$ is a valid value of $M_k$ for r-HMIP$_{DF}$.*

*Proof.*

$$(13) - (14) \implies M_k \geqslant \sum_{i \in N} \sum_{j \in N} \delta_{ijk}.$$

Also,

$$\delta_{ijk} \leqslant W_{ij} d_{ijk_1m_1} - W_{ij} d_{ijkm_2}$$

(since $\delta_{ijk}$ is the shadow price of constraint (6), which is obtained as the maximum possible change in the objective function (4) corresponding to a change in the right hand side of constraint (6) by a unit).
Hence, $\overline{M}_k = \sum_{i \in N} \sum_{j \in N} \delta_{ijk} = \sum_{i \in N} \sum_{j \in N} W_{ij} \left(d_{ijk_1m_1} - W_{ij} d_{ijkm_2}\right)$ is a valid value of $M_k$ for the $r$-HMIP$_{DF}$ formulation. □

*3.3. Single-level Reduction using CACs*

In this subsection, we present an alternate way of reducing the bilevel $r$-HMIP to a single level optimization problem by exploiting the structure of the lower level LP. For a given solution to the upper level problem (i.e., for fixed values of $z_k$), the lower level problem separates into an independent min-cost routing problem for every O-D pair $(i, j)$. The minimum cost routing requirement between every O-D pair $(i, j)$ can also be ensured through additional constraints in absence of the objective function, which allows the reduction of $r$-HMIP$_{2L}$ to single level. Such constraints have been widely used in facility location problems to allocate customers to their closest facilities, and are popularly called as CACs. Gerrard and Church (1996) and Espejo et al. (2012) compare different CACs used in location problems, and study their theoretical properties. These constraints find applications in hazardous facility location problems (Song et al., 2013), facility location problems under competition (Dobson and Karmarkar, 1987), and facility interdiction problems (Liberatore et al., 2011), among others.

In hub-and-spoke networks satisfying (i) triangular inequality between every pair of nodes, and (ii) identical economies of scale (represented by a common discount factor $\delta$) on all inter-hub links, a path between any O-D pair can have at most two hubs. Hence, the total number of paths available in the network is polynomial, which can be enumerated in CACs. We present three different sets of CACs to reduce $r$-HMIP$_{2L}$ to single level:

1. The first set of CACs is an extension of its counterpart for facility location problems proposed by Church and Cohon (1976). These CACs are used by Lei (2013) to convert the bilevel $r$-HMIP to single level. They are stated as follows:

$$\sum_{(q,s)\in C_{ijkm}} X_{ijqs} + X_{ijkm} \geqslant z_k + z_m - 1 \qquad \forall\, i, j \in N; k, m \in H \tag{CAC1a}$$

where $C_{ijkm} = \{(q, s)|\ d_{ijqs} < d_{ijkm} \text{ or } (d_{ijqs} = d_{ijkm} \text{ and } (q < k \text{ or } (q = k \text{ and } s < m)))\}$. For a given O-D pair $(i, j)$, CAC1a ensures that the flow between them is routed only through a path that is no costlier than the path $i \to k \to m \to j$ as long as hubs $k$ and $m$ are open. CAC1a arbitrarily breaks any tie between paths having the same cost. Breaking ties for $r$-HMIP is not necessary, unlike in facility location problems without which it becomes infeasible. Hence, we can rewrite CAC1a as follows:

$$\sum_{(q,s)\in \hat{C}_{ijkm}} X_{ijqs} + X_{ijkm} \geqslant z_k + z_m - 1 \qquad \forall\, i, j \in N; k, m \in H \tag{CAC1b}$$

where $\hat{C}_{ijkm} = \{(q, s)|\ d_{ijqs} < d_{ijkm} \text{ or } (d_{ijqs} = d_{ijkm} \text{ and } (q \neq k \text{ or } (q = k \text{ and } s \neq m)))\}$.

2. The second set of CACs forbids assignment of flows between any O-D pair $(i, j)$ to a path costlier than the path $i \to k \to m \to j$ as long as hubs $k$ and $m$ are open. It is written as follows:

$$\sum_{(q,s)\in E_{ijkm}} X_{ijqs} \leqslant 2 - z_k - z_m \qquad \forall\, i, j \in N; k, m \in H \tag{CAC2}$$

where $E_{ijkm} = \{(q, s)|\ d_{ijqs} > d_{ijkm}\}$. CAC2 is similar to the constraint proposed by Wagner and Falkson (1975) for $p$-center problems.

3. The third set of CACs, like CAC1b, ensures the flows between a given O-D pair $(i, j)$ is routed only through a path $i \to q \to s \to j$ that is no costlier than the path $i \to k \to$

$m \to j$ as long as the hubs $k$ and $m$ are open. This is given by:

$$\sum_{q \in H} \sum_{s \in H} d_{ijqs} X_{ijqs} + (M - d_{ijkm})(z_k + z_m - 1) \leqslant M \qquad \forall\, i, j \in N; k, m \in H \qquad \text{(CAC3)}$$

where $M = \max_{i,j \in N}\{\sum_{k \in H} \sum_{m \in H} d_{ijkm}\}$.
In the above inequality, by fixing $z_k$ and $z_m$ to 1, the allocations $X_{ijqs}$ will be through paths shorter than $d_{ijkm}$. CAC3 is an adaptation of the CAC used by Berman et al. (2009) for facility location problems.

The single-level reformulation of $r$-HMIP with CACs takes the following form:

$$\begin{aligned}[r\text{-HMIP}_{CAC}] : \max_{\mathbf{z},\mathbf{X}} & \sum_{i \in N} \sum_{j \in N} \sum_{k \in H} \sum_{m \in H} W_{ij} d_{ijkm} X_{ijkm} \\ \text{s.t. } & (2), (3), (5) - (7) \\ & CAC1b \quad \text{or} \quad CAC2 \quad \text{or} \quad CAC3\end{aligned}$$

Note that all the single-level reformulations of $r$-HMIP$_{CAC}$ that use CAC1 (CAC1b), CAC2, and CAC3 have the same number of variables as the original bilevel $r$-HMIP$_{2L}$. In order to find the most efficient CACs among the proposed CACs for reduction, we study the dominance relationships between them in the following subsection.

### *3.4. Dominance Relationship between CACs*

The dominance relationship between CACs has been studied in the context of facility location problems (Espejo et al., 2012). A constraint dominates another constraint if the former yields a tighter LP relaxation than the latter. To prove that one constraint dominates the other, one needs to show that the LP feasible region of the dominating constraint is a subset of the LP feasible region of the dominated constraint. For studying the dominance relationship between CACs, the dominance rules we use are as follows: (i) a constraint dominates the another constraint if the former implies the latter, but not the other way round; (ii) if both constraints imply one another, we say that the constraints are equivalent.

**Proposition 2.** *CAC2 is equivalent to CAC1b*

*Proof.* CAC2 can be written as:

$$\begin{aligned} 1 - \sum_{(q,s) | d_{ijqs} \leqslant d_{ijkm}} X_{ijqs} &\leqslant 2 - z_k - z_m \qquad && \forall\, i, j \in N; k, m \in H \\ \text{or} \quad \sum_{(q,s) | d_{ijqs} \leqslant d_{ijkm}} X_{ijqs} &\geqslant z_k + z_m - 1 && \forall\, i, j \in N; k, m \in H. \end{aligned}$$

By separating $X_{ijkm}$ term, the above inequality can be rewritten as follows:

$$\sum_{(q,s) \in \hat{C}_{ijkm}} X_{ijqs} + X_{ijkm} \geqslant z_k + z_m - 1 \qquad \forall\, i, j \in N; k, m \in H.$$

Hence, *CAC2* $\implies$ *CAC1b*. Similarly, one can prove that *CAC1b* $\implies$ *CAC2*.
Therefore, *CAC2* is equivalent to *CAC1b*. □

**Proposition 3.** *CAC2 dominates CAC3*

*Proof.* To prove CAC2 dominates CAC3, we show that CAC2 $\implies$ CAC3 while CAC3 $\centernot\implies$ CAC2.
We first prove CAC2 $\implies$ CAC3. For this we rewrite CAC2 as:

$$\sum_{(q,s)\in E_{ijkm}} X_{ijqs} + z_k + z_m \leqslant 2 \qquad \forall\, i,j \in N; k,m \in H$$

Hence, the following is a relaxation of CAC2:

$$X_{ijqs} + z_k + z_m \leqslant 2 \qquad \forall i,j \in N; k,m \in H; (q,s) \in E_{ijkm} \qquad (CAC2\text{-rel})$$

It is evident that $CAC2 \implies CAC2$-rel, while $CAC2$-rel $\centernot\implies CAC2$. Therefore, $CAC2$ dominates $CAC2$-rel. To show that $CAC2$ dominates $CAC3$, we need to prove that $CAC2$-rel either dominates or is equivalent to $CAC3$.

Multiplying both sides of CAC2-rel by $d_{ijqs}$ and summing over $(q,s) \in E_{ijkm}$, we get:

$$\sum_{(q,s)\in E_{ijkm}} d_{ijqs}X_{ijqs} + \sum_{(q,s)\in E_{ijkm}} d_{ijqs}z_k + \sum_{(q,s)\in E_{ijkm}} d_{ijqs}z_m \leqslant 2\sum_{(q,s)\in E_{ijkm}} d_{ijqs} \qquad \forall\, i,j \in N; k,m \in H.$$

Adding $\sum_{(q,s)|d_{ijqs}\leqslant d_{ijkm}} d_{ijqs}X_{ijqs} + (M - d_{ijkm} - \sum_{(q,s)\in E_{ijkm}} d_{ijqs})(z_k + z_m - 1)$ to both sides of the above inequality, where $M$ is $\max_{ij}\sum_{k,m} W_{ij}d_{ijkm}$, we get

$$\begin{aligned}\sum_{q\in H}\sum_{s\in H} d_{ijqs}X_{ijqs} + (M - d_{ijkm})(z_k + z_m - 1) \leqslant \sum_{(q,s)\in E_{ijkm}} d_{ijqs} + \sum_{(q,s)|d_{ijqs}\leqslant d_{ijkm}} d_{ijqs}X_{ijqs}\\ +(M - d_{ijkm})(z_k + z_m - 1) - \sum_{(q,s)\in E_{ijkm}} d_{ijqs}(z_k + z_m - 1) \quad \forall\, i,j \in N; k,m \in H.\end{aligned} \qquad (CAC2\text{-rel2})$$

In the above inequality, the maximum value of the RHS (which corresponds to $z_k$, $z_m = 1$) is always bounded by $M$ since $\sum_{(q,s)|d_{ijqs}\leqslant d_{ijkm}} d_{ijqs}X_{ijqs} \leqslant d_{ijkm}$. Therefore, we get the following inequality:

$$\sum_{q\in H}\sum_{s\in H} d_{ijqs}X_{ijqs} + (M - d_{ijkm})(z_k + z_m - 1) \leqslant M \qquad \forall\, i,j \in N; k,m \in H,$$

which proves that CAC2-rel $\implies$ CAC3. Since CAC2 dominates CAC2-rel, CAC2 dominates CAC3.
Next, we prove CAC3 $\centernot\implies$ CAC2. This follows immediately from the following relations we proved above:

CAC2 $\implies$ CAC2-rel $\implies$ CAC3
But,
CAC2-rel $\centernot\implies$ CAC2
Therefore,
CAC3 $\centernot\implies$ CAC2.

