INDIAN INSTITUTE OF MANAGEMENT
AHMEDABAD • INDIA

Research and Publications

Alternate Second Order Conic Programming Reformulations for Hub Location with Capacity Selection under Demand Uncertainty

Sneha Dhyani
Sachin Jayaswal
Ankur Sinha
Navneet Vidyarthi

W.P. No. 2018-12-04
December 2018

INDIAN INSTITUTE OF MANAGEMENT
AHMEDABAD – 380015
INDIA

Alternate Second Order Conic Programming Reformulations for Hub Location with Capacity Selection under Demand Uncertainty

Sneha Dhyani
Sachin Jayaswal
Ankur Sinha
Navneet Vidyarthi

Abstract

In this paper, we study the single allocation hub location problem with capacity selection in the presence of congestion at hubs. Accounting for congestion at hubs leads to a non-linear mixed integer program, for which we propose 18 alternate mixed integer second order conic program (MISOCP) reformulations. Based on our computational studies, we identify the best MISOCP-based reformulation, which turns out to be $20-60$ times faster than the state-of-the-art. Using the best MISOCP-based reformulation, we are able to exactly solve instances up to 50 nodes in less than half-an-hour. We also theoretically examine the dimensionality of the second order cones associated with different formulations, based on which their computational performances can be predicted. Our computational results corroborate our theoretical findings. Such insights can be helpful in the generation of efficient MISOCPs for similar classes of problems.
Keywords: Hub-and-Spoke Network; Congestion; Capacity Selection; Stochastic Demand; Single Allocation; Second Order Conic Programming

1 Introduction

Hub-and-spoke is a widely studied network structure, which finds applications in supply chain networks, airline networks, telecommunications, postal deliveries, etc. The key idea behind a hub-and-spoke network is to route all flows through intermediate facilities, called hubs, where they are aggregated before being sent to their respective destinations. Hubs serve as centres to collect, sort, break-bulk or switch modes of travel while transferring flows. The main cost advantage in a hub-and-spoke network comes from the economies of scale in inter-hub transfers achieved due to aggregation of flows. Since hubs perform activities like sorting and breaking bulk for large volumes of goods, any variation in the demand or service rate adversely affects the service quality at the hubs due to congestion. One of the ways often sought to alleviate congestion at hubs is capacity expansion. Considering capacity expansion post network design is often expensive or may even be infeasible in some cases. Therefore, considering capacity selection in anticipation of congestion due to uncertainties at the design phase itself may result in a very different network design, which might be a better alternative (1).

In this paper, we study a hub location problem with capacity selection in the presence of congestion due to demand uncertainty. The problem is relevant for a network designer, who plans to transport goods among multiple origins and destinations at a minimum cost. The designer needs to identify locations, called hubs, through which she will route all the flows and benefit from the economies of scale due to large volumes flowing through fewer links. In light of the expected congestion due to variation in the arrival rates of incoming demand and service rates of hubs, the designer also needs to decide on the trade-off between installing a higher service capacity at the hub, hence minimizing the effect of congestion, or installing a lower capacity, and bearing the high cost due to congestion.

We model the hub-and-spoke network as spatially distributed M/G/1 queues, whose locations and capacities need to be selected in order to minimize the total cost. The total cost consists of the capacity installation cost, the transportation cost, and the congestion cost. The congestion term introduces non-linearity in the objective function, which makes the resulting hub location problem with capacity selection under congestion a non-linear mixed integer program (NLMIP). Hub location problems, even without capacity selection decision and congestion, is known to be NP-hard (2). Capacity selection decision along with the non-linearity introduced due to congestion makes the problem even more challenging. The objective of this paper is to solve the resulting problem efficiently. To this end, we present several alternate MISOCP-based reformulations of the problem, which are solved directly using the state-of-the-art solvers, and compare their computational performances against the Mixed Integer Linear Programming (MILP) based reformulations, obtained using Outer-Approximations (OA).

Through this paper, we make the following contributions to the literature on hub location problems. First, we propose two new NLMIP-based formulations for the hub location problem with capacity selection under congestion. Our new formulations are built on the basic model (without capacity selection and congestion) proposed by (3). We refer to these models as EK-based models. We compare our proposed formulations with two other NLMIP-based formulations from the literature (4; 1), which are based on the well-known model proposed by (5). We refer to these models as SK-based models. We subsequently show, through computational experiments, that the models proposed by us significantly outperform the latter two formulations. Second, we present nine different MISOCP-based reformulations for each of the SK-based and EK-based NLMIPs. From our extensive computational experiments using two of the well-known datasets, namely the Civil Aeronautics Board (CAB) dataset and the Australian Post (AP) dataset, we suggest the overall best formulation of the problem. The best reformulation, which is one of the EK-based MISOCPs, solves the problem 20-60 times faster as compared to the existing formulation/method in the literature. We further provide insights about the reformulations based on the computational results and the properties of the second order cones. These insights should be useful as a general guideline for the selection of a given MISOCP from among several alternatives.

The rest of the paper is organized as follows. In Section 2, we present a review of the literature on the hub location problem and its variants. The problem description, followed by its different NLMIP-based formulations are presented in Section 3. In Section 4, we present our alternate MISOCP-based reformulations of the problem, followed by extensive computational results in Section 5. Finally, the conclusions and directions for future research are presented in Section 6.

2 Literature Review

In the literature review, we first discuss the basic hub location models and their variants, followed by their capacitated versions. We then review the literature on hub location models with multiple capacity levels, wherein the capacity level choice is also a part of the decision. We finally discuss models with congestion at hubs.

2.1 Hub Location Models and Basics

Hub location problems have been broadly categorized as p-hub median, hub location with fixed costs, p-hub center and p-hub covering problems. Another classification is based on capacity restrictions. Both uncapacitated and capacitated versions are further classified based on allocation decisions of non hub nodes to hub nodes. Hence, there is- *Uncapacitated Hub Location Problem with Single Allocation* (UHLPSA), *Uncapacitated Hub Location Problem with Multiple Allocation* (UHLPMA), *Capacitated Hub Location Problem with Single Allocation* (CHLPSA) and *Capacitated Hub Location Problem with Multiple Allocation* (CHLPMA). Some extensions to the model have also been made by relaxing the assumption of constant discount factor during inter-hub transfer. (6), (7), (8), (9) modelled inter-hub discount as a function of flow between the hubs. Also, unlike most models that considered discount only between hubs, (10), (11), (12), considered the advantage of economies of scale on all links.

The first UHLPSA formulation proposed by (13) was hard as it was quadratic with integer variables. A linearized reformulation of this model was proposed by (14), which was later strengthened by (5). (5) used four-subscripted path based variable defined as x_{ijkm} for flows going from i to j through hubs k and m. (15) proposed a different formulation while observing a hub-and-spoke model for the Australian Post services. They defined a three-subscripted flow based variable x_{ikm} for flows originating from i and flowing through hubs k and m. Their model had lesser number of variables as well as it performed well in the execution time, and is therefore known to be the best formulation in terms of computational time for UHLPSA. Several heuristics were also developed by (16),(17), (18), (19), (20), where the Lagrangian based heuristic proposed by (19) is found to be computationally the most efficient. Similar advancements were seen in the literature related to uncapacitated multiple allocation p-hub median problems UHLPMA, when the first formulation was proposed by (21). Contributions were made in finding exact solutions using explicit enumeration algorithm (22) and Benders decomposition (23), or, heuristics solutions (24; 25; 26). Some contributions were also made by studying the polyhedral properties and tightening the lower bounds by adding valid inequalities (27). For detailed review of hub location problems, one can see the review papers by (28), (29), (30), (31).

2.2 Capacitated Hub Location Models

By its very own design, hubs face huge influx of demand, which often results into an unbalanced distribution of flows with few hubs getting very large volumes while others remaining less utilized. The capacitated version was studied as one of the measures to deal with this issue for both single allocation (32; 33; 15; 34; 35; 14; 36) and multiple allocation models

(25; 27; 26; 37; 38). While most of the studies (14; 37; 25) allowed capacity restrictions only on incoming flows from non hubs to hubs, (11) also considered capacity restrictions on flows on the inter-hub links. Some other models, for example, the arc based formulation proposed by (38), and the modular link capacity based formulation by (39), considered capacity restrictions on both incoming flows as well as flows on inter-hub links. Capacitated versions mostly imposed capacity constraints on hubs alone, but studies by (39) and (40) proposed new formulations and considered capacity restrictions on both hubs as well as inter-hub arcs. (38), and (26) modelled capacity restrictions on hubs as well as on other arcs in general. Another measure to balance the flow in the network was studied by (34), who relaxed the capacity constraints and included a processing time/service time criteria in their objective function.

2.3 Hub Location Models with Multiple Capacity Levels

The above literature on capacitated hub location models considered capacity as given. (41) formulated the capacitated single allocation hub location problem with multiple capacity levels (CSAHLPM), which allowed hubs to choose from different capacity levels that could be procured at different costs. They compared and proposed several valid inequalities and refinements for the multiple capacity level model for each of the CHLPSA formulations proposed by (14), (5) and (15). (42) later imposed an additional balancing requirement constraints to balance the number of assignment of non hub nodes to hubs. In both the studies, they proposed two different formulations, which differed in terms of decision variables. One of the formulations used two-subscripted variables for assignment of non hub nodes to hub nodes and assignment of capacity levels to hubs. The other formulation used a three-subscripted variable for both non hub and capacity level assignment. They found that although the three-subscripted model had a tighter LP relaxation, the two-subscripted formulation performed better computationally. In the multiple capacity level literature, (43) studied the splittable multi-capacity hub location problem (MCHLP) and the non-splittable multi-capacity hub location problem (NMCHLP), depending on whether or not commodities are allowed to be split over multiple paths. They used the formulation proposed by (44) for UHLPMA, and solved it using Bender's decomposition. Another variant of the hub location problem with multiple capacity levels was studied by (45) in a multi-period setting. They proposed a deterministic model, which in addition to the hub location and flow routing decisions, also decided on the initial and incremental capacity levels based on the time periods. They extended the model to a two-stage stochastic setting, where hub location and initial capacity related decisions are made at the first stage, and the non hub node allocation, flow routing and incremental capacity decisions are made at the second stage.

2.4 Hub Location Model with Congestion

Despite offering a balanced flow in the network, CHLPSA does not account for the exponential increase of service delay at the hubs when incoming flows reach its capacity, specially due to variable demand (4). (46) were the first to address the issue of service delays by assuming peak hour arrivals and departure at airport hubs leading to M/D/c queues. They imposed a probabilistic constraint on the number of planes waiting for landing, and used

Tabu search heuristic as a solution method. (47) proposed an extension to the above model by considering heterogeneity among customers, thereby imposing different service level constraint for each customer class. (48) were the first to incorporate congestion in the objective by modelling congestion as a power law function. They used path based formulation with single allocation as proposed by (5), and through piece-wise linearization of the congestion function along with Lagrangian heuristic based on (19), solved the problem exactly. (49) solved the above problem for multiple allocation by using the formulation proposed by (44) along with congestion function modelled as a power law function. They solved the problem using generalized Benders decomposition. Both (48) and (49) assumed capacities at hubs high enough to be treated as uncapacitated. By modelling congestion at hubs in the objective function using a power law function, the model mostly tries to evenly distribute flows among the hubs. However, (4) argue that congestion at a hub arises due to the relative demand compared to its capacity. Therefore, they used a capacitated model, with capacity level to install at hubs as additional decision variables, and incorporated congestion by modelling each potential hub as an M/M/1 queue. The problem was solved using an OA-based reformulation and Lagrangian heuristics. (50) used both power law function and M/M/1 queue approximation to model congestion, and proposed a hybrid algorithm using OA and Benders decomposition to solve the resulting problem. (51) modelled congestion from both user perspective and network owner perspective for a UHLPSA, and solved the problem using generalized Benders decomposition. (52) used the power law model proposed by (48), and solved it using second order conic program formulations, strengthened with perspective cuts. They were the first to use the flow based formulation proposed by (15) for hub location with congestion. (1) and (53) are the two more recent works that use queuing based hub location models. (1) use M/G/1 queue to model congestion at the hubs, and use the path based formulation, as proposed by (5), to model HLP with multiple capacity levels and congestion. They propose an exact OA-based and a genetic algorithm based method to solve the resulting model. (53) models congestion using M/M/x queues, where the number of servers is also a decision variable (to capture multiple capacity levels), and solve the resulting problem using a hybrid of particle swarm optimization and genetic algorithm. They also use the path based formulation for HLP with congestion.

3 Model Formulation

Consider a complete graph $G = (N, A)$, where $N = \{1, \ldots, |N|\}$ represents the set of nodes corresponding to origins/destinations for some traffic/flow W_{ij} between each pair of origin $i \in N$ and destination $j \in N$. It is required that flows must be transported at minimum cost. Since a direct link between each Origin-Destination (O-D) pair increases the cost in the network, a subset of nodes $H \in N$ is chosen as hubs through which all the flows are routed. We assume that all the flows originating at a given node always traverses through the same first hub, irrespective of its final destination, which makes it a single allocation problem. Single allocation hub location problems are relevant to situations in which sorting at the source is not possible (or too costly), so that all shipments are transported from the origin as a whole to the allocated hub, as is typically done in postal or parcel networks (54).

To encourage flows between hubs, a discount, due to economies of scale, is offered for inter-hub transfers, which helps reduce the overall cost. However, high flows and variability in demand and service rates result in congestion at hubs, which affects its service quality.

Poor service quality is modelled as increase in cost for every extra unit of flow entering the hub. One way to address this issue is to install higher capacities that can accommodate larger flows and reduce congestion. Since there is a fixed cost for every capacity level, a trade-off exists between the congestion cost and the capacity installation cost. Therefore, while locating hubs, a decision on capacity level also needs to be considered with an overall objective to minimize the sum total of the flow costs, the hub setup costs, and the congestion costs. We refer to the resulting problem as single allocation hub location problem with multiple capacity levels and congestion (SHLPCC).

We use the following standard modelling assumptions: (i) distances follow triangle inequality; (ii) discount is offered only on inter-hub flows; (iii) inter-hub discount is the same for every pair of hubs; (iv) inter-hub discount is constant, independent of the volume of flows; (v) arcs do not have any capacity restrictions, and do not require any set up cost; (vi) demands originate at the nodes according to a Poisson process, and service times at the hubs follow a general distribution. Assumption of Poisson arrivals and general distribution of service times result in the following expression, using the Pollaczek-Khintchine formula, for the expected number of units of flow ($E[N_k]$) at hub k. Note that the incoming flow at a hub also follows a Poisson distribution (due to the superposition of Poisson distributions).

$$E[N_k] = \left(\frac{1+c_k^2}{2}\right)\frac{\lambda_k^2}{\mu_k(\mu_k-\lambda_k)} + \frac{\lambda_k}{\mu_k} \iff 1/2\left\{\left(1+c_k^2\right)\frac{\lambda_k}{\left(\mu_k-\lambda_k\right)} + \left(1-c_k^2\right)\frac{\lambda_k}{\mu_k}\right\}, \quad (1)$$

where λ_k is the mean of the incoming Poisson flows, μ_k is the mean service rate, and c_k is the coefficient of variation of service times.

Sets and Indices		
i	:	Index for origin nodes, $i \in N$;
j	:	Index for destination nodes, $j \in N$;
k, m	:	Index for potential hubs, $k, m \in N$;
l	:	Index for capacity levels, $l \in L$;
Parameters		
p	:	Number of hubs to be opened;
W_{ij}	:	Mean demand rate from origin i to destination j;
d_{ij}	:	Distance from origin i to destination j where $d_{ij} = d_{ik} + d_{km} + d_{mj}$;
χ	:	Collection cost - cost per unit flow per unit distance from non hub node to hub node;
δ	:	Distribution cost - cost per unit flow per unit distance from hub node to non hub node;
α	:	Transfer cost/Discount - cost per unit flow per unit distance between two hubs. $0 \leq \alpha \leq 1, \alpha < \chi, \alpha < \delta$;
γ_k^l	:	Capacity (service rate) at hub k with capacity level l;
c_{kl}	:	Coefficient of variation of service times of a hub k with capacity levels l;
Q_k^l	:	Fixed cost for installing hub at node k with capacity level l;
$O_i = \sum_j W_{ij}$	:	Total flow originating at i;
$D_i = \sum_j W_{ji}$	:	Total flow reaching i;
θ	:	Congestion cost per unit user at hub k;

We use two different ways to assign capacity levels to hubs, similar to (42), for (15) base model. In the first scheme, we use one set of binary variables for assignment of non hub nodes to hubs, and another set for assignment of capacity levels to hubs. We refer to this scheme as two-subscripted capacity allocation variable scheme using which the complete NLMIP formulation is presented in Section 3.1. In the second scheme, we use a single set of three-subscripted variables for assignment of non hub nodes to hubs, along with the selection of capacity levels. We refer to this scheme as three-subscripted capacity allocation variable scheme using which the complete NLMIP formulation is described in Section 3.2.

3.1 SHLPCC model based on two-subscripted capacity allocation variable

For SHLPCC with two-subscripted capacity allocation variable for EK-based models, we define the following decision variables.

$$x_{ikm} = \text{Amount of flow originating at } i \text{ and flowing through hubs } k \text{ and } m, \text{ in that order.}$$

$$z_{ik} = \begin{cases} 1, & \text{if node } i \text{ is assigned to hub } k \\ 0, & \text{otherwise.} \end{cases}$$

$$y_{kl} = \begin{cases} 1, & \text{if hub } k \text{ is assigned capacity level } l \\ 0, & \text{otherwise.} \end{cases}$$

Note, z_{kk} assigns a hub k to itself, thus eliminating the use of a separate variable for locating hubs. A hub can have only one capacity level hence $\sum_l y_{kl} = 1$. Therefore, the service rate (μ_k) and the coefficient of variation of service times (c_k), which depend on capacity level, are related to y_{kl} as follows

$$\mu_k = \sum_l \gamma_k^l y_{kl}, \quad c_k^2 = \sum_l c_{kl}^2 y_{kl} \quad \forall k \in N.$$

The mean flow at hub k is given by

$$\lambda_k = \sum_i \sum_m x_{ikm} = \sum_i O_i z_{ik} \quad \forall k \in N.$$

Note that λ_k in the above expression captures only the flows directly entering hub k from the origin nodes, but not the flows entering hub k via some other hub. This is appropriate in situations where the flows require processing (e.g., collecting, sorting, batching, etc.) only at the first hub in their path from their origin to their destination, but do not require further processing at the second hub (25; 47). However, in situations where the flows need further processing before distribution, the mean flow at a hub k is given by

$$\lambda_k = \sum_i \sum_m x_{ikm} + \sum_i \sum_{m \neq k} x_{imk}.$$

Here, the first term includes the flows for which node k is the first hub, while the second term includes the flows that are routed through some other hub m before entering hub k (37; 49).

Substituting λ_k, μ_k and c_k in equation (1) with the appropriate decision variables z_{ik} and y_{kl}, the congestion term modifies to:

$$E[N_k(y,z)] = 1/2\left\{\left(1+\sum_l c_{kl}^2 y_{kl}\right)\frac{\sum_i\sum_j W_{ij}z_{ik}}{\left(\sum_l \gamma_k^l y_{kl} - \sum_i\sum_j W_{ij}z_{ik}\right)} + \left(1-\sum_l c_{kl}^2 y_{kl}\right)\frac{\sum_i\sum_j W_{ij}z_{ik}}{\sum_l \gamma_k^l y_{kl}}\right\}$$

The SHLPCC formulation with two-subscripted capacity allocation variables is given as:

$$\begin{aligned}
\min \quad & \sum_i\sum_k d_{ik}(\chi O_i + \delta D_i)z_{ik} + \sum_i\sum_k\sum_m \alpha d_{km}x_{ikm} \\
& + \sum_k\sum_l Q_k^l y_{kl} + \sum_k \theta E[N_k(y,z)] && (2)\\
\text{s.t.} \ & \sum_k z_{ik} = 1 && \forall i \in N \quad (3)\\
& z_{ik} \le z_{kk} && \forall i \in N, k \in N \quad (4)\\
& \sum_k z_{kk} = p && (5)\\
& \sum_m x_{ikm} - \sum_m x_{imk} = O_i z_{ik} - \sum_j W_{ij} z_{jk} && \forall i \in N, k \in N \quad (6)\\
& \sum_m x_{ikm} \le O_i z_{ik} && \forall i \in N, k \in N \quad (7)\\
& \sum_i\sum_j W_{ij} z_{ik} \le \sum_l \gamma_k^l y_{kl} && \forall k \in N \quad (8)\\
& \sum_l y_{kl} = z_{kk} && \forall k \in N \quad (9)\\
& y_{kl} \in 0,1 && \forall k \in N, l \in L \quad (10)\\
& x_{ikm} \ge 0 && \forall i \in N, k \in N, m \in N \quad (11)\\
& z_{ik} \in 0,1 && \forall i \in N, k \in N \quad (12)
\end{aligned}$$

The first component of the objective function (2) includes the cost of transferring flows from non hub nodes to hub nodes. The second component is the cost of inter-hub transfer, and the third component is the cost of locating hubs at certain capacity levels. Constraint set(3) ensures single-allocation of non hub nodes to hubs. Constraint set (4) prevents assignment of a node to another node unless that latter is a hub. Constraint (5) is to enforce the selection of only p-hubs, and constraint set (6) represents the flow balance constraints. Constraint set (7) prevents any traffic originating at node i from flowing via hub k unless the node i is allocated to hub k. Constraint set (8) ensures that the total flow at hub k does not exceed its installed capacity, which is required for the stability of the queueing system at the open hubs. Note that the queue stability constraint ideally requires a strict inequality in the constraint set (8). However, we exploit the knowledge that the constraint set (8) can never be binding at optimality (since that will make the congestion term in the objective function tend to infinity, which can never be optimal), and retain the $\le$ sign in constraint set (8). Constraint set (9) ensures that a hub k, given by variable z_{kk}, can have only one capacity allocation.

3.1.1 Partial Linearization

(2)-(12) is an NLMIP, which is challenging to solve. Next, we suggest its partial linearization to convert it into a form with a linear objective function and one set of non-linear constraints, which can be further reformulated using outer-approximation and MISOCPs. To this end, we introduce ρ_k and s_k as additional sets of variables, which are defined using the following relations:

$$\rho_k = \frac{\lambda_k}{\mu_k} = \frac{\sum_i \sum_j W_{ij} z_{ik}}{\sum_l \gamma_k^l y_{kl}} \qquad \forall k \in N$$

$$s_k = \frac{\sum_i \sum_j W_{ij} z_{ik}}{\left(\sum_l \gamma_k^l y_{kl} - \sum_i \sum_j W_{ij} z_{ik}\right)} \qquad \forall k \in N \tag{13}$$

$$0 \le \rho_k \le 1, s_k \ge 0 \qquad \forall k \in N. \tag{14}$$

Using the newly defined variables, $E[N_k(y,z)]$ in (2) can be rewritten as:

$$E[N_k(y)] = 1/2\left\{\left(1+\sum_l c_{kl}^2 y_{kl}\right)s_k+\left(1-\sum_l c_{kl}^2 y_{kl}\right)\rho_k\right\} = 1/2\left\{s_k+\rho_k+\sum_l c_{kl}^2 y_{kl}(s_k-\rho_k)\right\}.$$

Introducing auxiliary variables L_{kl} and V_{kl} $\forall k, l$, and defined as $L_{kl} = \rho_k y_{kl}$ and $V_{kl} = s_k y_{kl}$, $E[N_k(y)]$ can be further restated as:

$$E[N_k] = 1/2\left\{s_k + \rho_k + \sum_l c_{kl}^2 (V_{kl} - L_{kl})\right\},$$

where L_{kl} and V_{kl} are non-linear, for which we use the following standard linearization:

$$\sum_l V_{kl} = s_k \qquad \forall k \in N \tag{15}$$

$$V_{kl} \le M y_{kl} \qquad \forall k \in N, l \in L \tag{16}$$

$$V_{kl} \ge 0 \qquad \forall k \in N, l \in L \tag{17}$$

$$\sum_l L_{kl} = \rho_k \qquad \forall k \in N \tag{18}$$

$$L_{kl} \le y_{kl} \qquad \forall k \in N, l \in L \tag{19}$$

$$0 \le L_{kl} \le 1 \qquad \forall k \in N, l \in L. \tag{20}$$

Since $\rho_k = \frac{\lambda_k}{\mu_k} = \frac{\sum_i \sum_j W_{ij} z_{ik}}{\sum_l \gamma_k^l y_{kl}}$, we have:

$$\sum_i \sum_j W_{ij} z_{ik} = \rho_k \sum_l \gamma_k^l y_{kl} = \sum_l \gamma_k^l y_{kl} \rho_k = \sum_l \gamma_k^l L_{kl} \qquad \forall k \in N. \tag{21}$$

Using the above transformations, (2)-(12) can be written as follows:

$$\textbf{[EK-2s]} \quad \min \sum_i \sum_k d_{ik}(\chi O_i + \delta D_i) z_{ik} + \sum_i \sum_k \sum_m \alpha d_{km} x_{ikm} + \sum_k \sum_l Q_k^l y_{kl} + \theta/2 \sum_k s_k + \rho_k + \sum_l c_{kl}^2 (V_{kl} - L_{kl})$$

$$\text{s.t.} \quad (3) - (21).$$

The only non-linearity in EK-2s appears in (13), which can be handled using the well-known OA-based method, as detailed in Appendix B.1. We refer to this OA-based method for the two-subscripted capacity allocation based variable as EK-OA-2s. In Section 4.1, we propose five alternate MISOCP-based reformulations of EK-2s.

3.2 SHLPCC model based on three-subscripted capacity allocation variable

For SHLPCC with three-subscripted capacity allocation variable for the EK-based model, we define the following decision variables.

$$x_{ikm} = \text{Amount of flow with origin at } i \text{ that goes through hubs } k \text{ and } m$$

$$t_{ik}^{l} = \begin{cases} 1, & \text{if node } i \text{ is assigned to hub } k \text{ which has capacity level } l. \\ 0, & \text{otherwise.} \end{cases}$$

The new variable t_{ik}^{l} is related to the previous variables y_{kl} and z_{ik} as follows:

$$z_{ik} = \sum_{l} t_{ik}^{l} \quad \forall i, k. \qquad y_{kl} = t_{kk}^{l} \quad \forall k, l \qquad z_{kk} = \sum_{l} t_{kk}^{l} \tag{22}$$

Since we do not have variable y_{kl}, as in the case with two-subscripted capacity based allocation scheme, we do not encounter complexities arising from $\gamma_k^l y_{kl}$. Therefore, we work with the following two forms of the expected number of users, $E[N_k(t)]$.

$$E[N_k(t)] = \frac{(1 + c_{kl}^2)(\sum_i \sum_j W_{ij} t_{ik}^l)^2}{2(\gamma_k^l)(\gamma_k^l - \sum_i \sum_j W_{ij} t_{ik}^l)} + \frac{\sum_i \sum_j W_{ij} t_{ik}^l}{\gamma_k^l} \tag{23}$$

$$E[N_k(t)] = 1/2 \left\{ \left(1 + c_{kl}^2\right) \frac{\sum_i \sum_j W_{ij} t_{ik}^l}{\left(\gamma_k^l - \sum_i \sum_j W_{ij} t_{ik}^l\right)} + \left(1 - c_{kl}^2\right) \frac{\sum_i \sum_j W_{ij} t_{ik}^l}{\gamma_k^l} \right\} \tag{24}$$

The SHLPCC formulation with three-subscripted capacity allocation variable is given as:

[EK-3s]

$$\begin{aligned}
\min \quad & \sum_i \sum_k d_{ik}(\chi O_i + \delta D_i) \sum_l t^l_{ik} + \sum_i \sum_k \sum_m \alpha d_{km} x_{ikm} \\
& + \sum_k \sum_l Q^l_k t^l_{kk} + \sum_k \sum_l \theta E[N_k(t)] && (25) \\
\text{s.t.} \quad & \sum_k \sum_l t^l_{ik} = 1 && \forall i, j \in N \quad (26) \\
& t^l_{ik} \le t^l_{kk} && \forall i, k \in N, l \in L \quad (27) \\
& \sum_k \sum_l t^l_{kk} = p && (28) \\
& \sum_m x_{ikm} - \sum_m x_{imk} = O_i \sum_l t^l_{ik} - \sum_j W_{ij} \sum_l t^l_{jk} && \forall i, k \in N \quad (29) \\
& \sum_m x_{ikm} \le O_i \sum_l t^l_{ik} && \forall i, k \in N \quad (30) \\
& \sum_i \sum_j W_{ij} t^l_{ik} \le \gamma^l_k && \forall k \in N, l \in L \quad (31) \\
& \sum_l t^l_{kk} \le 1 && \forall k \in N \quad (32) \\
& x_{ikm} \ge 0 && \forall i, k, m \in N \quad (33) \\
& t^i_{kl}, t^i_{kk} \in 0, 1 && \forall i, k \in N, l \in L \quad (34)
\end{aligned}$$

Constraints (26) - (34) are straightforward conversions of constraints (3) - (12) by using the relationships shown in (22). Note that constraint (27) is a disaggregated version of constraint (4), as $z_{ik} \le z_{kk} \iff \sum_l t^l_{ik} \le \sum_l t^l_{kk}$, which can be disaggregated $\forall\ l \in L$. For the non-linear terms, $\frac{(\sum_i \sum_j W_{ij} t^l_{ik})^2}{\gamma^l_k - \sum_i \sum_j W_{ij} t^l_{ik}}$ and $\frac{\sum_i \sum_j W_{ij} t^l_{ik}}{\gamma^l_k - \sum_i \sum_j W_{ij} t^l_{ik}}$ in (23) and (24), we propose MISOCP-based reformulations in Section 4.2. The OA-based method for EK-3s, which we refer to as EK-OA-3s, is discussed in Appendix B.2.

We do a similar study on SHLPCC with the two-subscripted and the three-subscripted capacity allocation variables for the SK-based model, and propose MISOCP-based reformulations in Appendix A. In the following sections, we discuss our MISOCP-based reformulations and analysis for EK-2s and EK-3s models.

4 MISOCP-based Reformulations

In this section, we briefly describe second order conic programs. In Section 4.1, we propose MISOCP-based reformulations for EK-2s, followed by MISOCP-based reformulations for EK-3s in Section 4.2.

A Second Order Conic Program is a convex optimization problem of the following form:

$$\begin{aligned}
\min_{x \in \mathbb{R}^n} \quad & f^T x \\
\text{s.t.} \quad & \left\| A_i x + b_i \right\| \le c_i^T x + d_i, \quad \forall i = 1 \dots N && (35)
\end{aligned}$$

Each of the constraints in (35) is a second-order cone constraint of dimension n_i. Here, $f \in \mathbb{R}^n$, $A_i \in \mathbb{R}^{(n_i-1)n}$, $b_i \in \mathbb{R}^{n_i-1}$, $c_i \in \mathbb{R}^n$, $d_i \in \mathbb{R}$. Note that when $A_i = 0, \forall i = 1, \ldots, N$, the above SOCP reduces to an LP, while $c_i = 0, \forall i = 1, \ldots N$ reduces it to a Quadratically Constrained Quadratic Programming (QCQP).

For any positive-semi definite matrix Q ($Q \succcurlyeq 0$), i.e. when $x^T Q x \geq 0$ for all x, CPLEX accepts second order conic constraints in the following two forms:

$$x^T Q x \leq y^2 \qquad y \geq 0 \tag{Form-1}$$

$$x^T Q x \leq yz \qquad y, z \geq 0 \tag{Form-2}$$

Form-2 constraint, also called hyperbolic constraint, can be written as SOCs by the following two transformations:

$$2x^T Q x + y^2 + z^2 \leq (y+z)^2 \qquad y, z \geq 0 \tag{Form-2.1}$$

$$4x^T Q x + (y-z)^2 \leq (y+z)^2 \qquad y, z \geq 0 \tag{Form-2.2}$$

From our initial experiments, and as also illustrated by (55), constraints of the (Form-2.1) perform the best. We have, therefore, represented all the constraint of Form-2 as (Form-2.1) in this paper.