□

### *3.5. Reduced Formulations of CACs*

In the following subsection, we present refinements of CAC1 and CAC2 that lead to fewer constraints. These reduced formulations of CAC1 and CAC2 are based on constraint dominance principles.

**Proposition 4.** *For a given O-D pair $(i, j)$ and hubs $k$ and $m$ $(m \neq k)$ between them, $CAC1b_{ijkm}$ dominates $CAC1b_{ijmk}$ when $d_{ijkm} < d_{ijmk}$.*

*Proof.* To prove the above proposition, we show that $CAC1b_{ijkm} \implies CAC1b_{ijmk}$ while $CAC1b_{ijmk} \not\Rightarrow CAC1b_{ijkm}$ when $d_{ijkm} < d_{ijmk}$.
For this, we rewrite $CAC1b_{ijkm}$ and $CAC1b_{ijmk}$ as follows:

$$CAC1b_{ijkm}: \sum_{(q,s)\in \hat{C}_{ijkm}\cup\{(k,m)\}} X_{ijqs} - z_k - z_m \geqslant -1$$
$$CAC1b_{ijmk}: \sum_{(q,s)\in \hat{C}_{ijmk}\cup\{(m,k)\}} X_{ijqs} - z_k - z_m \geqslant -1.$$

Also, $\hat{C}_{ijkm} \cup \{(k,m)\} \subset \hat{C}_{ijmk} \cup \{(m,k)\}$ when $d_{ijkm} < d_{ijmk}$. Hence, $CAC1b_{ijkm} \implies CAC1b_{ijmk}$ while $CAC1b_{ijmk} \not\Rightarrow CAC1b_{ijkm}$ when $d_{ijmk} < d_{ijmk}$. □

Based on Proposition 4, we propose a new formulation for *CAC1*, which is given below. For this, we define a set $H'_{ij} = \{H''_{ijkm} | k, m \in H, k \leqslant m\}$ for each O-D pair $(i, j)$, where

$$H''_{ijkm} = \begin{cases} (k,m) \text{ if } d_{ijkm} \leqslant d_{ijmk} \\ (m,k) \text{ if } d_{ijkm} > d_{ijmk}. \end{cases}$$

The new CAC can then be written as follows:

$$\sum_{(q,s)\in \hat{C}_{ijkm}} X_{ijqs} + X_{ijkm} \geqslant z_k + z_m - 1 \qquad \forall\, i, j \in N; (k,m) \in H'_{ij} \tag{rCAC1}$$

The reduced constraint set rCAC1 has $|N|^2((p^2 + p)/2)$ constraints whereas, CAC1 has $|N|^2 p^2$ constraints.

**Proposition 5.** *For a given O-D pair $(i, j)$ and hubs $k$ and $m$ $(m \neq k)$ between them, $CAC2_{ijkm}$ dominates $CAC2_{ijmk}$ when $d_{ijkm} < d_{ijmk}$.*

*Proof.* To prove the above proposition, we show that $CAC2b_{ijkm} \implies CAC2b_{ijmk}$ while $CAC2b_{ijmk} \not\Rightarrow CAC2b_{ijkm}$ when $d_{ijkm} < d_{ijmk}$.
For this, we rewrite $CAC2b_{ijkm}$ and $CAC2b_{ijmk}$ as follows:

$$CAC2b_{ijkm}: \sum_{(q,s)\in E_{ijkm}} X_{ijqs} + z_k + z_m \leqslant 2$$
$$CAC2b_{ijmk}: \sum_{(q,s)\in E_{ijmk}} X_{ijqs} + z_k + z_m \leqslant 2.$$

Also, $E_{ijkm} \supset E_{ijmk}$ when $d_{ijkm} < d_{ijmk}$. Hence, $CAC2b_{ijkm} \implies CAC2b_{ijmk}$ while $CAC2b_{ijmk} \not\Rightarrow CAC2b_{ijkm}$ when $d_{ijmk} < d_{ijmk}$. □

Based on Proposition 5, we propose a new formulation for CAC, which is given below:

$$\sum_{(q,s)\in E_{ijkm}} X_{ijqs} \leqslant 2 - z_k - z_m \qquad \forall\, i,j \in N; (k,m) \in H'_{ij}. \tag{rCAC2}$$

The reduced constraint set rCAC2 has $|N|^2((p^2+p)/2)$ constraints whereas, CAC2 has $|N|^2p^2$ constraints.

*3.6. Advantages of CAC2 over CAC1*

We outline the advantages of CAC2 over CAC1 at two stages of the solution process - the presolve stage and the branch-and-bound stage. CAC2 has certain structural properties that help in solving the single-level reformulation of HMIP faster than the one with CAC1. These properties are also valid for rCAC2 since it is a tighter version of CAC2.

**Advantages at presolve stage:** Presolve procedure is executed by a commercial solver prior to solving the optimization problem in order to reduce the size of the given problem by removing the redundant variables and constraints. Probing is a process that is carried out at the presolve step wherein logical consequences are investigated by setting the binary variables at their bounds (Savelsbergh, 1994). We show that CAC2 and rCAC2 together with constraint (6) eliminate a large number variables by probing.

**Proposition 6.** *For a given O-D pair (i,j) and hub k, $X_{ijkm}$ variables that appear common in (6) and $CAC2_{ijkk}$ can be set to zero.*

*Proof.* For a given O-D pair $(i,j)$ and hub $k$, consider the following possible two cases:

- $z_k = 0$: variables that appear in both constraint (6) and $CAC2_{ijkk}$ are set to zero by constraint (6).
- $z_k = 1$: variables that appear in both constraint (6) and $CAC2_{ijkk}$ are set to zero by $CAC2_{ijkk}$.

Since the variables that appear in both constraint 6 and CAC2 are set to zero in either case, they can be eliminated from the model prior to solving. □

Thus, CAC2 and rCAC2 eliminate a large number of variables by probing. Despite CAC1 being equivalent to CAC2, probing using the former is not straightforward. However, CAC1 also eliminates some variables at presolve stage, although not as many as CAC2, as evident later from computational results in Section4.

**Advantages at branch-and-bound stage:** In a branch-and-bound algorithm, the LP relaxation of the MIP problem is solved at the root node. Further, branching is done by setting the integer variables to its bounds that have taken a fractional value in the optimal solution to the relaxed problem. In our problem, branching is done by setting $z_k$ variables to zero and one. When a $z_k$ variable is set to one, some $X_{ijkm}$ variables are set to zero because of the $CAC2$ formulation. These variables can be eliminated from the model to reduce its size. Similarly, when $z_k$ is set to zero, some $X_{ijkm}$ variables are eliminated because of constraint (6), which again reduces the model size.

## 4. Computational Comparison of Formulations

We present the results of our computational experiments to compare the computational efficiencies of the different single level reformulations. For our experiments, we use instances derived from the Civil Aeronautics Board (CAB) dataset containing $|N| = 25$ and $p \in \{7, 10\}$. The set of $p$ hubs in the existing hub-and-spoke network for each of the instances is obtained by solving a corresponding uncapacitated $p$-hub median problem (Ebery et al., 2000). All the computational experiments are performed on a workstation with a 2.60GHz Intel Xeon - e5 processor and 24 GB RAM, and all the instances are solved using CPLEX 12.6.

In Table 1, we present the results for $r \in \{1, 2, 3, 4, 5, 6, 7, 8\}$. The discount factors for collection ($\alpha$) and distribution ($\gamma$) are both set at 1.0, while the discount factor for transshipment ($\delta$) is varied in the set $\{0.9, 0.5, 0.1\}$. The first set of columns lists the problem parameters: $|N|$, $p$, $r$ and $\delta$. The second set of columns reports the size (number of constraints and variables) and the CPU time for the single-level dual reformulation, $r$-HMIP$_{DF}$. We also report the size of the formulation (number of constraints and variables) before and after presolve operation for the single-level $r$-HMIP for different variants of CACs. The column "Original Size" reports the number of constraints and variables in the formulations *before* the presolve operation, which is the same for the three unreduced variants of CACs (CAC1, CAC2, CAC3). For example, for problem instances with $N = 25$ and $p = 7$, all the three formulations, $r$-HMIP$_{CAC1}$, $r$-HMIP$_{CAC2}$ and $r$-HMIP$_{CAC3}$, consist of 35,626 constraints and 30,633 variables. Similarly, for $N = 25$ and $p = 10$, the number of constraints and variables are 69,376 and 62,511, respectively. The number of constraints and variables *after* the presolve operation and the total CPU time (in seconds) to solve the problem optimally for each of the CAC formulations are reported under their respective columns "Cons.", "Vars." and "CPU(s)", respectively.