4.1 MISOCP-based Reformulations for EK-2s

In the model EK-2s (based on two-subscripted capacity allocation variables), the non-linear term (13) includes two decision variables y_{kl} and z_{ik}. In this section, we propose MISOCP-based reformulations based on both the variables. The first SOC is based on z_{ik} variables, the next three SOCs are based on y_{kl} variables, while the fifth SOC is based on the relationship between the traffic intensity ρ_k and s_k.

EK-MISOCP1: For our first reformulation, since $s_k \geq \frac{\sum_i \sum_j W_{ij} z_{ik}}{\sum_l \gamma_k^l y_{kl} - \sum_i \sum_j W_{ij} z_{ik}}$, and z_{ik} is binary, we have $\sum_i \left(\sum_j W_{ij} \right) z_{ik}^2 \leq s_k (\sum_l \gamma_k^l y_{kl} - \sum_i \sum_j W_{ij} z_{ik})$. We introduce variable t_k such that $(\sum_l \gamma_k^l y_{kl} - \sum_i \sum_j W_{ij} z_{ik}) \geq t_k$. The constraint holds with inequality for the minimization problem. On substituting, we get the following set of constraints,

$$\sum_i \sum_j W_{ij} z_{ik}^2 \leq s_k t_k \quad \forall k \in N, \tag{36}$$

$$\sum_l \gamma_k^l y_{kl} - \sum_i \sum_j W_{ij} z_{ik} \geq t_k \quad \forall k \in N, \tag{37}$$

$$t_k \geq 0 \qquad \forall k \in N. \tag{38}$$

Note that (37) dominates (8), hence (8) is eliminated. Since (36) are hyperbolic constraint, it is transformed to (Form-2.1) as:

$$2 \sum_i \sum_j W_{ij} z_{ik}^2 + s_k^2 + t_k^2 + \leq (s_k + t_k)^2 \quad \forall k \in N. \tag{39}$$

Our first EK-MISOCP-based reformulation is then given as follows:

$$\begin{aligned}
\textbf{[EK-MISOCP1]} \quad \min \quad & \sum_i \sum_k d_{ik}(\chi O_i + \delta D_i) z_{ik} + \sum_i \sum_k \sum_m \alpha d_{km} x_{ikm} \\
& + \sum_k \sum_l Q_k^l y_{kl} + \theta/2 \sum_k s_k + \rho_k + \sum_l c_{kl}^2 (V_{kl} - L_{kl}) \\
\text{s.t.} \quad & (3) - (7), (9) - (12), (14) - (21), (37) - (39).
\end{aligned}$$

EK-MISOCP1 has $2N$ additional constraints and N additional variables, out of which N constraints are SOCs, each of dimension $N + 3$.

EK-MISOCP2: We next propose SOC constraints based on y_{kl} variables. Similar to the previous formulation, we have variable s_k such that,

$$\frac{\sum_i \sum_j W_{ij} z_{ik}}{\sum_l \gamma_k^l y_{kl} - \sum_i \sum_j W_{ij} z_{ik}} \leq s_k \iff \sum_i \sum_j W_{ij} z_{ik} \leq s_k (\sum_l \gamma_k^l y_{kl} - \sum_i \sum_j W_{ij} z_{ik}).$$

Adding $(\sum_l \gamma_k^l y_{kl} - \sum_i \sum_j W_{ij} z_{ik})$ on both sides we get $\sum_l \gamma_k^l y_{kl} \leq (1 + s_k)(\sum_l \gamma_k^l y_{kl} - \sum_i \sum_j W_{ij} z_{ik})$. We introduce auxiliary variables t_k and τ_k such that $(\sum_l \gamma_k^l y_{kl} - \sum_i \sum_j W_{ij} z_{ik}) \geq t_k$ and $\tau_k = 1 + s_k$. Also, using the fact that y_{kl} is binary, we have the following constraints

$$\sum_l \gamma_k^l y_{kl}^2 \leq \tau_k t_k \qquad \forall k \in N, \tag{40}$$

$$\sum_l \gamma_k^l y_{kl} - \sum_i \sum_j W_{ij} z_{ik} \geq t_k \qquad \forall k \in N, \tag{41}$$

$$\tau_k = 1 + s_k \qquad \forall k \in N, \tag{42}$$

$$t_k, \tau_k \geq 0 \qquad \forall k \in N. \tag{43}$$

Hyperbolic constraint (40) are transformed to

$$2 \sum_l \gamma_k^l y_{kl}^2 + \tau_k^2 + t_k^2 \leq (\tau_k + t_k)^2 \qquad \forall k \in N. \tag{44}$$

The second EK-MISOCP-based reformulation is, therefore:

$$\begin{aligned}
\textbf{[EK-MISOCP2]} \quad \min \quad & \sum_i \sum_k d_{ik}(\chi O_i + \delta D_i) z_{ik} + \sum_i \sum_k \sum_m \alpha d_{km} x_{ikm} \\
& + \sum_k \sum_l Q_k^l y_{kl} + \theta/2 \sum_k s_k + \rho_k + \sum_l c_{kl}^2 (V_{kl} - L_{kl}) \\
\text{s.t.} \quad & (3) - (7), (9) - (12), (14) - (21), (41) - (44).
\end{aligned}$$

EK-MISOCP2 has $3N$ additional constraints and $2N$ additional variables, including N SOCs, each of dimension L+3.

EK-MISOCP3: Disaggregating (40) results in higher number of SOCs of smaller dimensions, which gives us a new set of constraints as follows:

$$\gamma_k^l y_{kl}^2 \leq \tau_k t_k \quad \forall k \in N, l \in L. \tag{45}$$

Transforming to (Form-2.1), we have

$$2\gamma_k^l y_{kl}^2 + \tau_k^2 + t_k^2 \leq (\tau_k + t_k)^2 \quad \forall k \in N, l \in L. \tag{46}$$

Our third EK-MISOCP-based reformulation is, therefore,

$$
\begin{aligned}
\textbf{[EK-MISOCP3]} \quad \min \quad & \sum_i \sum_k d_{ik}(\chi O_i + \delta D_i) z_{ik} + \sum_i \sum_k \sum_m \alpha d_{km} x_{ikm} \\
& + \sum_k \sum_l Q_k^l y_{kl} + \theta/2 \sum_k s_k + \rho_k + \sum_l c_{kl}^2 (V_{kl} - L_{kl}) \\
\text{s.t.} \quad & (3) - (7), (9) - (12), (14) - (21), (41) - (43), (46).
\end{aligned}
$$

EK-MISOCP3 has $N + NL$ additional constraints and $2N$ additional variables, out of which NL constraints are SOCs, each of dimension 4.

EK-MISOCP4: Another slight variation results in SOCs of smaller dimension as compared to (40). We define p_k, such that,

$$p_k^2 = \sum_l \gamma_k^l y_{kl}^2 \iff p_k = \sum_l \sqrt{\gamma_k^l} y_{kl} \quad \forall k \in N. \tag{47}$$

Substituting in (40) we get,

$$p_k^2 \le \tau_k t_k \quad \forall k \in N, \tag{48}$$

$$p_k \ge 0 \quad \forall k \in N. \tag{49}$$

Transformation to (Form-2.1) results in the following constraint

$$2p_k^2 + \tau_k^2 + t_k^2 \le (\tau_k + t_k)^2 \quad \forall k \in N. \tag{50}$$

Our fourth EK-MISOCP-based reformulation is as follows:

$$
\begin{aligned}
\textbf{[EK-MISOCP4]} \quad \min \quad & \sum_i \sum_k d_{ik}(\chi O_i + \delta D_i) z_{ik} + \sum_i \sum_k \sum_m \alpha d_{km} x_{ikm} \\
& + \sum_k \sum_l Q_k^l y_{kl} + \theta/2 \sum_k s_k + \rho_k + \sum_l c_{kl}^2 (V_{kl} - L_{kl}) \\
\text{s.t.} \quad & (3) - (7), (9) - (12), (14) - (21), (41) - (43), (47), (49), (50).
\end{aligned}
$$

EK-MISOCP4 has $4N$ additional constraints and $3N$ additional variables, respectively, out of which there are N SOCs, each of dimension 4.

EK-MISOCP5: From the definition of s_k and ρ_k, as used earlier, we can write $\frac{\rho_k}{1-\rho k} \le s_k \iff \rho_k \le s_k(1 - \rho_k)$. Adding $(1 - \rho_k)$ on both the sides we get the following hyperbolic constraint

$$1 \le (1 + s_k)(1 - \rho_k) \qquad \forall k \in N. \tag{51}$$

As stated in Section 4, CPLEX does not accept the above constraints directly unless both expressions $(1+s_k)$ and $(1-\rho_k)$ are explicitly stated non-negative. Therefore, we substitute

$$\tau_k = 1 + s_k \qquad \forall k \in N, \tag{52}$$

$$\psi_k = 1 - \rho_k \qquad \forall k \in N, \tag{53}$$

$$\tau_k, \psi_k \ge 0 \qquad \forall k \in N, \tag{54}$$

which results in the following hyperbolic constraint

$$1 \le \tau_k \psi_k \qquad \forall k \in N. \tag{55}$$

The above hyperbolic constraint is converted to (Form-2.1) as

$$2 + \tau_k^2 + \psi_k^2 \le (\tau_k + \psi_k)^2 \quad \forall k \in N. \tag{56}$$

This is given to the solver directly. Our fifth EK-MISOCP-based reformulation is, therefore,

$$\begin{aligned}
\textbf{[EK-MISOCP5]} \quad \min \quad & \sum_i \sum_k d_{ik}(\chi O_i + \delta D_i) z_{ik} + \sum_i \sum_k \sum_m \alpha d_{km} x_{ikm} \\
& + \sum_k \sum_l Q_k^l y_{kl} + \theta/2 \sum_k s_k + \rho_k + \sum_l c_{kl}^2 (V_{kl} - L_{kl}) \\
\text{s.t.} \quad & (3) - (12), (52) - (54), (56).
\end{aligned}$$

EK-MISOCP5 has $3N$ additional constraints and $2N$ additional variables with N SOC constraints, each of dimension 3.

4.2 MISOCP-based Reformulations for EK-3s

In the model EK-3s (based on three-subscripted capacity allocation variables), the non-linear term in the objective function can expressed in two alternate forms, given by (23) and (24), using which we obtain alternate MISOCP-based reformulations.

EK-MISOCP6: From (23), we have

$$E[N_k(t)] = \frac{(1 + c_{kl}^2)(\sum_i \sum_j W_{ij} t_{ik}^l)^2}{2(\gamma_k^l)(\gamma_k^l - \sum_i \sum_j W_{ij} t_{ik}^l)} + \frac{\sum_i \sum_j W_{ij} t_{ik}^l}{\gamma_k^l}.$$

We introduce variable $r_{kl}, t_{kl} \ge 0$, such that

$$\frac{(\sum_i \sum_j W_{ij} t_{ik}^l)^2}{\gamma_k^l - \sum_i \sum_j W_{ij} t_{ik}^l} \le r_{kl} \iff (\sum_i \sum_j W_{ij} t_{ik}^l)^2 \le r_{kl}(\gamma_k^l - \sum_i \sum_j W_{ij} t_{ik}^l) \quad \forall k, l.$$

Also, $\gamma_k^l - \sum_i \sum_j W_{ij} t_{ik}^l \ge t_{kl} \;\; \forall k, l$, therefore we have the following set of constraints

$$(\sum_i \sum_j W_{ij} t_{ik}^l)^2 \le r_{kl} t_{kl} \qquad \forall k \in N, l \in L, \tag{57}$$

$$\gamma_k^l - \sum_i \sum_j W_{ij} t_{ik}^l \ge t_{kl} \quad \forall k \in N, l \in L, \tag{58}$$

$$t_{kl}, r_{kl} \ge 0 \qquad \forall k \in N, l \in L. \tag{59}$$

Constraint (57), which is hyperbolic constraint, is transformed to (Form-2.1) as:

$$2(\sum_i \sum_j W_{ij} t_{ik}^l)^2 + r_{kl}^2 + t_{kl}^2 \le (r_{kl} + t_{kl})^2 \qquad \forall k \in N, l \in L. \tag{60}$$

Notice that because of minimization objective, both r_{kl} and t_{kl} hold with equality at the optimum. Also (58) dominates (31), hence (31) is eliminated. Our sixth EK-MISOCP-based formulation is as follows:

$$\text{[EK-MISOCP6]} \quad \min \sum_i \sum_k d_{ik}(\chi O_i + \delta D_i) \sum_l t^l_{ik} + \sum_i \sum_k \sum_m \alpha d_{km} x_{ikm} + \sum_k \sum_l Q^l_k t^l_{kk} + \theta \sum_k \sum_l \frac{(1+c^2_{kl})}{2(\gamma^l_k)} r_{kl} + \frac{\sum_i \sum_j W_{ij} t^l_{ik}}{\gamma^l_k}$$
$$\text{s.t.} \quad (26)-(30), (32)-(34), (58)-(60).$$

EK-MISOCP6 has $2NL$ additional constraints and $2NL$ additional variables, out of which there are NL number of SOCs of dimension 4.

EK-MISOCP7: From (24) we have,

$$E[N_k(t)] = 1/2\left\{\left(1+c^2_{kl}\right)\frac{\sum_i \sum_j W_{ij} t^l_{ik}}{\left(\gamma^l_k - \sum_i \sum_j W_{ij} t^l_{ik}\right)} + \left(1-c^2_{kl}\right)\frac{\sum_i \sum_j W_{ij} t^l_{ik}}{\gamma^l_k}\right\}.$$

Introducing variable s_{kl} such that, $\frac{(\sum_i \sum_j W_{ij} t^l_{ik})}{\gamma^l_k - \sum_i \sum_j W_{ij} t^l_{ik}} \le s_{kl}$ and $(\gamma^l_k - \sum_i \sum_j W_{ij} t^l_{ik}) \ge t_{kl}$, we have the following constraints

$$\sum_i \sum_j W_{ij}(t^l_{ik})^2 \le s_{kl} t_{kl} \quad \forall k \in N, l \in L, \tag{61}$$
$$\gamma^l_k - \sum_i \sum_j W_{ij} t^l_{ik} \ge t_{kl} \quad \forall k \in N, l \in L, \tag{62}$$
$$t_{kl}, s_{kl} \ge 0 \qquad \forall k \in N, l \in L. \tag{63}$$

Hyperbolic constraint (61) is transformed to Form-2.1 as

$$2\sum_i \sum_j W_{ij}(t^l_{ik})^2 + s^2_{kl} + t^2_{kl} \le (s_{kl} + t_{kl})^2 \quad \forall k \in N, l \in L. \tag{64}$$

Seventh EK-MISOCP-based reformulation is as follows:

$$\text{[EK-MISOCP7]} \quad \min \sum_i \sum_k d_{ik}(\chi O_i + \delta D_i) \sum_l t^l_{ik} + \sum_i \sum_k \sum_m \alpha d_{km} x_{ikm} + \sum_k \sum_l Q^l_k t^l_{kk} + \theta/2 \sum_k \sum_l \left(1+c^2_{kl}\right) s_{kl} + \left(1-c^2_{kl}\right)\frac{\sum_i \sum_j W_{ij} t^l_{ik}}{\gamma^l_k}$$
$$\text{s.t.} \quad (26)-(30), (32)-(34), (62)-(64).$$

EK-MISOCP7 has $2NL$ additional constraints and $2NL$ additional variables, out of which there are NL SOCs, each of dimension $N+3$.