Table 1: Comparison between different single level reformulations of $r$-HMIP using CAB dataset

| $N$ | $p$ | $r$ | $\delta$ | $r$-HMIP$_{DF}$ | | | Original Size - CACs | | $r$-HMIP$_{CAC1}$ | | | $r$-HMIP$_{CAC2}$ | | | $r$-HMIP$_{CAC3}$ | | | $r$-HMIP$_{rCAC1}$ | | | $r$-HMIP$_{rCAC2}$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Cons. | Var. | CPU(s) | Cons | Var. | Cons. | Var. | CPU(s) | Cons. | Var. | CPU(s) | Cons. | Var. | CPU(s) | Cons. | Var. | CPU(s) | Cons. | Var. | CPU(s) |
| 25 | 7 | 1 | 0.9 | 4390 | 5015 | 7 | 35626 | 30633 | 9911 | 8619 | 4 | 1329 | 845 | 2 | 34986 | 30058 | 43 | 6611 | 8619 | 3 | 1329 | 845 | **2** |
| | | 2 | | | | 6 | | | 27593 | 23775 | 15 | 10383 | 4655 | 5 | 35094 | 30058 | 47 | 18377 | 23479 | 8 | 10365 | 4655 | **4** |
| | | 3 | | | | 7 | | | 27593 | 23775 | 17 | 10383 | 4655 | 6 | 35094 | 30058 | 69 | 18377 | 23479 | 12 | 10365 | 4655 | **5** |
| | | 4 | | | | 5 | | | 27593 | 23775 | 22 | 10383 | 4655 | 5 | 35094 | 30058 | 98 | 18377 | 23479 | 16 | 10365 | 4655 | **5** |
| | | 5 | | | | 2 | | | 27593 | 23775 | 13 | 10383 | 4655 | 3 | 35094 | 30058 | 83 | 18377 | 23479 | 10 | 10365 | 4655 | **2** |
| | | 1 | 0.5 | | | 9 | | | 19519 | 17015 | 7 | 3713 | 2405 | 3 | 34964 | 30058 | 46 | 12815 | 17015 | 5 | 3713 | 2405 | **3** |
| | | 2 | | | | 10 | | | 29809 | 25835 | 18 | 11220 | 6567 | 7 | 35094 | 30058 | 48 | 19645 | 25529 | 12 | 10904 | 6567 | **6** |
| | | 3 | | | | 10 | | | 29809 | 25835 | 30 | 11220 | 6567 | **7** | 35094 | 30058 | 51 | 19645 | 25529 | 19 | 10904 | 6567 | 8 |
| | | 4 | | | | 8 | | | 29809 | 25835 | 28 | 11220 | 6567 | 11 | 35094 | 30058 | 80 | 19645 | 25529 | 19 | 10904 | 6567 | **8** |
| | | 5 | | | | 5 | | | 29809 | 25835 | 9 | 11220 | 6567 | 3 | 35094 | 30058 | 88 | 19645 | 25529 | 6 | 7817 | 4922 | **3** |
| | | 1 | 0.1 | | | 12 | | | 26581 | 23197 | 11 | 6303 | 4143 | 3 | 34996 | 30058 | 47 | 17061 | 23197 | 7 | 6303 | 4143 | **3** |
| | | 2 | | | | 12 | | | 31579 | 27481 | 28 | 13176 | 8669 | 12 | 35094 | 30058 | 52 | 20280 | 27167 | 16 | 12956 | 8669 | **9** |
| | | 3 | | | | 13 | | | 31579 | 27481 | 43 | 13176 | 8669 | 10 | 35094 | 30058 | 70 | 20280 | 27167 | 25 | 12956 | 8669 | **10** |
| | | 4 | | | | 10 | | | 31579 | 27481 | 36 | 13176 | 8669 | **10** | 35094 | 30058 | 73 | 20280 | 27167 | 20 | 12956 | 8669 | 11 |
| | | 5 | | | | 7 | | | 31579 | 27481 | 8 | 13176 | 8669 | 3 | 35094 | 30058 | 38 | 20280 | 27167 | 5 | 12956 | 8669 | **3** |
| | 10 | 1 | 0.9 | 6271 | 6896 | 13 | 69376 | 62511 | 28475 | 25786 | 14 | 1743 | 1102 | 7 | 68135 | 61285 | 90 | 17441 | 25786 | 9 | 1743 | 1102 | **7** |
| | | 2 | | | | 14 | | | 58805 | 53083 | 48 | 14045 | 6741 | 12 | 68241 | 61285 | 163 | 35957 | 52397 | 25 | 14027 | 6741 | **11** |
| | | 3 | | | | 10 | | | 58805 | 53083 | 61 | 14045 | 6741 | 13 | 68241 | 61285 | 393 | 35957 | 52397 | 34 | 14027 | 6741 | **12** |
| | | 4 | | | | 21 | | | 58805 | 53083 | 74 | 14045 | 6741 | 16 | 68241 | 61285 | 425 | 35957 | 52397 | 45 | 14027 | 6741 | **13** |
| | | 5 | | | | 22 | | | 58805 | 53083 | 67 | 14045 | 6741 | 15 | 68241 | 61285 | 520 | 35957 | 52397 | 40 | 14027 | 6741 | **13** |
| | | 6 | | | | 16 | | | 58805 | 53083 | 82 | 14045 | 6741 | 12 | 68241 | 61285 | 659 | 35957 | 52397 | 44 | 14027 | 6741 | **12** |
| | | 7 | | | | 15 | | | 58805 | 53083 | 96 | 14045 | 6741 | 14 | 68241 | 61285 | 2504 | 35957 | 52397 | 33 | 14027 | 6741 | **11** |
| | | 8 | | | | 12 | | | 58805 | 53083 | 49 | 14045 | 6741 | 8 | 68241 | 61285 | 914 | 35957 | 52397 | 40 | 14027 | 6741 | **7** |
| | | 1 | 0.5 | | | 17 | | | 45559 | 41340 | 22 | 5301 | 3430 | 8 | 68099 | 61285 | 83 | 27655 | 41340 | 14 | 5301 | 3430 | **7** |
| | | 2 | | | | 29 | | | 62389 | 56487 | 64 | 17091 | 9895 | 18 | 68241 | 61285 | 150 | 37917 | 55765 | 33 | 16807 | 9895 | **17** |
| | | 3 | | | | 25 | | | 62389 | 56487 | 84 | 17091 | 9895 | 22 | 68241 | 61285 | 288 | 37917 | 55765 | 52 | 16807 | 9895 | **19** |
| | | 4 | | | | 42 | | | 62389 | 56487 | 115 | 17091 | 9895 | 25 | 68241 | 61285 | 325 | 37917 | 55765 | 58 | 16807 | 9895 | **23** |
| | | 5 | | | | 32 | | | 62389 | 56487 | 104 | 17091 | 9895 | 21 | 68241 | 61285 | 572 | 37917 | 55765 | 64 | 16807 | 9895 | **20** |
| | | 6 | | | | 23 | | | 62389 | 56487 | 109 | 17091 | 9895 | **19** | 68241 | 61285 | 526 | 37917 | 55765 | 59 | 16807 | 9895 | 20 |
| | | 7 | | | | 17 | | | 62389 | 56487 | 115 | 17091 | 9895 | 18 | 68241 | 61285 | 738 | 37917 | 55765 | 60 | 16807 | 9895 | **18** |
| | | 8 | | | | 15 | | | 62389 | 56487 | 47 | 17091 | 9895 | 9 | 68241 | 61285 | 1198 | 37917 | 55765 | 28 | 16807 | 9895 | **8** |
| | | 1 | 0.1 | | | 27 | | | 56371 | 51118 | 36 | 9639 | 6468 | 9 | 68203 | 61285 | 96 | 34049 | 51118 | 20 | 9639 | 6468 | **8** |
| | | 2 | | | | 50 | | | 64741 | 58721 | 62 | 21443 | 14681 | 25 | 68241 | 61285 | 137 | 39119 | 57964 | 41 | 21055 | 14681 | **21** |
| | | 3 | | | | 37 | | | 64741 | 58721 | 83 | 21443 | 14681 | **26** | 68241 | 61285 | 250 | 39119 | 57964 | 53 | 21055 | 14681 | 27 |
| | | 4 | | | | 58 | | | 64741 | 58721 | 178 | 21443 | 14681 | 37 | 68241 | 61285 | 358 | 39119 | 57964 | 93 | 21055 | 14681 | **35** |
| | | 5 | | | | 42 | | | 64741 | 58721 | 152 | 21443 | 14681 | 45 | 68241 | 61285 | 526 | 39119 | 57964 | 66 | 21055 | 14681 | **45** |
| | | 6 | | | | 49 | | | 64741 | 58721 | 127 | 21443 | 14681 | 39 | 68241 | 61285 | 518 | 39119 | 57964 | 71 | 21055 | 14681 | **39** |
| | | 7 | | | | 38 | | | 64741 | 58721 | 99 | 21443 | 14681 | 31 | 68241 | 61285 | 468 | 39119 | 57964 | 53 | 21055 | 14681 | **30** |
| | | 8 | | | | 15 | | | 64741 | 58721 | 31 | 21443 | 14681 | 11 | 68241 | 61285 | 299 | 39119 | 57964 | 19 | 21055 | 14681 | **10** |

$\alpha = 1, \gamma = 1$

From Table 1, we make the following observations:

- The results clearly highlight the computational inefficiency of $r$-HMIP$_{CAC3}$, as highlighted by its relatively higher CPU times compared to the other formulations. This is mainly due to a weak LP relaxation of $r$-HMIP$_{CAC3}$, besides the large problem size post-presolve.
- A comparison of the results between $r$-HMIP$_{CAC1}$ and $r$-HMIP$_{CAC2}$ shows that the latter speeds up the computation by a factor of 2 to 4 for instances with $p = 7$, and by a factor of 2 to 7 for instances with $p = 10$, as highlighted by their relative CPU times. This gain comes largely from the elimination of a substantially larger proportion of the variables and constraints post-presolve in $r$-HMIP$_{CAC2}$ compared to $r$-HMIP$_{CAC1}$.
- The size of the reduced formulation $r$-HMIP$_{rCAC1}$ post-presolve is significantly smaller, resulting in a reduction in its CPU times by a factor of 1.5 to 2, compared to its original $r$-HMIP$_{CAC1}$ formulation.
- The reduced formulation $r$-HMIP$_{rCAC2}$ post-presolve has almost the same size as its original formulation $r$-HMIP$_{CAC2}$, hence yielding only marginal additional savings in CPU times.
- The dual based formulation $r$-HMIP$_{DF}$ is computationally more efficient than $r$-HMIP$_{CAC1}$ and $r$-HMIP$_{CAC3}$, but not as efficient as $r$-HMIP$_{CAC2}$ and $r$-HMIP$_{rCAC2}$.
- Compared to $r$-HMIP$_{rCAC1}$, $r$-HMIP$_{rCAC2}$ results in savings in CPU time by a factor of 1.5 to 6.

In summary, the results for small instances in Table 1 show that $r$-HMIP$_{DF}$, $r$-HMIP$_{CAC2}$ and $r$-HMIP$_{rCAC2}$ are computationally more efficient formulations, as highlighted by their *significantly* lower CPU times, compared to the rest. Amongst these three, $r$-HMIP$_{rCAC2}$ requires the lowest CPU times for 35 out of the 39 problem instances, and $r$-HMIP$_{CAC2}$ requires the lowest CPU times for the remaining 4 instances. Although CPU time for $r$-HMIP$_{DF}$ is not the lowest for any of these 39 instances, it is not significantly worse than the lowest CPU time for any of these instances.