EK-MISOCP8: We propose another reformulation by using s_k and introducing variable u^l_{ijk} such that,

$$\sum_i \sum_j u^l_{ijk} \le s_{kl} \quad \forall k \in N, l \in L \quad \text{, and} \quad \frac{W_{ij} t^l_{ik}}{\gamma^l_k - \sum_i \sum_j W_{ij} t^l_{ik}} \le u^l_{ijk}.$$

Substituting $\gamma_k^l - \sum_i \sum_j W_{ij} t_{ik}^l \geq t_{kl}$, we get the following set of constraints

$$W_{ij}(t_{ik}^l)^2 \leq u_{ijk}^l t_{kl} \quad \forall i,j,k \in N, l \in L, \tag{65}$$

$$\sum_i \sum_j u_{ijk}^l \leq s_{kl} \quad \forall k \in N, l \in L, \tag{66}$$

$$\gamma_k^l - \sum_i \sum_j W_{ij} t_{ik}^l \geq t_{kl} \quad \forall k \in N, l \in L, \tag{67}$$

$$u_{ijkl} \geq 0 \quad \forall i,j,k \in N, l \in L, \tag{68}$$

$$t_{kl}, s_{kl} \geq 0 \quad \forall k \in N, l \in L. \tag{69}$$

Transformation of hyperbolic constraint (65) to Form-2.1 gives us,

$$2W_{ij}(t_{ik}^l)^2 + (u_{ijk}^l)^2 + t_{kl}^2 \leq (u_{ijk}^l + t_{kl})^2 \forall i,j,k \in N, l \in L. \tag{70}$$

Our eighth EK-MISOCP-based reformulation is as follows:

[EK-MISOCP8]

$$\min \sum_i \sum_k d_{ik}(\chi O_i + \delta D_i) \sum_l t_{ik}^l + \sum_i \sum_k \sum_m \alpha d_{km} x_{ikm}$$
$$+ \sum_k \sum_l Q_k^l t_{kk}^l + \theta/2 \sum_k \sum_l \left(1 + c_{kl}^2\right) s_{kl} + \left(1 - c_{kl}^2\right) \frac{\sum_i \sum_j W_{ij} t_{ik}^l}{\gamma_k^l}$$
$$\text{s.t.} \quad (26) - (30), (32) - (34), (66) - (70).$$

EK-MISOCP8 has $3NL$ additional constraints and $N^3L + 2NL$ additional variables, out of which there are N^3L SOCs, each of dimension $N + 3$.

EK-MISOCP9: Multiplying both sides with γ_k^l, and adding $(\sum_i \sum_j W_{ij} t_{ik}^l)^2$ to the constraint $(\sum_i \sum_j W_{ij} t_{ik}^l) \leq s_{kl}(\gamma_k^l - \sum_i \sum_j W_{ij} t_{ik}^l) \quad \forall k,l$, we have $(\sum_i \sum_j W_{ij} t_{ik}^l)^2 \leq (s_{kl}\gamma_k^l - \sum_i \sum_j W_{ij} t_{ik}^l)(\gamma_k^l - \sum_i \sum_j W_{ij} t_{ik}^l) \quad \forall k,l.$

Define t_{kl} and v_{kl}, such that,

$$v_{kl} = s_{kl}\gamma_k^l - \sum_i \sum_j W_{ij} t_{ik}^l, \quad \forall k \in N, l \in L, \tag{71}$$

$$t_{kl} = \gamma_k^l - \sum_i \sum_j W_{ij} t_{ik}^l \quad \forall k \in N, l \in L, \tag{72}$$

$$t_{kl}, v_{kl}, s_{kl} \geq 0 \quad \forall k \in N, l \in L. \tag{73}$$

The constraint becomes

$$(\sum_i \sum_j W_{ij} t_{ik}^l)^2 \leq t_{kl} v_{kl} \qquad \forall k,l, \tag{74}$$

which is transformed to Form-2.1 as:

$$2(\sum_i \sum_j W_{ij} t_{ik}^l)^2 + v_{kl}^2 + t_{kl}^2 \leq (v_{kl} + t_{kl})^2 \qquad \forall k,l. \tag{75}$$

Our ninth EK-MISOCP-based formulation is, therefore:

[EK-MISOCP9]

$$\min \sum_i \sum_k d_{ik}(\chi O_i + \delta D_i) \sum_l t_{ik}^l + \sum_i \sum_k \sum_m \alpha d_{km} x_{ikm}$$
$$+ \sum_k \sum_l Q_k^l t_{kk}^l + \theta/2 \sum_k \sum_l \left(1 + c_{kl}^2\right) s_{kl} + \left(1 - c_{kl}^2\right) \frac{\sum_i \sum_j W_{ij} t_{ik}^l}{\gamma_k^l}$$
$$\text{s.t.} \quad (26) - (30), (32) - (34), (71) - (73), (75).$$

EK-MISOCP9 introduces $3NL$ additional constraints and $3NL$ additional variables including NL SOCs, each of dimension 4.

Table 1 presents a summary of all MISOCPs including number (#) of constraints, variables, and SOCs along with their dimensions.

Table 1: Summary of MISOCP-based Reformulations

EK	SOCs			Model Constraints			
MISOCP	Eqn #	#	Size	Eqn #	#	Binary Var.	Conti. Var.
1	39	N	N+3	3-7, 9, 15- 16, 18- 19, 21, 37, 39	$3N^2 + 7N + 2NL + 1$	$N^2 + NL$	$N^3 + 3N + 2NL$
2	44	N	L+3	3-7, 9, 15- 16, 18- 19, 21, 41-42, 44	$3N^2 + 8N + 2NL + 1$	$N^2 + NL$	$N^3 + 4N + 2NL$
3	46	NL	4	3-7, 9, 15- 16, 18- 19, 21, 41-42, 46	$3N^2 + 7N + 3NL + 1$	$N^2 + NL$	$N^3 + 4N + 2NL$
4	50	N	4	3-7, 9, 15- 16, 18- 19, 21, 41-42, 47, 50	$3N^2 + 9N + 2NL + 1$	$N^2 + NL$	$N^3 + 5N + 2NL$
5	56	N	4	3-9, 15- 16, 18- 19, 21, 52-53, 56	$3N^2 + 9N + 2NL + 1$	$N^2 + NL$	$N^3 + 4N + 2NL$
6	60	NL	4	26-30, 32, 58, 60	$3N^2 + N^2L + 2NL + N + 1$	N^2L	$N^3 + 2NL$
7	64	NL	N+3	26-30, 32, 62, 64	$3N^2 + N^2L + 2NL + N + 1$	N^2L	$N^3 + 2NL$
8	70	N^3L	4	26-30, 32, 66-67, 70	$3N^2 + N^3L + N^2L + 2NL$ + N+1	N^2L	$N^3 + N^3L + 2NL$
9	75	NL	4	26-30, 32, 71-72, 75	$3N^2 + N^2L + 3NL + N + 1$	N^2L	$N^3 + 3NL$

5 Computational Experiments

For our computational experiments, we use two well-known datasets, CAB (56) and AP (3). In Section 5.1, we perform experiments on instances generated from the CAB dataset for all MISOCP-based reformulations (SK-2s, SK-3s, EK-2s and EK-3s based reformulations), which are solved directly using the solver. We compare their performances against the two MILP reformulations, which are obtained using outer-approximation (OA). All the computational experiments are run on a workstation with a 2.20 GHz Intel Xeon E5-2630 processor and 64 GB RAM. All MISOCP-based reformulations are solved directly using CPLEX 12.7.1. For the two OA-based methods, the MILP at each iteration is solved using CPLEX 12.7.1.

CPLEX uses one of the following two alternate parameter settings to solve MISOCPs: *miqcpstrat 1* and *miqcpstrat 2*. In *miqcpstrat 1*, it uses an SOCP based branch and bound algorithm, where at each node the continuous relaxation is solved using an interior point algorithm, specifically designed for SOCPs. In *miqcpstrat 2*, it uses the fact that the SOCPs are NLPs, which can be approximated using outer-approximations. To allow CPLEX to choose the best strategy from among the two, we set the parameter to *miqcpstrat 0*.

5.1 Experiments based on CAB dataset

Using CAB dataset, we run our experiments on instances with the number of nodes ($|N|$) as 10, 15, 20 and 25, and the number of hubs (p) as 3 and 4. We use two values of the inter-hub discount factor (α) as 0.4 and 0.8, and two values of congestion cost(θ) as 20 and 50. The coefficient of variation (c_k) of the service times at hub k is varied as 0, 1 and 2 to model (M/D/1), (M/M/1) and (M/G/1) queues, respectively. This gives us a total of 96 instances for our computational experiments. For each of these instances, the three capacity levels ($l = 1, 2, 3$) at any hub are set as $\frac{\sum_i \sum_j W_{ij}}{p} + \beta A_l \sum_i \sum_j W_{ij}$, where A_l = -1, 0, 1 for

$l = 1, 2, 3$, respectively, and β is set to 0.21, 0.22, 0.23 and 0.24 for $|N|$= 10, 15, 20 and 25. Fixed cost Q_k^l is set as 150, 200 and 250 for the three capacity levels. A similar scheme for setting the capacity levels and the corresponding costs is used by (1). The optimality tolerance (ϵ) for OA is set to 10^{-6}. For each instance, a maximum CPU time limit of 14,400 seconds (4 hours) is used.

5.1.1 Results and Analysis:

From our initial experiments, as reported in Table 5 and Table 8 in Appendix A, we find that the results from the SK-based models are far inferior to those based on EK-based models. Hence, in the remaining experiments, we restrict our analysis to only the EK-based models. We compare the performances of all the proposed EK-based MISOCP reformulations against EK-OA-2s, EK-OA-3s, obtained using OA, and strengthened using perspective reformulation (57). Reformulations using OA is presented in Appendix B. Our initial computational experiments clearly suggested EK-MISOCP8 to be the worst among all formulations, and is, therefore, excluded from further analysis. We discuss the reasons for the inferior performance of EK-MISOCP8 in Section 5.3. Tables 2, 3, 4, and 5 present the computational performances (in terms of CPU time) of eight of the nine MISOCPs (excluding EK-MISOCP8), and the two OA-based reformulations for $|N|$=10, 15, 20 and 25. In each of these tables, the column "Hub" reports the set of p hubs opened in the resulting solution, while columns "Cap" and "Intensity" report the corresponding capacity level and the traffic intensity at each of the open hubs, respectively. The flow cost, location (capacity installation) cost, congestion cost, and sum of all the three costs are reported under the columns "FC", "LC", "CC", and "TC". The last row in each table reports the number of instances for which each of the formulations performs the best, based on which EK-MISOCP4 and EK-MISOCP5 seem to be the best two formulations, exhibiting the best performances for 46 and 31 (out of a total of 96) instances, respectively. Comparison among the rest of the formulations is not so obvious from the tables. Further, for the remaining 19 instances where neither EK-MISOCP4 nor EK-MISOCP5 is the best, it is not immediately obvious from the tables how well they perform vis-à-vis the rest. Hence, for a better comparative analysis, we use performance profiles (58). For this, let n_s and n_p be the number of formulations and the number of test instances, respectively, and S and P be their respective sets. Let $t_{p,s}$ be the CPU time taken to solve the instance $p \in P$ using formulation $s \in S$. Then, performance ratio $r_{p,s}$ is calculated as:

$$r_{p,s} = \frac{t_{p,s}}{\min_{s \in S}\{t_{p,s}\}}$$

Treating $t_{p,s}$ as a random variable, a performance profile $\rho_s(\tau)$ gives its cumulative probability distribution at τ. In other words, it gives the probability with which the CPU time taken by a given formulation s does not exceed 2^τ times the CPU time required by the best among all the formulations under study. Mathematically, it can be stated as:

$$\rho_s(\tau) = \frac{1}{n_p}\Big\{p \in P : \log_2(r_{p,s}) \leq \tau\Big\}.$$

Specifically, $\rho_s(\tau = 0)$ gives the probability that s is the best among all the formulations.

Figure 1 presents four different performance profiles of the different EK-MISOCP and OA-based reformulations, one corresponding to each value of $|N|$. For $|N| = 10$, Figure 1a

Table 2: Comparison of EK-MISOCPs against EK-OA-2s and EK-OA-3s for N=10 (CAB dataset)

c	θ	Hub Located	Cap Level	Intensity	Costs: FC, LC, CC	Costs: TC	EK OA-2s	EK OA-3s	CPU time (sec) MISOCPs: 1	2	3	4	5	6	7	9
					p=3, $\alpha = 0.4$											
0	20	4,6,7	3,3,2	0.7, 0.7, 0.8	575, 700, 97	1373	2.5	1.8	2.4	2.3	1.0	1.0	**1.0**	1.9	2.3	1.7
	50	4,6,7	3,3,3	0.7, 0.7, 0.5	575, 750, 177	1503	1.7	1.6	3.8	3.1	5.9	1.1	**0.7**	1.1	2.8	0.9
1	20	4,6,7	3,3,2	0.7, 0.7, 0.8	575, 700, 152	1427	2.9	2.1	3.9	3.2	3.7	**0.9**	1.2	1.8	2.7	1.1
	50	4,6,7	3,3,3	0.7, 0.7, 0.5	575, 750, 263	1588	1.7	**4**	4.1	1.0	2.6	1.3	**0.9**	1.2	3.7	0.9
2	20	4,6,7	3,3,3	0.7, 0.7, 0.5	575, 750, 208	1533	2	3.5	4.5	1.1	2.6	**1.0**	2.0	1.3	3.3	1.4
	50	4,6,7	3,3,3	0.7, 0.7, 0.5	575, 750, 519	1845	3.6	5.2	5.5	2.9	4.3	1.6	**1.1**	1.7	6.8	1.4
					p=3, $\alpha = 0.8$											
0	20	4,7,9	3,2,3	0.7, 0.6, 0.8	717, 700, 96	1513	2.0	2.8	4.7	2.7	1.2	**1.2**	1.2	1.5	1.8	1.6
	50	4,7,9	3,3,3	0.7, 0.5, 0.7	721, 750, 177	1649	2.2	2.4	5.0	4.4	1.3	1.2	1.6	**1.0**	1.8	1.3
1	20	4,7,9	3,2,3	0.7, 0.6, 0.8	717, 700, 150	1567	3.0	2.5	4.3	2.4	1.9	1.0	**0.9**	1.2	2.3	1.4
	50	4,7,9	3,3,3	0.7, 0.5, 0.7	721, 750, 263	1734	3.2	3.5	2.6	2.9	1.7	**0.9**	1.4	1.9	3.6	1.5
2	20	4,7,9	3,3,3	0.7, 0.5, 0.7	721, 750, 208	1679	3.5	3.8	3.0	1.4	1.9	**1.0**	1.6	1.6	3.4	1.6
	50	4,7,9	3,3,3	0.7, 0.5, 0.7	721, 750, 519	1991	4.1	6.8	4.0	4.0	1.8	1.6	**1.4**	1.8	5.4	1.5
					p=4, $\alpha = 0.4$											
0	20	3,4,6,7	2,3,3,3	0.4, 0.7, 0.7, 0.6	500, 950, 89	1539	1.7	2.2	2.6	2.8	1.7	0.9	1.5	**0.9**	2.2	1.5
	50	3,4,6,7	2,3,3,3	0.4, 0.7, 0.7, 0.6	500, 950, 222	1672	2.6	2.8	2.0	1.3	2.1	1.2	1.1	1.4	2.5	**1.1**
1	20	3,4,6,7	2,3,3,3	0.4, 0.7, 0.7, 0.6	500, 950, 130	1581	2.6	2.4	4.0	3.4	1.4	**0.8**	1.2	1.1	1.6	1.4
	50	2,4,7,9	3,3,3,3	0.5, 0.5, 0.6, 0.6	529, 1000, 239	1768	3.0	3.8	3.1	3.1	1.5	**1.1**	1.3	1.2	2.8	1.2
2	20	2,4,7,9	3,3,3,3	0.5, 0.5, 0.6, 0.6	529, 1000, 174	1703	2.6	4.5	3.5	1.8	4.2	1.2	**1.1**	1.5	3.5	1.9
	50	2,4,7,9	3,3,3,3	0.5, 0.5, 0.6, 0.6	529, 1000, 434	1963	2.6	6.6	2.0	2.5	3.4	1.4	**1.2**	1.2	7.2	1.4
					p=4, $\alpha = 0.8$											
0	20	4,5,6,7	3,2,3,2	0.7, 0.5, 0.8, 0.8	672, 900, 133	1705	2.0	3	4.3	1.4	3.3	**0.9**	1.1	1.0	2.4	1.2
	50	3,4,7,9	2,3,3,3	0.4, 0.7, 0.6, 0.7	693, 950, 222	1865	2.6	5	3.5	**1.1**	2.6	1.4	1.1	1.8	2.0	1.3
1	20	3,4,7,9	2,3,3,3	0.4, 0.7, 0.6, 0.7	693, 950, 130	1773	2.2	5	3.5	1.1	2.6	1.4	**1.0**	1.4	3.7	1.3
	50	4,5,6,7	3,3,3,3	0.5, 0.5, 0.6, 0.6	704, 1000, 239	1942	4.6	3.8	4.9	1.9	3.5	1.2	**1.1**	1.4	2.6	1.2
2	20	4,5,6,7	3,3,3,3	0.5, 0.5, 0.6, 0.6	704, 1000, 173	1877	7.1	4.2	5.0	4.6	3.5	1.7	**1.1**	1.6	3.2	1.8
	50	4,5,6,7	3,3,3,3	0.5, 0.5, 0.6, 0.6	704, 1000, 434	2137	5.4	5	5.8	4.2	3.2	1.9	1.5	**1.3**	6.2	1.8
				No. of best performing instances			0	0	0	1	0	8	11	3	0	1