In Table 2, we present additional computational results using 12 larger problem instances derived from the AP dataset with 100 and 200 nodes, for the three most efficient formulations, namely $r$-HMIP$_{CAC2}$, $r$-HMIP$_{rCAC2}$, and $r$-HMIP$_{DF}$, as highlighted above. The discount factors for collection ($\alpha$), transshipment ($\delta$), and distribution ($\gamma$) are set to 3, 0.75, and 2, respectively. Like the computational results reported in Table 1, the set of p hubs in the existing hub-and-spoke network for each of these 12 instances is obtained by solving a corresponding uncapacitated $p$-hub median problem (Ebery et al., 2000). CPLEX is able to solve all the 12 instances using the $r$-HMIP$_{DF}$ formulation. However, CPLEX is able to solve only 4 of them using the $r$-HMIP$_{CAC2}$ and $r$-HMIP$_{rCAC2}$ formulations, and using CPU times that are up to an order-of-magnitude larger compared to $r$-HMIP$_{DF}$. The remaining 8 instances, indicated by "Memory" against their CPU times, cannot be solved due to memory limitations of the hardware used in the experiments.

In the following section, we further exploit the structure of $r$-HMIP$_{DF}$, HMIP$_{CAC2}$ and $r$-HMIP$_{rCAC2}$ using Benders decomposition. Benders decomposition has been successfully applied to a variety of problems arising in the context of Facility Location (Vatsa and Jayaswal, 2016)

Table 2: Comparison between dual based and CAC based single level reformulations of $r$-HMIP using AP dataset

| Parameters | | | CPLEX | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | $r$-HMIP$_{DF}$ | | $r$-HMIP$_{CAC2}$ | | $r$-HMIP$_{rCAC2}$ | |
| $\|N\|$ | $p$ | $r$ | Gap | CPU(s) | Gap | CPU(s) | Gap | CPU(s) |
| 100 | 10 | 5 | 0 | 1,476 | 0 | 8,881 | 0 | 8,315 |
| | | 6 | 0 | 935 | 0 | 6,039 | 0 | 5,070 |
| | | 7 | 0 | 891 | 0 | 4,299 | 0 | 3,996 |
| | | 8 | 0 | 328 | 0 | 3,395 | 0 | 3,255 |
| | 15 | 5 | 0 | 6,328 | - | Memory | - | Memory |
| | | 6 | 0 | 8,913 | - | Memory | - | Memory |
| | | 7 | 0 | 8,092 | - | Memory | - | Memory |
| | | 8 | 0 | 6,035 | - | Memory | - | Memory |
| 200 | 10 | 5 | 0 | 21,948 | - | Memory | - | Memory |
| | | 6 | 0 | 22,505 | - | Memory | - | Memory |
| | | 7 | 0 | 19,190 | - | Memory | - | Memory |
| | | 8 | 0 | 6,787 | - | Memory | - | Memory |

$\alpha = 3$, $\delta = 0.75$, $\gamma = 2$.
"Memory" denotes insufficient memory to solve the problem

and Hub Location (de Camargo et al., 2008, 2009; Contreras et al., 2011a). Since the single level reformulations of $r$-HMIP bear similarity with the hub location models, we expect Benders decomposition to be successful here as well.

## 5. Benders Decomposition for HMIP

Benders decomposition is a well-known method of partitioning an MIP into *an integer master problem* and a *linear sub-problem* (Benders, 1962). In this section, we present Benders decomposition for the three efficient HMIP formulations: $r$-HMIP$_{DF}$, $r$-HMIP$_{CAC2}$ and $r$-HMIP$_{rCAC2}$. We also discuss the computational gain from Benders decomposition of $r$-HMIP$_{rCAC2}$ relative to that from the Benders decomposition of $r$-HMIP$_{CAC2}$, which is otherwise not possible when solving the two formulations directly using CPLEX.

### 5.1. Benders Decomposition of $r$-$HMIP_{DF}$

Let $Z = \{z_k \in \{0,1\} | \sum_{k\in H} z_k = p - r\}$. For any fixed binary solution $\bar{\mathbf{z}} \in Z$, the resulting problem in the space of $\boldsymbol{\phi}$, $\boldsymbol{\delta}$ and $\mathbf{V}$ variables, which we refer to as the *primal sub-problem* (PS), can be stated as:

$$[PS_{DF}] : \max_{\boldsymbol{\phi},\boldsymbol{\delta},\mathbf{V}} \sum_{i\in N}\sum_{j\in N} \phi_{ij} - \sum_{k\in H} V_k \tag{17}$$

$$\text{s.t. } \phi_{ij} - \delta_{ijk} \leqslant W_{ij} d_{ijkm} \qquad \forall i,j \in N; \forall k,m \in H, k = m \tag{18}$$

$$\phi_{ij} - \delta_{ijk} - \delta_{ijm} \leqslant W_{ij} d_{ijkm} \qquad \forall i,j \in N; \forall k,m \in H, k \neq m \tag{19}$$

$$V_k \leqslant M_k \bar{z}_k \qquad \forall k \in H \tag{20}$$

$$V_k \geqslant \sum_{i\in N}\sum_{j\in N} \delta_{ijk} - M_k(1 - \bar{z}_k) \qquad \forall k \in H \tag{21}$$

$$\phi_{ij} \text{ unbounded}; \quad \delta_{ijk} \geqslant 0; V_k \geqslant 0; \qquad \forall i,j \in N; \forall k \in H \tag{22}$$

Let $\chi_{ijkk}$, $\chi_{ijkm}$ $(k \neq m)$, $\tau_k$, and $\omega_k$ be dual variables associated with constraints (18), (19), (20), and (21), respectively. The dual of the sub-problem can be stated as follows:

$$
\begin{aligned}
[DS_{DF}] : \min_{\boldsymbol{\chi},\boldsymbol{\tau},\boldsymbol{\omega}} & \sum_{i \in N}\sum_{j \in N}\sum_{k \in H}\sum_{m \in H} W_{ij} d_{ijkm} \chi_{ijkm} \\
& + \sum_{k \in H} M_k \bar{z}_k \tau_k + \sum_{k \in H} M_k (1 - \bar{z}_k)\omega_k && (23) \\
\text{s.t.} \; & \sum_{k \in H}\sum_{m \in H} \chi_{ijkm} = 1 && \forall i, j \in N \quad (24) \\
& \sum_{m \in H} \chi_{ijkm} + \sum_{m \in H \setminus \{k\}} \chi_{ijmk} \leqslant \omega_k && \forall i, j \in N; \forall k \in H \quad (25) \\
& \tau_k - \omega_k \geqslant -1 && \forall k \in H \quad (26) \\
& \chi_{ijkm}, \tau_k, \omega_k \geqslant 0 && \forall i, j \in N; \forall k, m \in H \quad (27)
\end{aligned}
$$

The above dual minimization problem is bounded since all the variables in the problem ($\chi_{ijkm}$, $\tau_k$ and $\omega_k$) are positive. Denoting the set of all extreme points of $DS_{DF}$ by $EP_{DF}$, the master problem for $r$-HMIP$_{DF}$ is as follows:

$$
\begin{aligned}
[MP_{DF}] : \max_{\mathbf{z},\theta} \; & \theta && (28) \\
\text{s.t.} \; & \sum_{k \in H} z_k = p - r && (29) \\
& \theta \leqslant \sum_{i \in N}\sum_{j \in N}\sum_{k \in H}\sum_{m \in H} W_{ij} d_{ijkm} \bar{\chi}_{ijkm} + \sum_{k \in H} M_k \bar{\tau}_k z_k + \\
& \quad \sum_{k \in H} M_k (1 - z_k)\bar{\omega}_k && \forall (\bar{\chi}_{ijkm}, \bar{\tau}_k, \bar{\omega}_k) \in EP_{DF} \quad (30) \\
& \theta \geqslant 0, z_k \in \{0, 1\} && \forall k \in H \quad (31)
\end{aligned}
$$

Solving the master problem with the addition of Benders cut (30) gives the upper bound, while the sub-problem gives a lower bound to the original problem. The algorithm terminates when the difference between upper and the best lower bound falls within a pre-specified tolerance $\epsilon$.

### *5.2. Benders Decomposition of r-HMIP$_{CAC2}$*

Let $Z = \{z_k \in \{0, 1\} | \sum_{k \in H} z_k = p - r\}$. For any fixed binary solution $\bar{\mathbf{z}} \in Z$, the resulting problem in the space of $\mathbf{X}$ variables, which we refer to as the *primal sub-problem* (PS), can be stated as:

$$
\begin{aligned}
[PS_{CAC2}] : \max_{\mathbf{X}} \; & \sum_{i \in N}\sum_{j \in N}\sum_{k \in H}\sum_{m \in H} W_{ij} d_{ijkm} X_{ijkm} && (32) \\
\text{s.t.} \; & \sum_{k \in H}\sum_{m \in H} X_{ijkm} = 1 && \forall\, i, j \in N \quad (33) \\
& \sum_{m \in H} X_{ijkm} + \sum_{m \in H \setminus \{k\}} X_{ijmk} \leqslant \bar{z}_k && \forall\, i, j \in N, k \in H \quad (34) \\
& \sum_{(q,s) \in E_{ijkm}} X_{ijqs} \leqslant 2 - \bar{z}_k - \bar{z}_m && \forall\, i, j \in N, k, m \in H \quad (35) \\
& X_{ijkm} \geqslant 0 && \forall\, i, j \in N, k, m \in H \quad (36)
\end{aligned}
$$

Associating $\phi^1_{ij}, \delta^1_{ijk}$ and $\beta^1_{ijkm}$ for the constraints (33), (34) and (35), respectively, we get the following dual sub-problem:

$$[DS_{CAC2}]:$$

$$\min_{\boldsymbol{\phi}^1,\boldsymbol{\lambda}^1,\boldsymbol{\beta}^1} \sum_{i\in N}\sum_{j\in N}\phi^1_{ij} + \sum_{i\in N}\sum_{j\in N}\sum_{k\in H}\lambda^1_{ijk}\bar{z}_k + \sum_{i\in N}\sum_{j\in N}\sum_{k\in H}\sum_{m\in H}\beta^1_{ijkm}(2-\bar{z}_k-\bar{z}_m) \tag{37}$$

$$\text{s.t.} \sum_{(q,s)|d_{ijqs}<d_{ijkm}}\beta^1_{ijqs} + \lambda^1_{ijk} + \phi^1_{ij} \geqslant W_{ij}d_{ijkm} \qquad \forall\, i,j\in N, k,m\in H, k=m \tag{38}$$

$$\sum_{(q,s)|d_{ijqs}<d_{ijkm}}\beta^1_{ijqs} + \lambda^1_{ijk} + \lambda^1_{ijm} + \phi^1_{ij} \geqslant W_{ij}d_{ijkm} \qquad \forall\, i,j\in N, k,m\in H, k\neq m \tag{39}$$

$$\phi^1_{ij} \text{ unbounded}; \beta^1_{ijkm}, \lambda^1_{ijk} \geqslant 0 \qquad \forall\, i,j\in N, k,m\in H \tag{40}$$

The above dual minimization problem is bounded since the variables $\lambda^1_{ijk}$ and $\beta^1_{ijkm}$ are positive, and the constraints (38) and (39) ensure that the free variable $\phi^1_{ij}$ has a finite lower bound. Denoting the set of all extreme points of $DS_{CAC2}$ by $EP_{CAC2}$, the master problem for $r$-HMIP$_{CAC2}$ is as follows:

$$[MP_{CAC2}]: \max_{\mathbf{z},\theta}\ \theta \tag{41}$$

$$\text{s.t.} \sum_{k\in H} z_k = p-r \tag{42}$$

$$\theta \leqslant \sum_{i\in N}\sum_{j\in N}\overline{\phi^1}_{ij} + \sum_{i\in N}\sum_{j\in N}\sum_{k\in H}\overline{\lambda^1}_{ijk}z_k + \sum_{i\in N}\sum_{j\in N}\sum_{k\in H}\sum_{m\in H}\overline{\beta^1}_{ijkm}(2-z_k-z_m) \quad \forall(\overline{\phi^1}_{ij}, \overline{\lambda^1}_{ijk}, \overline{\beta^1}_{ijkm}) \in EP_{CAC2} \tag{43}$$

$$\theta \geqslant 0, z_k \in \{0,1\} \qquad \forall k\in H \tag{44}$$

Solving the master problem with the addition of Benders cut (43) gives the upper bound, while the sub-problem gives a lower bound to the original problem. The algorithm terminates when the difference between upper and the best lower bound falls within a pre-specified tolerance$\epsilon$. gap.

### 5.3. Benders Decomposition of $r$-$HMIP_{rCAC2}$

Let $Z = \{z_k \in \{0,1\}|\sum_{k\in H} z_k = p-r\}$. For any fixed binary solution $\bar{\mathbf{z}} \in Z$, the primal sub-problem for $r$-HMIP$_{rCAC2}$ can be written as:

$$[PS_{rCAC2}]: \max_{\mathbf{X}} \sum_{i\in N}\sum_{j\in N}\sum_{k\in H}\sum_{m\in H} W_{ij}d_{ijkm}X_{ijkm} \tag{45}$$

$$\text{s.t.} \sum_{k\in H}\sum_{m\in H} X_{ijkm} = 1 \qquad \forall\, i,j\in N \tag{46}$$

$$\sum_{m\in H} X_{ijkm} + \sum_{m\in H\setminus\{k\}} X_{ijmk} \leqslant \bar{z}_k \qquad \forall\, i,j\in N, k\in H \tag{47}$$

$$\sum_{(q,s)\in E_{ijkm}} X_{ijqs} \leqslant 2 - \bar{z}_k - \bar{z}_m \qquad \forall\, i,j \in N, (k,m) \in H'_{ij} \tag{48}$$

$$X_{ijkm} \geqslant 0 \qquad \forall\, i,j \in N, k,m \in H \tag{49}$$

Associating $\phi^2_{ij}$, $\lambda^2_{ijk}$ and $\beta^2_{ijkm}$ as the dual variables with the constraints (46), (47) and (48), respectively, we get the following dual sub-problem:

$[DS_{rCAC2}]$ :

$$\min_{\boldsymbol{\phi}^2,\boldsymbol{\lambda}^2,\boldsymbol{\beta}^2} \sum_{i\in N}\sum_{j\in N} \phi^2_{ij} + \sum_{i\in N}\sum_{j\in N}\sum_{k\in H} \lambda^2_{ijk}\bar{z}_k + \sum_{i\in N}\sum_{j\in N}\sum_{(k,m)\in H'_{ij}} \beta^2_{ijkm}(2 - \bar{z}_k - \bar{z}_m) \tag{50}$$

$$\text{s.t.} \sum_{(q,s)\in B^1_{ijkm}} \beta^2_{ijqs} + \lambda^2_{ijk} + \phi^2_{ij} \geqslant W_{ij}d_{ijkm} \qquad \forall\, i,j \in N, k,m \in H, k = m \tag{51}$$

$$\sum_{(q,s)\in B^1_{ijkm}} \beta^2_{ijqs} + \lambda^2_{ijk} + \lambda^2_{ijm} + \phi^2_{ij} \geqslant W_{ij}d_{ijkm} \qquad \forall\, i,j \in N, k,m \in H, k \neq m \tag{52}$$

$$\beta^2_{ijkm} \geqslant 0 \qquad \forall\, i,j \in N, (k,m) \in H'_{ij} \tag{53}$$

$$\phi^2_{ij} \text{ unbounded}; \lambda^2_{ijk} \geqslant 0 \qquad \forall\, i,j \in N, k \in H \tag{54}$$

where $B^1_{ijkm}= \{(q,s)|d_{ijqs} < d_{ijkm};\;\; \forall (q,s) \in H'_{ij}\}$.
The above dual minimization sub-problem is bounded since the variables $\lambda^2_{ijk}$ and $\beta^2_{ijkm}$ are positive, and the constraints (51) and (52) ensure that the free variable $\phi^2_{ij}$ has a finite lower bound. Denoting the set of all extreme points of $DS_{rCAC2}$ by $EP_{rCAC2}$, the master problem for $r$-HMIP$_{rCAC2}$ is as follows:

$$[MP_{rCAC2}] : \max_{\mathbf{z},\theta}\; \theta \tag{55}$$

$$\text{s.t.} \sum_{k\in H} z_k = p - r \tag{56}$$

$$\theta \leqslant \sum_{i\in N}\sum_{j\in N} \overline{\phi^2}_{ij} + \sum_{i\in N}\sum_{j\in N}\sum_{k\in H} \overline{\lambda^2}_{ijk} z_k + \sum_{i\in N}\sum_{j\in N}\sum_{(k,m)\in B^1_{ijkm}} \overline{\beta^2}_{ijkm}(2 - z_k - z_m) \quad \forall (\overline{\phi^2}_{ij}, \overline{\lambda^2}_{ijk}, \overline{\beta^2}_{ijkm}) \in EP_{rCAC2} \tag{57}$$

$$\theta \geqslant 0, z_k \in \{0,1\} \qquad \forall k \in H \tag{58}$$

Solving the master problem with the addition of Benders cut (57) gives the upper bound, while the sub-problem gives a lower bound to the original problem. The algorithm terminates when the difference between upper and the best lower bound falls within a pre-specified tolerance $\epsilon$.

**Remark:** As evident from the computational results in Table 1, rCAC2 does not enjoy significant computational advantage over CAC2 when solving the resulting single-level $r$-HMIP formulations with these CACs directly using CPLEX since both these models are approximately the same size (exactly the same number of variables and approximately the same number of constraints) after the presolve operation. However, when the single level $r$-HMIP is decomposed into a master problem consisting of only the binary variables ($z_k$) and a sub-problem consisting of only the continuous variables ($X_{ijkm}$), presolve loses its effect. Nonetheless, with Benders decomposition, the dual sub-problem (50)-(54) resulting from the use of rCAC2 has fewer variables ($\beta_{ijkm}$) compared to the dual sub-problem (37)-(40) resulting from the use of CAC2. With $|N|$ nodes and $p$ hubs, the use of rCAC2 and CAC2 result in $|N|^2(p^2+p)/2$ and $|N|^2p^2$ $\beta_{ijkm}$ variables, respectively. For example, a problem instance with $|N| = 200$ and $p = 10$ results in 2,200,000 $\beta_{ijkm}$ variables with the use of rCAC2, which is otherwise 4,000,000 with the use of CAC2. Thus, rCAC2 aids in the reduction in the size of the dual sub-problem by a factor of $0.5 + (0.5/p)$, which tends to 0.5 as $p$ becomes large. We expect Benders decomposition to benefit from this reduction in the size of the dual sub-problem with rCAC2, and provide a computational advantage over CAC2.

### *5.4. Computational Results*

We now repeat the computational experiments, using the same 12 instances as in Table 2, for the Benders decomposition of the three formulations, namely $r$-HMIP$_{DF}$, $r$-HMIP$_{CAC2}$, and $r$-HMIP$_{rCAC2}$. For ease of comparison of Benders decomposition with the direct solution using CPLEX, we also present the computational results already reported in Table 2. Recall that $r$-HMIP$_{CAC2}$ and $r$-HMIP$_{rCAC2}$ performed significantly worse than $r$-HMIP$_{DF}$ when solved directly using CPLEX. In that case, CPLEX was able to solve all the 12 instances using $r$-HMIP$_{DF}$. On the other hand, it could solve only 4 of them using $r$-HMIP$_{CAC2}$ and $r$-HMIP$_{rCAC2}$, while the remaining 8 instances could not be solved due to memory limitations of the hardware used in the experiments. With Benders decomposition, the performances of $r$-HMIP$_{CAC2}$ and $r$-HMIP$_{rCAC2}$ improve drastically, while that of $r$-HMIP$_{DF}$ becomes worse. With Benders decomposition, we are now able to solve all the 12 instances using the $r$-HMIP$_{rCAC2}$ formulation within the 10 hour (36,000 seconds) CPU time limit. On the other hand, we can now solve only 4 (of the 12) instances using $r$-HMIP$_{DF}$, while the remaining 8 instances cannot be solved due to memory limitations of the hardware.

Finally, the results highlight that only two formulations, namely $r$-HMIP$_{DF}$ and Benders decomposition of $r$-HMIP$_{rCAC2}$, could solve all the 12 large instances of the problem. Further, for the 100-node instances, $r$-HMIP$_{DF}$ outperforms (in terms of CPU time) the Benders decomposition of $r$-HMIP$_{rCAC2}$ on 5 out of 8 problem instances. On the remaining 3 instances, Benders decomposition of $r$-HMIP$_{rCAC2}$ performs better. Whereas for the 200-node problem instances, Benders decomposition outperforms $r$-HMIP$_{DF}$ by a factor of 4-8. Hence, $r$-HMIP$_{DF}$ performance better for smaller instances of the problem, while for larger instances, Benders decomposition of $r$-HMIP$_{rCAC2}$ is clearly better. This observation is consistent with the remarks made in the previous subsection regarding the computational advantage of Benders decomposition of $r$-HMIP$_{rCAC2}$ formulation.