Table 3: Comparison of EK-MISOCPs against EK-OA-2s and EK-OA-3s for N=15 (CAB dataset)

c	θ	Hub Located	Cap Level	Intensity	Costs: FC, LC, CC	Costs: TC	EK OA-2s	EK OA-3s	CPU time (sec) MISOCPs 1	2	3	4	5	6	7	9
						p=3, $\alpha = 0.4$										
0	20	4,12,13	3,2,3	0.9, 0.4, 0.7	929, 700, 114	1743	22.2	20.9	5.0	**2.7**	4.1	4.0	3.4	6.0	9.8	6.0
	50	4,12,13	3,2,3	0.8, 0.5, 0.7	964, 700, 236	1900	29.7	41.4	3.1	**2.9**	4.4	3.1	3.0	7.6	14.8	8.3
1	20	4,12,13	3,2,3	0.8, 0.4, 0.8	940, 700, 169	1809	16.8	33.8	3.5	**2.3**	3.0	3.4	5.8	5.9	18.8	5.6
	50	4,6,7	3,3,3	0.6, 0.6, 0.6	1011, 750, 229	1990	39.8	45.4	4.2	**3.5**	3.7	4.5	4.0	7.5	27.4	7.4
2	20	4,6,7	3,3,3	0.6, 0.6, 0.6	1011, 750, 175	1936	124.9	59.2	4.4	4.4	3.9	4.0	**3.8**	9.3	53.3	10.4
	50	4,6,7	3,3,3	0.6, 0.6, 0.6	1011, 750, 437	2198	30.6	53.3	8.8	6.1	9.0	5.9	4.9	14.5	195.6	10.2
						p=3, $\alpha = 0.8$										
0	20	4,5,11	2,3,3	0.7, 0.7, 0.6	1140, 700, 92	1932	24.0	34.1	5.6	3.5	4.9	**3.4**	5.1	6.8	8.1	6.4
	50	4,5,11	3,3,3	0.6, 0.6, 0.6	1138, 750, 159	2047	32.6	43.5	5.7	4.1	4.4	**2.7**	3.2	3.8	20.5	4.6
1	20	4,5,11	3,3,3	0.6, 0.6, 0.6	1138, 750, 91	1979	25	48.0	3.8	3.8	4.4	**3.6**	7.3	9.2	17.6	6.2
	50	4,5,11	3,3,3	0.6, 0.6, 0.6	1138, 750, 229	2116	29.1	41.7	5.8	4.3	5.9	4.0	**3.2**	4.9	34.5	6.4
2	20	4,5,11	3,3,3	0.6, 0.6, 0.6	1138, 750, 174	2062	29.0	44.8	5.1	4.4	6.0	4.9	**3.4**	8.1	25.1	5.3
	50	4,5,11	3,3,3	0.6, 0.6, 0.6	1138, 750, 436	2324	46	589.6	12.1	7.0	9.7	6.9	**4.5**	15.3	621.3	9.5
						p=4, $\alpha = 0.4$										
0	20	4,6,7,12	3,3,2,2	0.7, 0.8, 0.7, 0.5	814, 900, 121	1835	12.9	16.4	4.8	3.3	4.7	**2.4**	3.2	4.3	9.9	3.7
	50	4,6,12,13	3,3,2,3	0.6, 0.6, 0.5, 0.7	832, 950, 210	1991	28.2	33.6	5.5	3.4	3.7	2.8	**2.6**	4.9	15.5	5.5
1	20	4,5,7,12	3,3,2,2	0.8, 0.8, 0.7, 0.5	816, 900, 186	1902	30.1	45.2	4.9	3.7	3.8	**3.1**	3.8	6.9	11.0	4.6
	50	4,6,12,13	3,3,2,3	0.6, 0.6, 0.5, 0.7	832, 950, 302	2084	31.1	70.7	7.3	5.1	4.5	3.3	**3.0**	6.0	22.2	5.9
2	20	4,6,12,13	3,3,2,3	0.6, 0.6, 0.5, 0.7	832, 950, 231	2013	44.1	36.6	6.0	4.3	6.8	**3.5**	3.7	7.8	80.8	6.8
	50	4,6,12,13	3,3,3,3	0.6, 0.6, 0.4, 0.6	854, 1000, 450	2304	22.0	36.6	21.6	6.5	11.6	8.0	**5.2**	10.3	1704.1	11.8
						p=3, $\alpha = 0.8$										
0	20	4,5,7,8	3,3,2,2	0.8, 0.8, 0.5, 0.7	1044, 900, 119	2064	26.7	22.8	4.1	**2.8**	3.7	3.4	3.4	4.5	8.8	4.2
	50	1,4,6,7	2,3,3,3	0.5, 0.7, 0.6, 0.6	1073, 950, 210	2233	27.3	34.1	5.5	4.5	3.7	**2.9**	3.3	6.6	20.2	7.4
1	20	4,5,7,8	3,3,2,2	0.8, 0.8, 0.5, 0.7	1044, 900, 186	2130	36.2	37.9	4.9	3.4	3.8	**2.7**	4.0	4.8	14.3	5.2
	50	4,8,9,13	3,3,3,3	0.6, 0.4, 0.6, 0.6	1063, 1000, 244	2306	27.9	27.2	6.7	6.8	5.4	**4.2**	4.9	6.1	52.2	7.5
2	20	4,8,9,13	3,3,3,3	0.6, 0.4, 0.6, 0.6	1063, 1000, 180	2243	28.3	38.7	9.3	6.2	6.2	4.1	**3.6**	6.8	56.8	7.0
	50	4,8,9,13	3,3,3,3	0.6, 0.4, 0.6, 0.6	1063, 1000, 450	2513	40.3	45.9	33.8	11.3	13.6	7.4	**5.8**	20.3	346.7	13.5
			No. of best performing instances				0	0	0	5	0	9	10	0	0	0

Table 4: Comparison of EK-MISOCPs against EK-OA-2s and EK-OA-3s for N=20 (CAB dataset)

c	θ	Hub Located	Cap Level	Intensity	Costs: FC, LC, CC	Costs: TC	EK OA-2s	EK OA-3s	CPU time (sec) MISOCPs: 1	2	3	4	5	6	7	9
					p=3, $\alpha = 0.4$											
0	20	4,12,17	3,2,3	0.8, 0.3, 0.8	848, 700, 104	1652	34.1	49.9	12.3	10.8	14.2	**9.2**	13.4	11.2	19.3	10.0
	50	4,12,17	3,2,3	0.8, 0.4, 0.8	865, 700, 238	1803	80.3	122.8	**8.6**	13.6	13.1	10.9	10.9	16.6	27.4	12.8
1	20	4,12,17	3,2,3	0.8, 0.4, 0.8	865, 700, 152	1717	66.7	155.9	14.7	12.4	10.1	**7.9**	15.6	13.7	34.4	21.7
	50	5,7,17	3,3,3	0.7, 0.5, 0.7	908, 750, 236	1894	99.2	215.1	19.5	17.8	17.3	17.4	**14.0**	17.5	118.5	19.6
2	20	5,7,17	3,3,3	0.7, 0.5, 0.7	908, 750, 183	1841	83.1	257.0	19.9	18.5	19.1	18.3	**15.0**	34.8	730.7	29.2
	50	4,7,17	3,3,3	0.6, 0.6, 0.6	939, 750, 411	2099	84.9	227.5	40.7	33.6	34.7	36.5	**19.9**	70.5	6484.1	60.1
					p=3, $\alpha = 0.8$											
0	20	1,4,17	2,3,3	0.6, 0.7, 0.7	1097, 700, 85	1882	85.4	231.9	18.7	**15.0**	17.5	15.9	15.4	18.9	74.6	22.4
	50	1,4,17	2,3,3	0.6, 0.7, 0.7	1097, 700, 213	2010	109.7	301.4	27.7	18.7	19.2	15.2	**14.0**	32.6	164.4	34.6
1	20	1,4,17	2,3,3	0.6, 0.7, 0.7	1097, 700, 130	1927	97.7	252.6	21.9	18.0	17.9	**15.6**	17.5	22.3	131.1	22.8
	50	6,11,17	3,3,3	0.7, 0.5, 0.6	1103, 750, 227	2080	110.3	288.9	32.6	20.9	29.5	23.0	**15.8**	42.2	565.8	39.9
2	20	6,11,17	3,3,3	0.7, 0.5, 0.6	1103, 750, 174	2027	143.7	243.3	41.1	36.1	28.7	24.0	**21.3**	51.1	1143.5	54.8
	50	4,17,20	3,3,3	0.6, 0.6, 0.6	1121, 750, 415	2286	512.1	316.5	106.0	32.9	74.9	58.7	**22.8**	123.1	*	168.5
					p=4, $\alpha = 0.4$											
0	20	1,4,12,17	2,3,2,3	0.7, 0.7, 0.4, 0.8	733, 900, 119	1752	49.7	57.3	20.0	16.5	12.1	**7.6**	10.2	9.6	27.2	10.5
	50	1,4,12,17	3,3,2,3	0.5, 0.7, 0.4, 0.7	744, 950, 212	1906	89.3	88.4	15.1	15.4	15.8	12.0	**11.7**	12.1	32.5	14.0
1	20	1,4,12,17	3,3,2,3	0.4, 0.7, 0.4, 0.8	727, 950, 138	1815	86.2	104.1	40.7	33.6	34.7	36.5	**19.9**	70.5	6484.1	60.1
	50	4,12,13,17	3,2,3,3	0.6, 0.4, 0.6, 0.7	752, 950, 296	1998	80.0	138.0	16.7	21.2	23.7	14.7	**11.8**	23.4	85.3	22.7
2	20	4,12,13,17	3,2,3,3	0.6, 0.4, 0.6, 0.7	752, 950, 228	1929	90.1	134.6	27.1	23.1	20.3	15.3	**14.4**	22.2	149.5	26.3
	50	4,7,17,20	3,3,3,3	0.5, 0.5, 0.6, 0.5	843, 1000, 391	2234	112.2	158.8	150.8	35.5	55.0	60.6	**20.9**	103.0	7720.8	130.0
					p=4, $\alpha = 0.8$											
0	20	1,4,8,17	2,3,2,3	0.6, 0.7, 0.5, 0.8	1016, 900, 114	2030	14400	268.7	22.0	14.2	21.1	**12.1**	19.6	26.4	53.6	23.1
	50	4,8,17,20	3,2,3,3	0.6, 0.5, 0.7, 0.6	1025, 950, 201	2177	89.8	329.7	38.4	27.3	20.4	21.1	**18.7**	27.2	179.2	23.4
1	20	4,8,17,20	3,2,3,3	0.6, 0.5, 0.7, 0.6	1025, 950, 114	2090	166.4	359.1	40.8	25.0	27.3	23.6	**17.5**	31.7	129.6	29.8
	50	4,7,17,20	3,3,3,3	0.4, 0.5, 0.6, 0.6	1031, 1000, 229	2261	166.0	232.7	83.1	33.8	33.2	29.7	**22.9**	57.6	1094.6	75.3
2	20	4,8,17,20	3,2,3,3	0.6, 0.5, 0.7, 0.6	1025, 950, 216	2192	193.9	232.5	95.4	35.5	43.8	36.7	**21.6**	56.1	1461.9	54.8
	50	4,11,17,20	3,3,3,3	0.5, 0.5, 0.6, 0.5	1045, 1000, 391	2435	222.1	165.3	499.7	167.5	87.5	78.8	**26.3**	93.1	9901.5	115.9
			No. of best performing instances				0	0	1	1	0	5	17	0	0	0

* represents optimal solution not found in the given CPU time limit of 4 hours.

Table 5: Comparison of EK-MISOCPs against EK-OA-2s and EK-OA-3s for N=25 (CAB dataset)

c	θ	Hub Located	Cap Level	Intensity	Costs: FC, LC, CC	Costs: TC	EK OA-2s	EK OA-3s	CPU time (sec) MISOCPs 1	2	3	4	5	6	7	9
						p=3, $\alpha = 0.4$										
0	20	4,12,18	3,2,3	0.7, 0.5, 0.8	902, 700, 85	1687	96.8	131.9	28.5	30.7	**19.9**	44.8	35.2	42.5	61.1	31.5
	50	4,12,17	3,2,3	0.7, 0.5, 0.7	903, 700, 209	1813	144.5	182.2	34.2	40.0	37.3	31.5	31.0	41.1	82.6	**29.6**
1	20	4,12,17	3,2,3	0.7, 0.5, 0.7	903, 700, 129	1732	155.4	198.4	29.0	37.3	42.2	32.6	**24.2**	48.0	85.9	38.8
	50	4,12,17	3,2,3	0.7, 0.6, 0.7	913, 700, 312	1925	210.6	343.7	46.0	53.0	64.7	36.3	**27.9**	49.4	486.2	66.5
2	20	4,12,17	3,2,3	0.7, 0.6, 0.7	913, 700, 253	1866	222.1	379.2	50.3	53.8	61.3	52.3	**33.4**	51.8	994.7	83.7
	50	5,8,17	3,3,3	0.6, 0.5, 0.6	1022, 750, 410	2182	314.7	512.6	263.0	**88.7**	131.7	124.1	102.3	385.0	*	581.8
						p=3, $\alpha = 0.8$										
0	20	2,4,12	3,3,2	0.8, 0.7, 0.5	1166, 700, 85	1951	290.6	482.6	88.1	78.3	101.5	90.4	**65.5**	86.7	201.2	88.3
	50	4,12,18	3,2,3	0.7, 0.5, 0.7	1169, 700, 209	2078	245.0	663.0	119.6	96.4	100.3	79.9	**65.0**	128.8	283.0	144.5
1	20	4,12,18	3,2,3	0.7, 0.5, 0.7	1169, 700, 129	1997	214.0	698.9	81.9	91.7	102.4	91.0	**70.6**	88.0	443.6	86.5
	50	4,12,18	3,2,3	0.7, 0.5, 0.7	1169, 700, 322	2190	554.2	1250.9	148.3	228.5	178.6	141.5	**110.3**	208.3	13787.1	332.5
2	20	4,12,18	3,2,3	0.7, 0.5, 0.7	1169, 700, 263	2132	726.2	1045.4	243.2	166.7	165.6	134.2	**92.2**	291.8	*	757.6
	50	5,11,18	3,3,3	0.6, 0.6, 0.6	1250, 750, 390	2391	632.9	1079.6	812.7	479.5	382.7	354.5	**159.9**	1064.9	*	1038.2
						p=4, $\alpha = 0.4$										
0	20	1,4,12,17	2,3,2,3	0.6, 0.7, 0.6, 0.8	695, 900, 109	1704	115.3	116.4	**42.5**	57.1	63.0	55.2	56.7	63.3	82.2	54.8
	50	1,4,12,17	3,3,2,3	0.4, 0.6, 0.6, 0.7	795, 950, 219	1963	184.1	322.5	**54.3**	75.8	93.4	80.0	**75.8**	71.8	350.9	84.4
1	20	1,4,12,17	2,3,2,3	0.6, 0.7, 0.6, 0.8	792, 900, 165	1857	214.9	272.2	54.9	73.2	84.0	**44.4**	47.8	52.8	140.7	63.8
	50	1,4,12,17	3,3,3,3	0.4, 0.6, 0.4, 0.7	803, 1000, 247	2050	281.4	511.7	115.7	149.9	141.1	102.7	**75.4**	130.2	2370.5	108.3
2	20	1,4,12,17	3,3,3,3	0.4, 0.6, 0.4, 0.7	803, 1000, 185	1988	292.4	465.9	125.2	130.3	139.9	122.5	**95.3**	112.8	4437.5	129.5
	50	1,4,12,17	3,3,3,3	0.5, 0.5, 0.4, 0.7	824, 1000, 420	2244	391.7	527.1	2027.1	191.6	224.8	212.6	**105.2**	686.3	*	615.9
						p=4, $\alpha = 0.8$										
0	20	1,4,12,18	2,3,2,3	0.6, 0.7, 0.6, 0.8	1091, 900, 109	2100	259.6	795.4	**83.7**	89.9	139.8	85.9	88.7	140.5	311.6	102.1
	50	1,4,12,18	2,3,2,3	0.6, 0.7, 0.6, 0.8	1093, 900, 270	2262	649.7	3258.8	**215.9**	136.5	190.3	140.2	**134.8**	266.8	2344.9	226.4
1	20	1,4,12,18	2,3,2,3	0.6, 0.7, 0.6, 0.8	1093, 900, 163	2156	690.8	1349.7	124.8	137.4	112.3	90.1	**87.9**	189.7	865.1	244.1
	50	4,12,17,20	3,3,3,3	0.6, 0.3, 0.5, 0.6	1113, 1000, 226	2339	1977.0	2848.1	834.4	224.9	186.6	232.0	**123.8**	455.9	*	628.5
2	20	4,12,17,20	3,3,3,3	0.6, 0.3, 0.5, 0.6	1113, 1000, 165	2277	950.7	2558.5	1487.8	254.8	296.7	322.0	**148.7**	761.2	*	1297.7
	50	4,8,17,20	3,3,3,3	0.6, 0.4, 0.5, 0.6	1131, 1000, 392	2522	1641.3	2165.8	*	755.6	919.1	450.9	**153.6**	3911.4	*	3349.6
				No. of best performing instances			0	0	4	1	1	1	18	0	0	1