## 6. Hub Protection Problem

In hub protection problems, the defender seeks to protect/fortify a subset of hubs against interdiction. In this section, we study a $u$-hub median protection problem ($u$-HMPP) in a multiple allocation hub-and-spoke network with $p$ existing hubs. Figure 2 shows a schematic representation of $u$-HMPP studied in this paper. The problem is modeled as a Stackelberg game between a defender and an attacker. The attacker seeks to interdict a subset of $r$ hubs

Table 3: Comparison between Benders decomposition of dual based and CAC based single level reformulations of $r$-HMIP using AP dataset

| Parameters | | | CPLEX | | | | | | Benders Decomposition | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | $r$-HMIP$_{DF}$ | | $r$-HMIP$_{CAC2}$ | | $r$-HMIP$_{rCAC2}$ | | $r$-HMIP$_{DF}$ | | $r$-HMIP$_{CAC2}$ | | $r$-HMIP$_{rCAC2}$ | |
| $\|N\|$ | $p$ | $r$ | Gap | CPU(s) | Gap | CPU(s) | Gap | CPU(s) | Gap | CPU(s) | Gap | CPU(s) | Gap | CPU(s) |
| 100 | 10 | 5 | 0 | 1,476 | 0 | 8,881 | 0 | 8,315 | 0 | 2,334 | 0 | 1,212 | 0 | **1,100** |
| | | 6 | 0 | **935** | 0 | 6,039 | 0 | 5,070 | 0 | 1,978 | 0 | 1,290 | 0 | 1,034 |
| | | 7 | 0 | 891 | 0 | 4,299 | 0 | 3,996 | 0 | 1,175 | 0 | 562 | 0 | **444** |
| | | 8 | 0 | 328 | 0 | 3,395 | 0 | 3,255 | 0 | **195** | 0 | 310 | 0 | 233 |
| | 15 | 5 | 0 | **6,328** | - | Memory | - | Memory | - | Memory | 87% | 36,000 | 0 | 21,080 |
| | | 6 | 0 | **8,913** | - | Memory | - | Memory | - | Memory | 86% | 36,000 | 0 | 24,091 |
| | | 7 | 0 | **8,092** | - | Memory | - | Memory | - | Memory | 94% | 36,000 | 0 | 20,606 |
| | | 8 | 0 | **6,035** | - | Memory | - | Memory | - | Memory | 85% | 36,000 | 0 | 14,956 |
| 200 | 10 | 5 | 0 | 21,948 | - | Memory | - | Memory | - | Memory | 42% | 36,000 | 0 | **5,779** |
| | | 6 | 0 | 22,505 | - | Memory | - | Memory | - | Memory | 26% | 36,000 | 0 | **3,328** |
| | | 7 | 0 | 19,190 | - | Memory | - | Memory | - | Memory | 0 | 18,572 | 0 | **2,274** |
| | | 8 | 0 | 6,787 | - | Memory | - | Memory | - | Memory | 0 | 8,109 | 0 | **1,180** |

$\alpha = 3$, $\delta = 0.75$, $\gamma = 2$
"Memory" denotes insufficient memory to solve the problem

from an existing set $H$ of $p$ hubs with the objective to maximize the defender's optimal routing/transportation cost through the $p-r$ surviving hubs in the network post-interdiction, which is modeled as $r$-HMIP$_{2L}$, as studied extensively in Section 3 and 4. The defender, in anticipation of the attacker's action, seeks to protect a subset of $u$ hubs from the existing set of $p$ hubs in the network such that its post-interdiction optimal routing cost through the $p-r$ surviving hubs, which the attacker intends to maximize, is as small as possible.

While protection problems have been widely studied in the context of facility location (Church and Scaparra, 2007b; Scaparra and Church, 2008a,b; Aksen et al., 2010; Scaparra and Church, 2012; Aksen and Aras, 2012; Aksen et al., 2013), they have received little attention in the context of hub-and-spoke networks, with Lei (2013) and Ghaffarinasab and Atayi (2017), to the best of our knowledge, being the only two papers in the area. Lei (2013) presented a trilevel MIP formulation for multiple allocation $u$-HMPP, which is reduced to a bilevel MIP using CAC1. However, no solution method or computational results were reported. Ghaffarinasab and Atayi (2017) solved large instances of the $u$-HMPP using an implicit enumeration procedure. In this section, we study the same trilevel MIP formulation as Lei (2013). However, we reduce the trilevel formulation to bilevel using rCAC2, which has been shown to perform the best among the different CACs discussed in Section 3 and Section 4 above. This allows us to solve large instances of $u$-HMPP using an efficient algorithm based on implicit enumeration and Benders decomposition, as described in the following subsection.

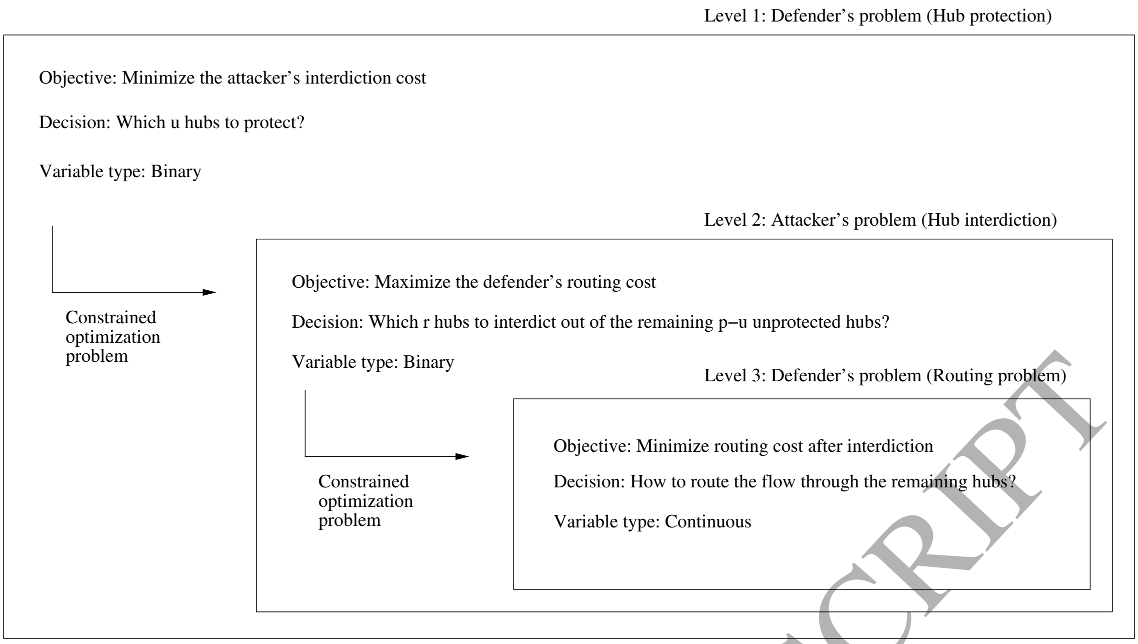


Figure 2: Hub protection problem as a tri-level MIP

### *6.1. Model Formulation*

To formulate $u$-HMPP, we define $y_k = 1$ if hub $k$ is protected, 0 otherwise. With this new set of variables, the trilevel formulation of $u$-HMPP can be stated as follows:

$$[u\text{-HMPP}_{3L}] : \min_{\mathbf{y}} \quad T_1 \tag{59}$$

$$\text{s.t.} \sum_{k \in H} y_k = u \tag{60}$$

$$y_k \in \{0, 1\} \tag{61}$$

$$T_1 = \max_{\mathbf{z}} T_2 \tag{62}$$

$$\text{s.t.}(2), (3)$$

$$y_k \leqslant z_k \qquad \forall\, k \in H \tag{63}$$

$$T_2 = \max_{\mathbf{x}} W_{ij} D_{ijkm} X_{ijkm} \tag{64}$$

$$(5) - (7)$$

The defender's problem of protecting a subset of $u$ out of an existing set of $p$ hubs is given by (59)-(61) at level 1. He takes this decision in anticipation of the attacker's decision, which is modeled as a bilevel $r$-HMIP, as described in Section 3, with the additional constraint (63) to ensure that a protected hub cannot be attacked. For feasibility, we ensure that $u + r \leqslant p$. As discussed in Section 3, the lower bilevel $r$-HMIP can be reduced to single level using CACs. Using the most efficient set of CACs, namely rCAC2, $u$-HMPP$_{3L}$ can be restated as the following bilevel program:

$$[u\text{-HMPP}_{2L}] : \min_{\mathbf{y}} \; T_1 \tag{65}$$

$$\text{s.t.}(60), \; (61)$$

$$T_1 = \max_{\mathbf{z},\mathbf{X}} \sum_{i\in N}\sum_{j\in N}\sum_{k\in H}\sum_{m\in H} W_{ij} d_{ijkm} X_{ijkm} \tag{66}$$
$$\text{s.t.}(5)-(7),(63), r\text{-CAC2}$$

*6.2. Implicit Enumeration Algorithm*

We present an implicit enumeration algorithm for solving $u$-HMPP, inspired by the solution method proposed by Scaparra and Church (2008a) for $r$-interdiction median problem with fortification ($r$-IMF). The algorithm is based on the proposition that the optimal solution to $r$-IMF will necessarily contain at least one of the facilities interdicted in $r$-interdiction median problem ($r$-IMP) since any other combination of protected facilities will not prevent the worst scenario for the defender.

The implicit enumeration algorithm for $u$-HMPP is described as follows. At the root node of the search tree, the algorithm solves an $r$-HMIP, yielding $r$ interdicted hubs. The root node is then branched into $r$ children nodes, each corresponding to the protection of a hub $k$ (by setting $y_k = 1$) out of the $r$ interdicted hubs. At each of these $r$ nodes, it solves a conditional hub interdiction problem (CHIP), which is an HMIP with the restriction that the protected hub $k$ cannot be interdicted (imposed using constraint set (63)). The solution to each of these CHIPs gives $r$ interdicted hubs. Each of these nodes is, in turn, branched into $r$ children nodes, each corresponding to the protection of a hub $k$ (by setting $y_k = 1$) out of the $r$ interdicted hubs, in addition to the hubs protected at its parent node. This procedure is repeated until the number of protected hubs on the path starting from the root node to the current node is $u$. Any node at which $u$ hubs are protected is called a leaf node. When each of the paths from the root node terminates in a leaf node, then the node with the lowest objective function value to its corresponding CHIP provides the solution to $u$-HMPP. If at any stage, the algorithm visits a node correspondng to a set of protected hubs, which is the same as that corresponding to some other already visited node, then the algorithm skips that node, since CHIPs at both the nodes are the same. At each node in the search tree, HMIP/CHIP is reduced to a single level MIP using rCAC2, which is solved using Benders decomposition.

To summarize the above procedure, we use $\ddot{y}$ and $\ddot{z}$ to denote the optimal solution vector to the protection and interdiction variables, respectively. Further, $\ddot{\theta}$ denotes the optimal objective function value to $u$-HMPP. Let $r_0$ denote the root node of the search tree, and $S$ denote the set of nodes in the tree to be visited. We define the following two sets associated with each node $n$: $C_n$ is the set of candidate hubs to be protected in the subsequent nodes on the subpath starting from node n; $F_n$ is the set of hubs protected on the path from root to node n. We also define a set $F$, which stores the set of protected hubs at all the nodes visited by the algorithm.

We use CHIP($F_n$) to denote CHIP with the additional restriction that the hubs in $F_n$ cannot be interdicted. Using the above notations, the implicit enumeration procedure can be outlined in Algorithm 1.

**Algorithm 1** Implicit Enumeration

**procedure** IMPLICIT ENUMERATION<br>
  $F \leftarrow \phi$<br>
  $F_{r_0} \leftarrow \phi$<br>
  $\ddot{y}_k \leftarrow 0 \quad \forall k \in H.$<br>
  Solve CHIP$(F_{r_0})$. $\hat{z} \leftarrow \{k| \, z_k = 0\}$; $\hat{\theta} \leftarrow$ obj. fun. value of CHIP$(F_{r_0})$<br>
  $\ddot{\theta} \leftarrow \hat{\theta}$; $C_{r_0} = \{k| \, \hat{z_k} = 0\}$; $S = \{r_0\}$<br>
  **while** $S \neq \phi$ **do**<br>
    select $n \in S$<br>
    **while** $C_n \neq \phi$ **do**<br>
      Select $k \in C_n$<br>
      $C_n \leftarrow C_n \backslash \{k\}$<br>
      Generate node $n_1$ with $F_{n_1} = F_n \cup \{k\}$<br>
      **while** $F_{n_1} \notin F$ **do**<br>
        $F = F \cup \{F_{n_1}\}$<br>
        Solve CHIP$(F_{n_1})$. $\hat{z} \leftarrow \{k| \, z_k = 0\}$; $\hat{\theta} \leftarrow$ obj. fun. value of CHIP$(F_n)$<br>
        **if** $|F_{n_1}| = q$ **then**<br>
          **if** $\hat{\theta} < \ddot{\theta}$ **then**<br>
            $\ddot{z} \leftarrow \hat{z}$; $\ddot{\theta} \leftarrow \hat{\theta}$<br>
            **for** $k \in H$ **do**<br>
              **if** $k \in F_{n_1}$ **then**<br>
                $\ddot{y_k} = 1$<br>
              **else** $\ddot{y_k} = 0$<br>
              **end if**<br>
            **end for**<br>
          **end if**<br>
        **else** $C_{n_1} = \{k| \, \hat{z_k} = 0\}$; $S = S \cup \{n_1\}$<br>
        **end if**<br>
      **end while**<br>
    **end while**<br>
  **end whilereturn** $\ddot{\theta}, \ddot{y}, \ddot{z}$<br>
**end procedure**

The proposed implicit enumeration procedure examines $r^0 + r^1 + r^2 + .... + r^u = \frac{r^{(u+1)}-1}{r-1}$ nodes, as opposed to $\binom{p}{q}$ CHIPs in complete enumeration of the set of $q$ protected hubs. To speed up the implicit enumeration algorithm, we next explore ways to solve HIPs/CHIPs efficiently.

**Remark:** Since the feasible region of the optimization problem at any child node is a strict subset of the corresponding feasible region of its parent node, the benders cuts generated in solving the parent node remains valid for its children also. Therefore, in the implicit enumeration algorithm, we retain those benders cuts of the parent nodes in each of their children nodes to ensure faster convergence of the corresponding optimization problem. This reduces the computational time for $u$-HMPP problem instances significantly.

### *6.3. Computational Results*

Table 4 provides a comparison of the computational performance of the implicit enumeration algorithm versus complete enumeration for 12 different problem instances derived from AP dataset by setting the parameters as: $N \in \{50, 100\}$, $p \in \{10\}$, $r \in \{5, 6, 7\}$, $u \in \{1, 2\}$. The discount factors are set to: $\alpha = 3, \delta = 0.75, \gamma = 2$. The results show that implicit enumeration algorithm is much faster, providing a reduction of 43% - 85% in CPU times, compared to complete enumeration.

Table 4: Computational results for hub protection problems with AP datasets

| $\|N\|$ | $p$ | $r$ | $u$ | Complete Enumeration CPU(s) | Implicit Enumeration CPU(s) | % Reduction in CPU Time |
|---|---|---|---|---|---|---|
| 50 | 10 | 5 | 1 | 845 | 313 | 63 |
| | | 5 | 2 | 2,922 | 485 | 84 |
| | | 6 | 1 | 656 | 312 | 53 |
| | | 6 | 2 | 2,039 | 689 | 67 |
| | | 7 | 1 | 413 | 210 | 50 |
| | | 7 | 2 | 1,320 | 485 | 64 |
| 100 | 10 | 5 | 1 | 3,820 | 1,454 | 61 |
| | | 5 | 2 | 13,176 | 2,063 | 85 |
| | | 6 | 1 | 3,103 | 1,320 | 58 |
| | | 6 | 2 | 8,908 | 1,919 | 79 |
| | | 7 | 1 | 1,956 | 1,119 | 43 |
| | | 7 | 2 | 5,356 | 1,955 | 64 |
| Min. | | | | 413 | 210 | 43 |
| Avg. | | | | 3,710 | 1027 | 64 |
| Max. | | | | 13,176 | 2063 | 85 |

$\alpha = 3, \delta = 0.75, \gamma = 2$

## 7. Conclusion

In this paper, we studied multiple allocation $r$-HMIP and multiple allocation $u$-HMPP for hub-and-spoke networks by formulating them as bilevel and trilevel MIP problems, respectively. For the bilevel $r$-HMIP, we explored two alternate ways to reduce it to a single-level optimization problem. The first approach uses the dual formulation of the lower level routing problem to reduce it to single-level, while the second approach exploits the structure of the solution to the lower level problem using CACs. We studied alternate forms of CACs, the dominance relation among them, and their computational performances. The results indicate that the best among our proposed alternate sets of CACs provides a computational advantage (in terms of reduced CPU times) by a factor of 7 compared to the set of CACs proposed in the literature. We further provided reduced versions of the alternate sets of CACs, one of which in conjunction with Benders decomposition was able to solve all the 12 large instances (with 100 and 200 nodes) of $r$-HMIP to optimality within the given CPU time limit (of 10 hours), and was also significantly faster than the dual based formulation for very large instances (with 200 nodes). The computational efficiency gained for $r$-HMIP by using CACs and Benders decomposition allowed us to further solve large instances of an otherwise intractable $u$-HMPP.

The current work opens up a number of exciting possibilities for future research. One of the natural extensions of the problems studied in this paper is the single allocation versions of $r$-HMIP and $u$-HMPP, which are more challenging to solve than their multiple allocation counterparts. The main challenge in solving the single allocation versions arises from the fact that the lower level problem in $r$-HMIP is an MIP, which eliminates the possibility of reducing it to single level using the dual method. Further, the CACs presented for multiple allocation $r$-HMIP are not directly applicable to its single allocation counterpart since the flows between any O-D pair $(i, j)$ in this case need not necessarily be routed through the least cost available path due to the restriction that any node should be assigned to a single hub. Another interesting extension of our study is to consider the possibility of interdiction at the design of the hub network itself. As discussed in Section 2, Parvaresh et al. (2013, 2014) are the only two papers to our knowledge to have studied this problem. However, both these papers use heuristic based solution approaches. We see exact solution approaches for this problem as an interesting research avenue. Yet another interesting extension is to incorporate uncertainties in the problem parameters (like demand, etc.). All these problems can further be extended to their capacitated versions.

**References**

**References**

Aksen, D., Akca, S. Ş., Aras, N., 2014. A bilevel partial interdiction problem with capacitated facilities and demand outsourcing. Computers & Operations Research 41, 346–358.

Aksen, D., Aras, N., 2012. A bilevel fixed charge location model for facilities under imminent attack. Computers & Operations Research 39 (7), 1364–1381.

Aksen, D., Aras, N., Piyade, N., 2013. A bilevel p-median model for the planning and protection of critical facilities. Journal of Heuristics 19 (2), 373–398.

Aksen, D., Piyade, N., Aras, N., 2010. The budget constrained r-interdiction median problem with capacity expansion. Central European Journal of Operations Research 18 (3), 269–291.

Albert, R., Jeong, H., Barabasi, A., 2002. Error and attack tolerance of complex networks. Nature 406, 378–382.

Alumur, S., Kara, B. Y., 2008. Network hub location problems: The state of the art. European journal of operational research 190 (1), 1–21.

An, Y., Zhang, Y., Zeng, B., 2015. The reliable hub-and-spoke design problem: Models and algorithms. Transportation Research Part B: Methodological 77, 103–122.

Assimakopoulos, N., 1987. A network interdiction model for hospital infection control. Computers in Biology and Medicine 17 (6), 413–422.

Audet, C., Hansen, P., Jaumard, B., Savard, G., 1997. Links between linear bilevel and mixed 0–1 programming problems. Journal of optimization theory and applications 93 (2), 273–300.

Azizi, N., Chauhan, S., Salhi, S., Vidyarthi, N., 2016. The impact of hub failure in hub-and-spoke networks: Mathematical formulations and solution techniques. Computers & Operations Research 65, 174–188.

Benders, J., 1962. Partitioning procedures for solving mixed integer programming problems. Numerische Mathematik 52, 238–252.

Berman, O., Drezner, Z., Tamir, A., Wesolowsky, G. O., 2009. Optimal location with equitable loads. Annals of Operations Research 167 (1), 307–325.

Brown, G., Carlyle, M., Salmern, J., Wood, K., 2006. Defending critical infrastructure. Interfaces 36 (6), 530–544.