* represents optimal solution not found in the given CPU time limit of 4 hours.

shows that EK-MISOCP5 and EK-MISOCP4 perform the best for around 46% (11 out of 24) and 33% (8 out of 24) of the instances, respectively. This is something that is already obvious from Table 2. However, what is less obvious from that table is the fact that EK-MISOCP5 never takes more than twice the CPU time taken by the best formulation for any of the 24 instances, which is only revealed by the performance profile. Even less obvious is another fact that at $\tau = 0.8$, EK-MISOCP4 is able to solve all the 24 instances, whereas EK-MISOCP5 can solve only 96% of the instances. Further, from Figures 1b, 1c, and 1d, we see that EK-MISOCP5 and EK-MISOCP4 are the best two formulations across all problem instances. We do similar analyses for all formulations with respect to the congestion factor θ and the coefficient of variation (c), the results for which are presented in Figure 2 and Figure 3, respectively. Once again, we observe that all MISOCPs dominate the OA-based reformulations, with EK-MISOCP5 being the best, followed by EK-MISOCP4. Moreover, EK-MISOCP5, which is the best formulation, solves all the instances 20-60 times faster as compared to SK-OA-2s, which is the existing best formulation/method in the literature.

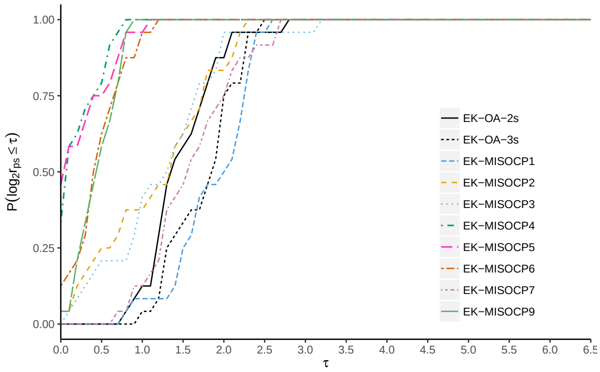
In the absence of congestion, hub location model with three-subscripted capacity allocation variable is known to dominate the model with two-subscripted capacity allocation variable in terms of LP relaxation (41). For CAB dataset, we observe from our computational results that this dominance of EK-3s does not necessarily translate into an advantage from a computational point of view owing to its larger model size. Specifically, all the formulations based on EK-2s, except EK-MISOCP1, perform better than EK-3s. However, in the presence of congestion, this conclusion may not be generalized to other datasets. Hence, for the next set of computational experiments, based on AP dataset, we proceed with EK-MISOCP4 and EK-MISOCP5 (the best two among EK-2s) and EK-MISOCP6 and EK-MISOCP9 (the best two among EK-3s).

5.2 Experiments based on AP dataset

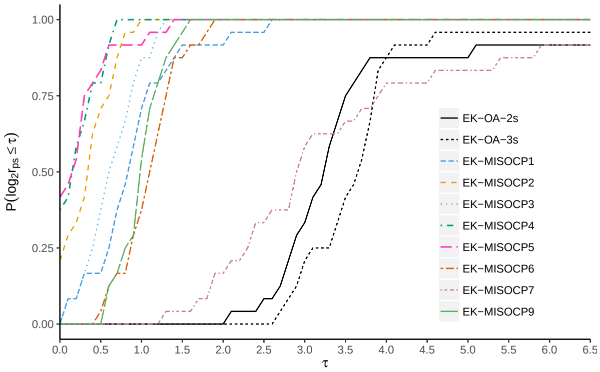
For the AP dataset, we limit our computational experiments only to EK-MISOCP4, EK-MISOCP5, EK-MISOCP6 and EK-MISOCP9 for $|N|$=25 and 50. The AP dataset specifies two different possible values for the capacity at each hub, referred to as tight (T) and loose (L). Following the scheme used by (43), we set the capacity (γ_k^l) and the fixed cost (Q_k^l) for level l at the potential hub node k as $\gamma_k^L = \Gamma_k^L$, $\gamma_k^l = 0.7 \times \gamma_k^{l+1}$, $Q_k^L = \Gamma_L^k$, and $Q_k^l = 0.9 \times Q_k^{l+1}\ \forall l = 1 \ldots L-1$, where Γ_k^L is the capacity for hub k provided in the dataset. Similar to experiments using CAB dataset, corresponding to each value of $|N|$, we run our experiments with the number of hubs (p) as 3 and 4, inter-hub discount factor (α) as 0.4 and 0.8, and the coefficient of variation (c_k) of the service times at hub k as 0, 1 and 2. In addition, we set the congestion cost(θ) as 20 and 50.

5.2.1 Results and Analysis:

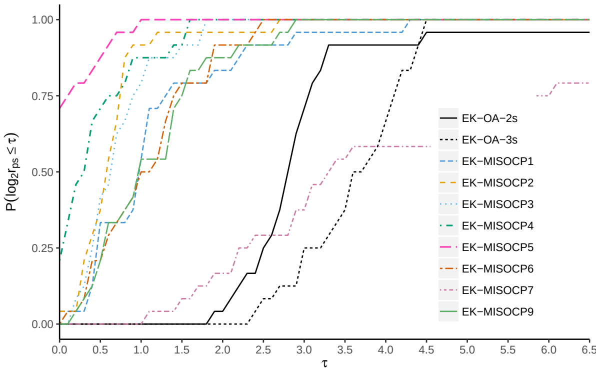
Tables 6 and 7 present the computational results for our experiments with the AP dataset. The column names in these tables are the same as the ones used for CAB dataset. For EK-MISOCP5, the column "CPU time (sec)/% Gap" reports the CPU time in seconds taken to solve the instance to optimality. For the other MISOCPs (4, 6, and 9), the maximum CPU time limit is set as the CPU time required to solve EK-MISOCP5 to optimality. For



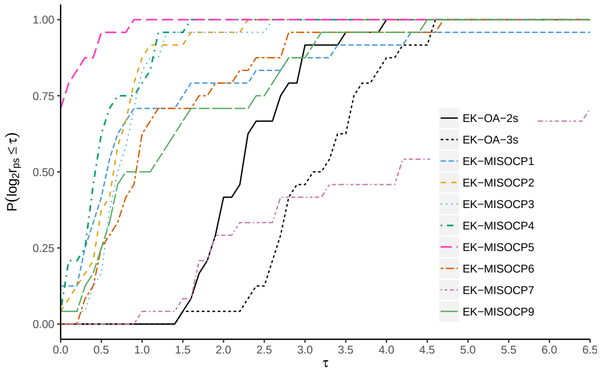
(a) N=10



(b) N=15



(c) N=20



(d) N=25

Figure 1: Performance Profile of EK-MISOCPs and OA-based method for N= 10,15,20,25

the instances that could not be solved to optimality within the time limit, we provide their optimality gap as reported by CPLEX. For a few instances that were solved within the time limit, we report the corresponding CPU times. For many of the instances, CPLEX was not able to find even a single integer feasible solution, which we denote in the table using an asterisk (*). As obvious from Tables 6 and 7, EK-MISOCP5 continues to outperform all other MISOCPs, followed by EK-MISOCP4. However, for smaller instances ($|N|$=25) with loose capacity (L), EK-3s models (EK-MISOCP6 and EK-MISOCP9) are able to beat EK-2s models (EK-MISOCP5 and EK-MISOCP4) in a few cases. This clearly suggests that for smaller instances continuous relaxation may play a role; however, for larger instances the performance is primarily dictated by the model size.

5.3 Observations

The difference in performances among the alternate MISOCPs can be partly understood from Table 1, which summarizes their properties. The number of SOC constraints ($|N|$) and their dimensions (4) are the smallest for EK-MISOCP5. It also has the least number of binary variables ($|N|^2+|N|\times|L|$) and continuous variables ($|N|^3+4|N|+2|N|\times|L|$). This explains the superiority of EK-MISOCP5 over all the other MISOCPs. EK-MISOCP4 also has the same number and dimension of SOC constraints, and the number of binary variables. However, its performance is slightly worse overall, as compared to EK-MISOCP5, due to a slightly higher number of continuous variables. Understandably, EK-MISOCP8 performs the worst since it has the largest number of SOC constraints ($|N|^3|L|$) and also the largest number of binary variables ($|N|^2|L|$) and continuous variables ($|N|^3 + |N|^3|L| + 2|N| \times |L|$). Though the dimension of the SOCs in EK-MISOCP8 is the smallest (4), its computational performance gets dictated largely by the other attributes, which are comparatively far worse. For similar reasons, the performance of EK-MISOCP7 is also very poor, only next to EK-MISOCP8. EK-MISOCP1, EK-MISOCP2, EK-MISOCP3, EK-MISOCP6, and EK-MISOCP9 are intermediate performers. Their performances can be explained based on the attributes discussed above. We hope the insights presented in this section will be useful in selecting an MISOCP from among several alternatives in other problem contexts as well.

6 Conclusions and Future Research Directions

In this paper, we have proposed several MISOCP-based reformulations for the hub location problem with capacity selection under congestion. The hub location problem is computationally challenging to solve; accounting for congestion at hubs adds another layer of difficulty by introducing non-linearity in the resulting model. The contribution of the paper lies in proposing several MISOCP-based reformulations for the NLMIP problem, which can be efficiently handled by the existing mixed integer programming solvers. All our MISOCP-based reformulations outperform the reformulations based on outer approximation on all the 96 instances generated from the CAB dataset. Further, based on our computational studies, we have identified the best MISOCP-based reformulation (EK-MISOCP5), which turns out to be several times faster than the existing best formulation/method in the literature. In particular, EK-MISOCP5 has allowed us to exactly solve instances from the AP dataset of the size of up to 50 nodes in less half-an-hour. The paper also provides insights into the properties of MISOCPs that determine their computational efficiency, which would be important in making the best selection from among various alternate formulations.

To the best of our knowledge, ours is the first application of MISOCP-based reformulation to hub location problem under congestion. Prompted by our success in the current study, we foresee its applications in other classes of hub location problems, where non-linearities may arise. For example, one potential application can be in the area of competitive hub location (59; 60), and more broadly in competitive facility location, where the non-linearity results from the market share function. On the methodology side, decomposition based techniques, like Lagrangean decomposition, Dantzig-Wolfe decomposition, and Benders decomposition, can be explored to solve the resulting MISOCP-based reformulation even more efficiently. Further, valid inequalities like polymatroid cuts (61), which characterize the convex hull corresponding to the mixed integer second order conic constraint, can be used in a branch-and-cut framework.

Table 6: Comparison among EK-MISOCP4, 5, 6 and 9 for |N| = 25 (AP dataset)

c	θ	Hub	Cap	Intensity	Costs		CPU time (sec)/% Gap			
							MISOCPs			
					FC, LC, CC	TC	4	5	6	9
					N=25L					
					p=3, $\alpha = 0.4$					
0	50	2,13,18	3,3,3	0.9, 0.8, 0.9	68263.4, 4529.85, 682.589	73475.9	1.5%	**12.872**	5.5%	5.5%
	100	2,13,18	3,3,3	0.9, 0.8, 0.9	68263.4, 4529.85, 1365.18	74158.5	**24.734**	25.52	7.8%	7.8%
	200	7,14,18	1,1,3	0.6, 0.4, 0.9	66671.8, 6843.55, 1831.22	75346.6	10.4%	**14.268**	*	5.7%
1	50	2,13,18	3,3,3	0.9, 0.8, 0.9	68263.4, 4529.85, 1235.26	74028.6	*	**9.624**	*	*
	100	7,14,18	1,1,3	0.6, 0.4, 0.9	66671.8, 6843.55, 1640.5	75155.9	2%	**33.06**	12.5%	12.5%
	200	8,17,18	2,1,3	0.6, 0.5, 0.8	68095.7, 7486.54, 1129.36	76711.6	48.435	62.64	49.231	**46.591**
2	50	7,14,18	1,1,3	0.6, 0.4, 0.9	66671.8, 6843.55, 1907.59	75423	2.2%	**37.048**	9.4%	6.5%
	100	8,17,18	2,1,3	0.6, 0.5, 0.8	68095.7, 7486.54, 1134.41	76716.6	3.8%	**34.7**	24.1%	24.1%
	200	8,17,18	3,1,3	0.5, 0.4, 0.8	68077.1, 7766.61, 1845.37	77689.1	11%	**42.232**	17.6%	17.4%
					p=3, $\alpha = 0.8$					
0	50	2,13,18	3,3,3	0.9, 0.8, 1	79579.2, 4529.85, 872.156	84981.2	2.7%	**36.672**	4.6%	5.3%
	100	8,17,18	1,1,3	0.8, 0.5, 0.8	78172.7, 6730.34, 578.558	85481.6	3%	**37.608**	5.3%	5.3%
	200	8,17,18	1,1,3	0.8, 0.5, 0.8	78172.7, 6730.34, 1157.12	86060.1	5.1%	**16.312**	*	*
1	50	8,17,18	1,1,3	0.8, 0.5, 0.8	78172.7, 6730.34, 474.379	85377.4	3.2%	**44.14**	4.9%	*
	100	8,17,18	1,1,3	0.8, 0.5, 0.8	78172.7, 6730.34, 948.756	85851.8	6.1%	**16.044**	*	*
	200	8,17,18	1,1,3	0.8, 0.5, 0.8	78172.7, 6730.34, 1897.51	86800.5	41.5%	**11.844**	*	*
2	50	8,17,18	1,1,3	0.8, 0.5, 0.8	78172.7, 6730.34, 1029.68	85932.7	3%	**52.296**	3.7%	5.8%
	100	8,17,18	2,1,3	0.6, 0.5, 0.8	78228.9, 7486.54, 1197.6	86913	6.1%	**14.792**	*	*
	200	8,17,18	3,1,3	0.4, 0.5, 0.8	78383.5, 7766.61, 1858.08	88008.2	54.512	65.824	**51.902**	17%
					p=4, $\alpha = 0.4$					
0	50	2,8,16,18	3,1,1,3	0.9, 0.6, 0.6, 0.9	61510.8, 5776.87, 530.304	67818	1.2%	**17.968**	4.2%	4.2%
	100	2,8,16,18	3,1,1,3	0.9, 0.6, 0.6, 0.9	61510.8, 5776.87, 1060.61	68348.3	2%	**13.924**	21.7%	*
	200	2,8,17,18	3,1,1,3	0.9, 0.6, 0.5, 0.8	60773.4, 7165.09, 1462.18	69400.6	3.7%	**18.5**	5.8%	6%
1	50	2,8,16,18	3,1,1,3	0.9, 0.6, 0.6, 0.9	61510.8, 5776.87, 909.356	68197.1	3.5%	**18.06**	3.5%	4.8%
	100	2,8,16,18	3,1,1,3	0.9, 0.6, 0.6, 0.9	61510.8, 5776.87, 1818.71	69106.4	4%	**23.996**	6%	6.1%
	200	2,8,17,18	3,1,1,3	0.9, 0.6, 0.5, 0.8	60773.4, 7165.09, 2382.41	70320.9	6.4%	**40.188**	7.8%	8.5%
2	50	2,8,17,18	3,1,1,3	0.9, 0.6, 0.5, 0.8	60773.4, 7165.09, 1285.78	69224.2	2.6%	**36.576**	6.5%	6.6%
	100	7,14,17,18	1,1,1,3	0.5, 0.4, 0.4, 0.7	60098.4, 9390.78, 786.621	70275.8	8%	**34.652**	9.8%	6.1%
	200	7,14,17,18	1,1,1,3	0.5, 0.4, 0.4, 0.7	60098.4, 9390.78, 1573.24	71062.4	61.668	69.608	**53.996**	56.637
					p=4, $\alpha = 0.8$					
0	50	2,8,17,18	3,1,1,3	0.9, 0.5, 0.5, 0.8	74143.6, 7165.09, 376.54	81685.2	1.8%	**53.9**	3.5%	3.7%
	100	2,8,17,18	3,1,1,3	0.9, 0.5, 0.5, 0.8	74143.6, 7165.09, 753.079	82061.7	2.7%	**41.792**	5.1%	8%
	200	7,14,17,18	1,1,1,3	0.5, 0.4, 0.4, 0.7	72648.8, 9390.78, 629.282	82668.9	2%	**53.464**	4.3%	4.5%
1	50	2,8,17,18	3,1,1,3	0.9, 0.5, 0.5, 0.8	74143.6, 7165.09, 619.434	81928.1	2.7%	**47.76**	4.1%	7.3%
	100	7,14,17,18	1,1,1,3	0.5, 0.4, 0.4, 0.7	72648.8, 9390.78, 432.636	82472.2	2.8%	**39.248**	4.8%	7.5%
	200	7,14,17,18	1,1,1,3	0.5, 0.4, 0.4, 0.7	72648.8, 9390.78, 865.272	82904.8	18.1%	**15.096**	42%	19.3%
2	50	7,14,17,18	1,1,1,3	0.5, 0.4, 0.4, 0.7	72648.8, 9390.78, 393.311	82432.9	3.7%	**47.904**	13.7%	4.9%
	100	7,14,17,18	1,1,1,3	0.5, 0.4, 0.4, 0.7	72648.8, 9390.78, 786.621	82826.2	2%	**68.824**	4.8%	6%
	200	7,14,17,18	1,1,1,3	0.5, 0.4, 0.4, 0.7	72648.8, 9390.78, 1573.24	83612.8	2.8%	**50.936**	4.9%	17.9%
				No. of best performing instances			1	32	2	1
					N=25T					
					p=3, $\alpha = 0.4$					
0	50	12,14,19	3,3,3	1, 0.7, 1	83485.5, 4545.57, 1376.68	89407.7	2.8%	**19.512**	*	*
	100	9,12,19	3,3,3	0.9, 0.8, 1	84249.1, 4394.85, 2087.51	90731.5	*	**8.14**	*	*
	200	9,12,25	3,3,3	0.7, 0.8, 0.9	84948.4, 4971.62, 1820.64	91740.7	3.1%	**20.96**	*	*
1	50	9,12,19	3,3,3	0.9, 0.8, 1	84249.1, 4394.85, 1951.3	90595.3	*	**8.068**	*	*
	100	9,12,25	3,3,3	0.7, 0.8, 0.9	84948.4, 4971.62, 1583.44	91503.5	4.1%	**19.326**	*	*
	200	9,12,25	3,3,3	0.7, 0.8, 0.9	84948.4, 4971.62, 3166.87	93086.9	2.3%	**18.256**	*	*
2	50	9,12,25	3,3,3	0.7, 0.8, 0.9	84948.4, 4971.62, 1801.39	91721.4	2.9%	**21.396**	*	*
	100	9,12,25	3,3,3	0.7, 0.8, 0.9	84948.4, 4971.62, 3602.77	93522.8	2.1%	**19.764**	*	*
	200	9,12,25	3,3,3	0.8, 0.8, 0.8	86542.2, 4971.62, 5433.94	96947.7	2%	**24.568**	*	*
					p=3, $\alpha = 0.8$					
0	50	12,14,19	3,3,3	1, 0.7, 1	92977.1, 4545.57, 1376.68	98899.4	**40.638**	41.11	40.2%	20.2%
	100	12,14,19	3,3,3	1, 0.7, 1	92977.1, 4545.57, 2753.35	100276	*	**24.216**	*	*
	200	12,14,19	3,3,3	0.9, 0.8, 1	94069.2, 4545.57, 3819.53	102434	4.2%	**34.826**	*	*
1	50	12,14,19	3,3,3	1, 0.7, 1	92977.1, 4545.57, 2621.03	100144	*	**30.952**	*	*
	100	12,14,19	3,3,3	0.9, 0.8, 1	94069.2, 4545.57, 3555.17	102170	4.3%	**30.578**	*	*
	200	12,14,19	3,3,3	0.9, 0.8, 0.9	96086.3, 4545.57, 4443.7	105076	6.8%	**37.612**	*	*
2	50	12,14,19	3,3,3	0.9, 0.8, 1	94069.2, 4545.57, 4245.7	102860	5.1%	**25.23**	*	*
	100	12,14,19	3,3,3	0.9, 0.8, 0.9	96086.3, 4545.57, 5160.54	105792	6.4%	**38.08**	*	*
	200	12,14,25	3,3,3	0.8, 0.6, 0.8	100341, 5122.34, 5399.39	110863	6.2%	**40.008**	*	*
					p=4, $\alpha = 0.4$					
0	50	2,12,14,19	2,3,2,3	0.7, 0.8, 0.9, 1	71993.2, 5046.17, 1015.4	78054.8	**31.451**	36.94	*	11.5%
	100	2,12,14,19	3,3,3,3	0.5, 0.8, 0.6, 1	71993.2, 5285.22, 1655.2	78933.7	2.6%	**27.284**	*	*
	200	2,12,14,19	3,3,3,3	0.5, 0.8, 0.6, 1	71993.2, 5285.22, 3310.4	80588.9	3.1%	**31.452**	*	21%
1	50	2,12,14,19	3,3,3,3	0.5, 0.8, 0.6, 1	71993.2, 5285.22, 1509.2	78787.7	6.1%	**19.97**	*	*
	100	2,12,14,19	3,3,3,3	0.5, 0.8, 0.6, 1	71993.2, 5285.22, 3018.4	80296.9	3.4%	**27.364**	*	*
	200	2,12,14,19	3,3,3,3	0.5, 0.8, 0.7, 0.9	74089.3, 5285.22, 3078.21	82452.7	30.3%	**36.924**	23.6%	23.9%
2	50	2,12,14,19	3,3,3,3	0.5, 0.8, 0.6, 1	71993.2, 5285.22, 3554.01	80832.5	6.2%	**30.398**	*	*
	100	2,12,14,19	3,3,3,3	0.5, 0.8, 0.7, 0.9	74089.3, 5285.22, 3412.22	82786.7	5.4%	**29.784**	*	*
	200	2,12,14,25	3,3,3,3	0.5, 0.7, 0.5, 0.8	75698.2, 5861.99, 4379.49	85939.7	6.9%	**29.778**	*	28.3%
					p=4, $\alpha = 0.8$					
0	50	2,12,13,19	2,3,3,3	0.7, 0.8, 0.9, 1	85729.8, 4675.97, 1083.56	91489.3	5.1%	**52.408**	11.6%	18.8%
	100	2,12,13,19	3,3,3,3	0.5, 0.8, 0.9, 1	85729.8, 4749.94, 2076.34	92556.1	32.7%	**43.4**	11.9%	9.1%
	200	12,13,19,23	3,3,3,3	0.9, 0.8, 0.9, 0.6	86603.4, 4914.24, 2336.89	93854.5	4.6%	**47.118**	15.4%	14.3%
1	50	2,12,13,19	3,3,3,3	0.5, 0.8, 0.9, 1	85729.8, 4749.94, 1915.65	92395.4	4.4%	**48.376**	15.7%	8.6%
	100	12,13,19,23	3,3,3,3	0.9, 0.8, 0.9, 0.6	86603.4, 4914.24, 2021.08	93538.7	27.7%	**29.52**	*	*
	200	12,13,19,23	3,3,3,3	0.9, 0.8, 0.9, 0.7	86908.9, 4914.24, 3715.95	95539.1	53.8%	**38.78**	18.1%	21%
2	50	12,13,19,23	3,3,3,3	0.9, 0.8, 0.9, 0.6	86603.4, 4914.24, 2289.49	93807.1	53.9%	**26.634**	*	*
	100	12,14,19,23	3,3,3,3	0.9, 0.5, 0.9, 0.7	86856.4, 5449.52, 3663.73	95969.7	6.8%	**51.56**	11.5%	13.2%
	200	12,14,19,23	3,3,3,3	0.9, 0.5, 0.9, 0.7	86856.4, 5449.52, 7327.46	99633.4	0.1%	**94.404**	10.5%	8%
				No. of best performing instances			2	34	0	0

* represents no solution in the given time limit.

Table 7: Comparison among EK-MISOCP4, 5, 6 and 9 for |N| = 50 (AP dataset)

c	θ	Hub	Cap	Intensity	Costs: FC, LC, CC	Costs: TC	CPU time (sec)/ % Gap — MISOCPs: 4	5	6	9
					N=50L					
					p=3, $\alpha = 0.4$					
	50	15,27,35	2,1,3	0.8, 0.8, 0.8	68783.3, 5656.71, 323.386	74763.4	3%	**116.774**	*	*
0	100	15,27,35	2,1,3	0.8, 0.8, 0.8	68783.3, 5656.71, 646.773	75086.8	3.2%	**338.784**	*	*
	200	15,28,35	2,1,3	0.7, 0.5, 0.8	68345.4, 6320.55, 913.826	75579.8	2.2%	**431.446**	28.9%	*
	50	15,27,35	2,1,3	0.8, 0.8, 0.8	68783.3, 5656.71, 529.873	74969.9	2.4%	**370.762**	*	*
1	100	15,28,35	2,1,3	0.7, 0.5, 0.8	68345.4, 6320.55, 714.416	75380.4	2.6%	**385.506**	*	*
	200	15,28,35	3,1,3	0.5, 0.5, 0.8	68345.4, 6515.28, 1118.55	75979.3	8.9%	**448.654**	*	9.4%
	50	15,28,35	2,1,3	0.7, 0.5, 0.8	68345.4, 6320.55, 743.462	75409.5	3.7%	**340.874**	70%	*
2	100	15,28,35	3,1,3	0.5, 0.5, 0.8	68345.4, 6515.28, 1131.45	75992.2	3.9%	**446.93**	*	*
	200	15,28,35	3,1,3	0.5, 0.5, 0.8	68345.4, 6515.28, 2262.91	77123.6	18.5%	**493.07**	10.5%	6.4%
					p=3, $\alpha = 0.8$					
	50	15,27,35	1,2,3	0.9, 0.7, 0.8	79532.5, 5465.73, 449.344	85447.6	2.7%	**537.106**	*	*
0	100	15,27,35	2,1,3	0.7, 0.8, 0.8	79495.2, 5656.71, 675.043	85827	3%	**280.716**	*	*
	200	15,27,35	2,2,3	0.7, 0.6, 0.8	79357.6, 5991.5, 1000.08	86349.1	3.6%	**452.032**	65.4%	*
	50	15,27,35	2,1,3	0.7, 0.8, 0.8	79495.2, 5656.71, 557.771	85709.7	3.7%	**258.748**	*	*
1	100	15,27,35	2,2,3	0.7, 0.6, 0.8	79357.6, 5991.5, 787.872	86136.9	3.1%	**649.472**	26.5%	*
	200	15,27,35	3,3,3	0.5, 0.4, 0.8	79357.6, 6310.22, 1128.85	86796.6	8.9%	**528.212**	*	*
	50	15,27,35	2,2,3	0.7, 0.6, 0.8	79357.6, 5991.5, 825.684	86174.8	3.4%	**367.834**	*	*
2	100	15,27,35	3,3,3	0.5, 0.4, 0.8	79357.6, 6310.22, 1152.41	86820.2	7.2%	**755.904**	6.7%	*
	200	15,27,35	3,3,3	0.5, 0.4, 0.8	79357.6, 6310.22, 2304.82	87972.6	5.3%	**685.908**	25.8%	7.8%
					p=4, $\alpha = 0.4$					
	50	6,27,32,35	1,2,3,2	0.7, 0.8, 0.6, 0.8	62898.9, 6414.66, 424.539	69738.1	*	**484.038**	*	*
0	100	6,27,32,35	1,3,3,2	0.7, 0.6, 0.6, 0.8	62898.9, 6538.65, 650.289	70087.8	6%	**955.1**	3.7%	3.1%
	200	6,27,32,35	1,3,3,3	0.7, 0.6, 0.6, 0.6	62898.9, 6850.95, 912.131	70662	10.5%	**281.668**	*	*
	50	6,27,32,35	1,3,3,2	0.7, 0.6, 0.6, 0.8	62898.9, 6538.65, 514.124	69951.7	5.1%	**780.996**	5.3%	8.3%
1	100	6,27,32,35	1,3,3,3	0.7, 0.6, 0.6, 0.6	62898.9, 6850.95, 664.834	70414.7	5.6%	**476.92**	45.9%	*
	200	6,27,32,35	1,3,3,3	0.7, 0.6, 0.6, 0.6	62898.9, 6850.95, 1329.67	71079.5	8.8%	**652.208**	51.1%	*
	50	6,27,32,35	1,3,3,3	0.7, 0.6, 0.6, 0.6	62898.9, 6850.95, 645.57	70395.4	*	**348.712**	*	*
2	100	6,27,32,35	1,3,3,3	0.7, 0.6, 0.6, 0.6	62898.9, 6850.95, 1291.14	71041	8.3%	**500.232**	*	*
	200	15,28,32,35	3,1,3,3	0.5, 0.5, 0.5, 0.6	62602.4, 7752.29, 1600.4	71955.1	*	**523.654**	*	*
					p=4, $\alpha = 0.8$					
	50	15,27,33,35	2,1,1,2	0.7, 0.8, 0.7, 0.8	74921.6, 6834.9, 377.662	82134.2	3.6%	**1440.12**	4.6%	4.7%
0	100	15,27,33,35	2,1,1,2	0.7, 0.8, 0.7, 0.8	74921.6, 6834.9, 755.323	82511.8	3.2%	**1127.72**	7.3%	4.9%
	200	15,27,33,35	2,2,1,2	0.7, 0.6, 0.7, 0.8	74831.4, 7169.68, 1243.22	83244.3	2.9%	**1609.54**	5.1%	3.5%
	50	15,27,33,35	2,1,1,2	0.7, 0.8, 0.7, 0.8	74921.6, 6834.9, 605.322	82361.8	2.9%	**1208.93**	5.6%	6.1%
1	100	15,27,33,35	2,2,1,2	0.7, 0.6, 0.7, 0.8	74831.4, 7169.68, 963.124	82964.2	4.4%	**1086.15**	4.9%	7.1%
	200	15,27,33,35	3,3,1,3	0.5, 0.4, 0.7, 0.5	74831.4, 7800.71, 1083.75	83715.8	4.9%	**938.108**	9.2%	7.7%
	50	15,27,33,35	2,2,1,2	0.7, 0.6, 0.7, 0.8	74831.4, 7169.68, 993.829	82994.9	3.7%	**1104.28**	5.5%	10.3%
2	100	15,27,33,35	3,3,1,3	0.5, 0.4, 0.7, 0.5	74831.4, 7800.71, 1028.23	83660.3	4%	**1067.78**	9%	9.2%
	200	15,27,32,35	3,3,3,3	0.5, 0.4, 0.5, 0.6	75366.6, 7547.24, 1503.58	84417.4	4.2%	**1224.02**	8.1%	6.7%
				No. of best performing instances			0	36	0	0
					N=50T					
					p=3, $\alpha = 0.4$					
	50	14,32,46	3,3,3	1, 1, 1	78946.8, 4108.17, 2600.49	85655.5	3.3%	**210.299**	*	*
0	100	25,32,46	3,3,3	0.9, 0.9, 1	80396.7, 4310.42, 2279.2	86986.3	28.7%	**77.007**	*	*
	200	25,32,46	3,3,3	0.9, 0.9, 1	80504.6, 4310.42, 4416.95	89232	23.9%	**120.364**	*	*
	50	25,32,46	3,3,3	0.9, 0.9, 1	80396.7, 4310.42, 2140.34	86847.5	4.5%	**253.749**	*	*
1	100	25,32,46	3,3,3	0.9, 0.9, 1	80504.6, 4310.42, 4139.33	88954.4	24.2%	**179.458**	*	*
	200	25,32,46	3,3,3	0.9, 0.9, 0.9	81082, 4310.42, 7221.97	92614.4	*	**149.601**	*	*
	50	25,32,46	3,3,3	0.9, 0.9, 0.9	81082, 4310.42, 4306.16	89698.6	18.6%	**239.085**	*	*
2	100	25,32,46	3,3,3	0.9, 0.9, 0.9	81082, 4310.42, 8612.34	94004.7	39.5%	**404.728**	*	*
	200	25,32,46	3,3,3	0.9, 0.9, 0.9	81082, 4310.42, 17224.7	102617	5.9%	**708.027**	*	*
					p=3, $\alpha = 0.8$					
	50	25,32,46	3,3,3	0.9, 0.9, 1	90540.8, 4310.42, 1139.6	95990.9	*	**215.879**	*	*
0	100	25,32,46	3,3,3	0.9, 0.9, 1	90540.8, 4310.42, 2279.21	97130.5	*	**226.852**	*	*
	200	25,32,46	3,3,3	0.9, 0.9, 1	90540.8, 4310.42, 4558.41	99409.7	*	**460.623**	*	*
	50	25,32,46	3,3,3	0.9, 0.9, 1	90540.8, 4310.42, 2140.34	96991.6	6.5%	**550.565**	*	*
1	100	25,32,46	3,3,3	0.9, 0.9, 1	90540.8, 4310.42, 4280.69	99131.9	2.5%	**320.46**	*	*
	200	25,32,46	3,3,3	0.9, 0.9, 0.9	91249.5, 4310.42, 7221.98	102782	1.2%	**845.82**	*	*
	50	25,32,46	3,3,3	0.9, 0.9, 0.9	91249.5, 4310.42, 4306.17	99866.1	4%	**211.736**	*	*
2	100	25,32,46	3,3,3	0.9, 0.9, 0.9	91249.5, 4310.42, 8612.34	104172	**895.252**	1167.57	*	*
	200	25,32,46	3,3,3	0.9, 0.9, 0.9	91249.5, 4310.42, 17224.6	112785	1.5%	**1219.12**	*	*
					p=4, $\alpha = 0.4$					
	50	14,32,35,38	3,2,3,3	0.9, 0.8, 0.9, 0.9	66111.9, 4848.88, 1070.9	72031.6	*	**102.578**	*	*
0	100	14,32,35,38	3,2,3,3	0.9, 0.8, 0.9, 0.9	66281.3, 4848.88, 1929.72	73059.9	1.3%	**188.701**	*	*
	200	14,32,35,38	3,3,3,3	0.9, 0.6, 0.9, 0.9	66516.6, 5008.57, 3185.12	74710.3	1.8%	**474.116**	*	*
	50	14,32,35,38	3,2,3,3	0.9, 0.8, 0.9, 0.9	66281.3, 4848.88, 1754.22	72884.4	1.2%	**320.526**	*	*
1	100	14,32,35,38	3,3,3,3	0.9, 0.6, 0.9, 0.9	66516.6, 5008.57, 2858.54	74383.7	1.5%	**378.696**	*	*
	200	14,32,38,46	3,3,3,3	0.8, 0.6, 0.8, 0.8	67753.9, 5245.61, 3259.03	76258.5	3.9%	**571.256**	*	*
	50	14,32,38,46	3,3,3,3	0.8, 0.5, 0.9, 0.9	67306.2, 5245.61, 2175.47	74727.3	1.3%	**535.173**	*	*
2	100	14,26,32,46	3,3,3,3	0.8, 0.8, 0.5, 0.8	68070.7, 5534.68, 2917.62	76523	7.3%	**441.907**	*	*
	200	14,26,32,46	3,3,3,3	0.7, 0.7, 0.6, 0.8	68502.3, 5534.68, 5194.29	79231.3	3.7%	**358.919**	*	*
					p=4, $\alpha = 0.8$					
	50	14,34,35,38	3,3,3,3	0.9, 0.9, 0.9, 1	78385.4, 4335.01, 1410.33	84130.7	2.4%	**511.334**	*	*
0	100	14,32,35,38	3,2,3,3	0.9, 0.8, 0.9, 0.9	78223.6, 4848.88, 2141.79	85214.3	2.5%	**1000.13**	*	*
	200	14,32,35,38	3,3,3,3	0.8, 0.6, 0.9, 0.9	78831, 5008.57, 3183.82	87023.4	6.6%	**927.888**	53.3%	*
	50	14,32,35,38	3,2,3,3	0.9, 0.8, 0.9, 0.9	78223.6, 4848.88, 1965.77	85038.2	4.1%	**617.19**	*	*
1	100	14,32,35,38	3,3,3,3	0.8, 0.6, 0.9, 0.9	78831, 5008.57, 2857.6	86697.2	5.1%	**778.654**	*	*
	200	14,32,38,46	3,3,3,3	0.8, 0.6, 0.8, 0.8	80251, 5245.61, 3262.68	88759.3	4.2%	**716.398**	*	*
	50	14,32,35,38	3,3,3,3	0.8, 0.7, 0.9, 0.8	79833.2, 5008.57, 2283.2	87125	3.4%	**870.678**	*	*
2	100	14,32,38,46	3,3,3,3	0.8, 0.6, 0.8, 0.8	80251, 5245.61, 3615.39	89112	6.1%	**939.039**	*	*
	200	14,26,32,46	3,3,3,3	0.7, 0.7, 0.6, 0.8	81345.8, 5534.68, 5161.03	92041.5	4.4%	**1578.46**	2.1%	4.1%
				No. of best performing instances			1	35	0	0