Campbell, J. F., 1994. Integer programming formulations of discrete hub location problems. European Journal of Operational Research 72 (2), 387–405.

Campbell, J. F., 1996. Hub location and the p-hub median problem. Operations Research 44 (6), 923–935.

Campbell, J. F., O'Kelly, M. E., 2012. Twenty-five years of hub location research. Transportation Science 46 (2), 153–169.

Cappanera, P., Scaparra, M. P., 2011. Optimal allocation of protective resources in shortest-path networks. Transportation Science 45 (1), 64–80.

Chaharsooghi, S., Momayezi, F., Ghaffarinasab, N., 2017. An adaptive large neighborhood search heuristic for solving the reliable multiple allocation hub location problem under hub disruptions. International Journal of Industrial Engineering Computations 8 (2), 191–202.

Church, R., Scaparra, M. P., 2007a. Analysis of facility systems reliability when subject to attack or a natural disaster. In: Critical Infrastructure. Springer Berlin Heidelberg, pp. 221–241.

Church, R. L., Cohon, J. L., 1976. Multiobjective location analysis of regional energy facility siting problems. Tech. rep., Brookhaven National Lab., Upton, NY (USA).

Church, R. L., Scaparra, M. P., 2007b. Protecting critical assets: the r-interdiction median problem with fortification. Geographical Analysis 39 (2), 129–146.

Church, R. L., Scaparra, M. P., Middleton, R. S., 2004. Identifying critical infrastructure: the median and covering facility interdiction problems. Annals of the Association of American Geographers 94 (3), 491–502.

Collado, R., Papp, D., 2012. Network interdiction - models, applications and unexplored directions. Research report RUTCOR 4 - 2012, Rutgers University.

Contreras, I., Cordeau, J.-F., Laporte, G., 2011a. Benders decomposition for large scale. Operations Research 59, 1477–1490.

Contreras, I., Cordeau, J.-F., Laporte, G., 2011b. Benders decomposition for large-scale uncapacitated hub location. Operations research 59 (6), 1477–1490.

Contreras, I., Cordeau, J.-F., Laporte, G., 2011c. The dynamic uncapacitated hub location problem. Transportation Science 45 (1), 18–32.

Contreras, I., Fernández, E., Marín, A., 2010. The tree of hubs location problem. European Journal of Operational Research 202 (2), 390–400.

Contreras, I., Tanash, M., Vidyarthi, N., 2016. Exact and heuristic approaches for the cycle hub location problem. Annals of Operations Research, 1–23.

Corley, H., Sha, D. Y., Sep. 1982. Most vital links and nodes in weighted networks. Oper. Res. Lett. 1 (4), 157–160.

Cormican, K. J., Morton, D. P., Wood, R. K., 1998. Stochastic network interdiction. Operations Research 46 (2), 184–197.

de Camargo, R. S., de Miranda Jr, G., Luna, H. P. L., 2008. Benders decomposition for the uncapacitated multiple allocation hub location problem. Computers & Operations Research 35 (4), 1047–1064.

de Camargo, R. S., de Miranda Jr, G., Luna, H. P. L., 2009. Benders decomposition for hub location problems with economies of scale. Transportation Science 43 (1), 86–97.

Dobson, G., Karmarkar, U. S., 1987. Competitive location on a network. Operations Research 35 (4), 565–574.

Ebery, J., Krishnamoorthy, M., Ernst, A., Boland, N., 2000. The capacitated multiple allocation hub location problem: Formulations and algorithms. European Journal of Operational Research 120 (3), 614–631.

Elhedhli, S., Hu, F. X., 2005. Hub-and-spoke network design with congestion. Computers & Operations Research 32 (6), 1615–1632.

Ernst, A. T., Krishnamoorthy, M., 1996. Efficient algorithms for the uncapacitated single allocation p-hub median problem. Location science 4 (3), 139–154.

Espejo, I., Marín, A., Rodríguez-Chía, A. M., 2012. Closest assignment constraints in discrete location problems. European Journal of Operational Research 219 (1), 49–58.

Farahani, R. Z., Hekmatfar, M., Arabani, A. B., Nikbakhsh, E., 2013. Hub location problems: A review of models, classification, solution techniques, and applications. Computers & Industrial Engineering 64 (4), 1096–1109.

Farley, J. D., 2003. Breaking al qaeda cells: A mathematical analysis of counterterrorism operations (a guide for risk assessment and decision making). Studies in Conflict & Terrorism 26 (6), 399–411.

Frangioni, A., 1995. On a new class of bilevel programming problems and its use for reformulating mixed integer problems. European journal of operational research 82 (3), 615–646.

Gerrard, R. A., Church, R. L., 1996. Closest assignment constraints and location models: Properties and structure. Location Science 4 (4), 251 – 270.

Ghaffarinasab, N., Atayi, R., 2017. An implicit enumeration algorithm for the hub interdiction median problem with fortification. European Journal of Operational Research.

Ghaffarinasab, N., Motallebzadeh, A., 2017. Hub interdiction problem variants: Models and metaheuristic solution algorithms. European Journal of Operational Research.

Gutfraind, A., Hagberg, A., Pan, F., 2009. Optimal interdiction of unreactive markovian evaders. In: Integration of AI and OR techniques in constraint programming for combinatorial optimization problems. Springer Berlin Heidelberg, pp. 102–116.

Hamacher, H. W., Labbé, M., Nickel, S., Sonneborn, T., 2004. Adapting polyhedral properties from facility to hub location problems. Discrete Applied Mathematics 145 (1), 104–116.

Israeli, E., Wood, R. K., 2002. Shortest-path network interdiction. Networks 40 (2), 97–111.

Jayaswal, S., Vidyarthi, N., 2013. Capacitated multiple allocation hub location with service level constraints for multiple consignment classes. Tech. Rep. WP2013-11-02, Indian Institute of Management Ahmedabad, Research and Publication Department.

Lei, T. L., 2013. Identifying critical facilities in hub-and-spoke networks: A hub interdiction median problem. Geographical Analysis 45 (2), 105–122.

Lewis, T. G., 2006. Critical Infrastructure Protection in Homeland Security: Defending a Networked Nation. Wiley-Interscience.

Liberatore, F., Scaparra, M. P., Daskin, M. S., 2011. Analysis of facility protection strategies against an uncertain number of attacks: the stochastic r-interdiction median problem with fortification. Computers & Operations Research 38 (1), 357–366.

Liberatore, F., Scaparra, M. P., Daskin, M. S., 2012. Hedging against disruptions with ripple effects in location analysis. Omega 40 (1), 21–30.

Lim, C., Smith, J. C., 2007. Algorithms for discrete and continuous multicommodity flow network interdiction problems. IIE Transactions 39 (1), 15–26.

Losada, C., Scaparra, M. P., Church, R. L., 2010. On a bi-level formulation to protect uncapacitated p-median systems with facility recovery time and frequent disruptions. Electronic Notes in Discrete Mathematics 36, 591–598.

Losada, C., Scaparra, M. P., Church, R. L., Daskin, M. S., 2012. The stochastic interdiction median problem with disruption intensity levels. Annals of Operations Research 201 (1), 345–365.

McMasters, A. W., Mustin, T. M., 1970. Optimal interdiction of a supply network. Naval Research Logistics Quarterly 17 (3), 261–268.

Morton, D. P., Pan, F., Saeger, K. J., 2007. Models for nuclear smuggling interdiction. IIE Transactions 39 (1), 3–14.

O'Kelly, M., 2015. Network hub structure and resilience. Networks and Spatial Economics 15, 235–251.

O'Kelly, M. E., 1986. The location of interacting hub facilities. Transportation science 20 (2), 92–106.

O'Kelly, M. E., Bryan, D., 1998. Hub location with flow economies of scale. Transportation Research Part B: Methodological 32 (8), 605–616.

Pan, F., Morton, D. P., 2008. Minimizing a stochastic maximum-reliability path. Networks 52 (3), 111–119.

Parvaresh, F., Hashemi Golpayegany, S., Moattar Husseini, S., Karimi, B., 2013. Solving the p-hub median problem under intentional disruptions using simulated annealing. Networks and Spatial Economics 13 (4), 445–470.

Parvaresh, F., Husseini, S. M., Golpayegany, S. H., Karimi, B., 2014. Hub network design problem in the presence of disruptions. Journal of Intelligent Manufacturing 25 (4), 755–774.

Savelsbergh, M., 1994. Preprocessing and probing techniques for mixed integer programming problems. ORSA Journal on Computing 6 (4), 445–454.

Scaparra, M. P., Church, R., 2012. Protecting supply systems to mitigate potential disaster a model to fortify capacitated facilities. International Regional Science Review 35 (2), 188–210.

Scaparra, M. P., Church, R. L., 2008a. A bilevel mixed-integer program for critical infrastructure protection planning. Computers & Operations Research 35 (6), 1905–1923.

Scaparra, M. P., Church, R. L., 2008b. An exact solution approach for the interdiction median problem with fortification. European Journal of Operational Research 189 (1), 76–92.

Sinha, A., Malo, P., Deb, K., 2017. A review on bilevel optimization: From classical to evolutionary approaches and applications. IEEE Transactions on Evolutionary Computation (Forthcoming) DOI 10.1109/TEVC.2017.2712906, 1–20.

Skorin-Kapov, D., Skorin-Kapov, J., O'Kelly, M., 1996. Tight linear programming relaxations of uncapacitated p-hub median problems. European Journal of Operational Research 94 (3), 582–593.

Smith, J. C., Lim, C., Sudargho, F., 2007. Survivable network design under optimal and heuristic interdiction scenarios. Journal of Global Optimization 38 (2), 181–199.

Song, B. D., Morrison, J. R., Ko, Y. D., 2013. Efficient location and allocation strategies for undesirable facilities considering their fundamental properties. Computers & Industrial Engineering 65 (3), 475–484.

Tanash, M., Contreras, I., Vidyarthi, N., 2017. An exact algorithm for the modular hub location problem with single assignments. Computers and Operations Research 85, 32–44.

Vatsa, A. K., Jayaswal, S., 2016. A new formulation and benders decomposition for the multi-period maximal covering facility location problem with server uncertainty. European Journal of Operational Research 251 (2), 404–418.

Wagner, J., Falkson, L., 1975. The optimal nodal location of public facilities with price-sensitive demand. Geographical Analysis 7 (1), 69–83.

Washburn, A., Wood, K., 1995. Two-person zero-sum games for network interdiction. Operations research 43 (2), 243–251.

Wood, R. K., 1993. Deterministic network interdiction. Mathematical and Computer Modelling 17 (2), 1–18.