* represents no solution in the given time limit.

References

[1] N. Azizi, N. Vidyarthi, S. S. Chauhan, Modelling and analysis of hub-and-spoke networks under stochastic demand and congestion, Annals of Operations Research 264 (1-2) (2018) 1–40.

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

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

[4] S. Elhedhli, H. Wu, A lagrangean heuristic for hub-and-spoke system design with capacity selection and congestion, INFORMS Journal on Computing 22 (2) (2010) 282–296.

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

[6] M. E. OKelly, D. Bryan, Hub location with flow economies of scale, Transportation Research Part B: Methodological 32 (8) (1998) 605–616.

[7] A. Kimms, Economies of scale in hub & spoke network design models: We have it all wrong, in: Perspectives on operations research, Springer, 2006, pp. 293–317.

[8] C. B. Cunha, M. R. Silva, A genetic algorithm for the problem of configuring a hub-and-spoke network for a ltl trucking company in brazil, European Journal of Operational Research 179 (3) (2007) 747–758.

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

[10] M. W. Horner, M. E. O'Kelly, Embedding economies of scale concepts for hub network design, Journal of Transport Geography 9 (4) (2001) 255–265.

[11] D. Bryan, Extensions to the hub location problem: Formulations and numerical examples, Geographical Analysis 30 (4) (1998) 315–330.

[12] I. Racunica, L. Wynter, Optimal location of intermodal freight hubs, Transportation Research Part B: Methodological 39 (5) (2005) 453–477.

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

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

[15] A. T. Ernst, M. Krishnamoorthy, Solution algorithms for the capacitated single allocation hub location problem, Annals of Operations Research 86 (1999) 141–159.

[16] J. G. Klincewicz, Heuristics for the p-hub location problem, European Journal of Operational Research 53 (1) (1991) 25–37.

[17] J. G. Klincewicz, Avoiding local optima in thep-hub location problem using tabu search and grasp, Annals of Operations research 40 (1) (1992) 283–302.

[18] D. Skorin-Kapov, J. Skorin-Kapov, On tabu search for the location of interacting hub facilities, European Journal of Operational Research 73 (3) (1994) 502–509.

[19] H. Pirkul, D. A. Schilling, An efficient procedure for designing single allocation hub and spoke systems, Management Science 44 (12-part-2) (1998) S235–S242.

[20] J. Ebery, Solving large single allocation p-hub problems with two or three hubs, European Journal of Operational Research 128 (2) (2001) 447–458.

[21] J. F. Campbell, Location and allocation for distribution systems with transshipments and transportion economies of scale, Annals of operations research 40 (1) (1992) 77–99.

[22] A. T. Ernst, M. Krishnamoorthy, An exact solution approach based on shortest-paths for p-hub median problems, INFORMS Journal on Computing 10 (2) (1998) 149–162.

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

[24] A. T. Ernst, M. Krishnamoorthy, Exact and heuristic algorithms for the uncapacitated multiple allocation p-hub median problem, European Journal of Operational Research 104 (1) (1998) 100–112.

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

[26] M. Sasaki, M. Fukushima, On the hub-and-spoke model with arc capacity conatraints, Journal of the Operations Research Society of Japan 46 (4) (2003) 409–428.

[27] N. Boland, M. Krishnamoorthy, A. T. Ernst, J. Ebery, Preprocessing and cutting for multiple allocation hub location problems, European Journal of Operational Research 155 (3) (2004) 638–653.

[28] J. G. Klincewicz, Hub location in backbone/tributary network design: A review, Location Science 6 (1-4) (1998) 307–335.

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

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

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

[32] T. Aykin, Lagrangian relaxation based approaches to capacitated hub-and-spoke network design problem, European Journal of Operational Research 79 (3) (1994) 501–523.

[33] T. Aykin, Networking policies for hub-and-spoke systems with application to the air transportation system, Transportation Science 29 (3) (1995) 201–221.

[34] M. da Graça Costa, M. E. Captivo, J. Clímaco, Capacitated single allocation hub location problema bi-criteria approach, Computers & Operations Research 35 (11) (2008) 3671–3695.

[35] I. Contreras, J. A. Díaz, E. Fernández, Lagrangean relaxation for the capacitated hub location problem with single assignment, OR spectrum 31 (3) (2009) 483–505.

[36] I. Contreras, J. A. Díaz, E. Fernández, Branch and price for large-scale capacitated hub location problems with single assignment, INFORMS Journal on Computing 23 (1) (2011) 41–55.

[37] A. Marín, Formulating and solving splittable capacitated multiple allocation hub location problems, Computers & operations research 32 (12) (2005) 3093–3109.

[38] I. Rodríguez-Martín, J. J. Salazar-González, Solving a capacitated hub location problem, European Journal of Operational Research 184 (2) (2008) 468–479.

[39] H. Yaman, Star p-hub median problem with modular arc capacities, Computers & Operations Research 35 (9) (2008) 3009–3019.

[40] A. Hoff, J. Peiró, Á. Corberán, R. Martí, Heuristics for the capacitated modular hub location problem, Computers & Operations Research 86 (2017) 94–109.

[41] I. Correia, S. Nickel, F. Saldanha-da Gama, Single-assignment hub location problems with multiple capacity levels, Transportation Research Part B: Methodological 44 (8-9) (2010) 1047–1066.

[42] I. Correia, S. Nickel, F. Saldanha-da Gama, Hub and spoke network design with single-assignment, capacity decisions and balancing requirements, Applied Mathematical Modelling 35 (10) (2011) 4841–4851.

[43] I. Contreras, J.-F. Cordeau, G. Laporte, Exact solution of large-scale hub location problems with multiple capacity levels, Transportation Science 46 (4) (2012) 439–459.

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

[45] I. Correia, S. Nickel, F. Saldanha-da Gama, A stochastic multi-period capacitated multiple allocation hub location problem: Formulation and inequalities, Omega 74 (2018) 122–134.

[46] V. Marianov, D. Serra, Location models for airline hubs behaving as m/d/c queues, Computers & Operations Research 30 (7) (2003) 983–1003.

[47] S. Jayaswal, N. Vidyarthi, Capacitated multiple allocation hub location with service level constraints for multiple consignment classes, Tech. rep., Indian Institute of Management Ahmedabad, Research and Publication Department (2013).

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

[49] R. S. de Camargo, G. Miranda Jr, R. P. M. Ferreira, H. Luna, Multiple allocation hub-and-spoke network design under hub congestion, Computers & Operations Research 36 (12) (2009) 3097–3106.

[50] R. S. de Camargo, G. de Miranda Jr, R. P. Ferreira, A hybrid outer-approximation/benders decomposition algorithm for the single allocation hub location problem under congestion, Operations Research Letters 39 (5) (2011) 329–337.

[51] R. S. de Camargo, G. Miranda, Single allocation hub location problem under congestion: Network owner and user perspectives, Expert Systems with Applications 39 (3) (2012) 3385–3391.

[52] R. Kian, K. Kargar, Comparison of the formulations for a hub-and-spoke network design problem under congestion, Computers & Industrial Engineering 101 (2016) 504–512.

[53] H. Hasanzadeh, M. Bashiri, A. Amiri, A new approach to optimize a hub covering location problem with a queue estimation component using genetic programming, Soft Computing 22 (3) (2018) 949–961.

[54] J. F. Meier, U. Clausen, Solving single allocation hub location problems on euclidean data, Transportation Science 52 (5) (2017) 1141–1155.

[55] A. Ahmadi-Javid, P. Hoseinpour, Convexification of queueing formulas by mixed-integer second-order cone programming: An application to a discrete location problem with congestion, arXiv preprint arXiv:1710.05794.

[56] M. E. O'kelly, A quadratic integer program for the location of interacting hub facilities, European Journal of Operational Research 32 (3) (1987) 393–404.

[57] O. Günlük, J. Linderoth, Perspective reformulation and applications, in: Mixed Integer Nonlinear Programming, Springer, 2012, pp. 61–89.

[58] E. D. Dolan, J. J. Moré, Benchmarking optimization software with performance profiles, Mathematical programming 91 (2) (2002) 201–213.

[59] V. Marianov, D. Serra, C. ReVelle, Location of hubs in a competitive environment, European Journal of Operational Research 114 (2) (1999) 363–371.

[60] A. Lüer-Villagra, V. Marianov, A competitive hub location and pricing problem, European Journal of Operational Research 231 (3) (2013) 734–744.

[61] A. Atamtürk, G. Berenguer, Z.-J. Shen, A conic integer programming approach to stochastic joint location-inventory problems, Operations Research 60 (2) (2012) 366–381.

7 SHLPCC for the SK-based model

For SK-based model, the flow variables x_{ikm} are replaced by path variables x_{ijkm}, where $x_{ijkm} = \begin{cases} 1, & \text{if flows from } i \text{ to } j \text{ are routed via hub } k \text{ and } m \\ 0, & \text{otherwise.} \end{cases}$
Definition of other variables and parameters remain same.

7.1 Two-subscripted capacity allocation variable

$$\textbf{[SK-2s]} \quad \min \quad \sum_i \sum_j \sum_k \sum_m F_{ijkm} x_{ijkm} + \sum_k \sum_l Q_k^l y_{kl} + \theta \sum_k 1/2E[N_k(y,z)]$$

$$\text{s.t.} \quad (3) - (5), (7) - (10), (12), (14) - (21)$$

$$\sum_m x_{ijkm} = z_{ik} \qquad \forall i, j, k \tag{76}$$

$$\sum_k x_{ijkm} = z_{jm} \qquad \forall i, j, m \tag{77}$$

$$x_{ijkm} \in 0, 1 \qquad \forall i, j, k, m, l \tag{78}$$

Here, $F_{ijkm} = W_{ij}(\chi d_{ik} + \alpha d_{km} + \delta d_{mj})$ is the total flow through path $i - j - k - m$. (76) and (77) connect the assignment variables and path variables.
[SK-MISOCP1] (3)-(5), (7), (9)-(10), (12), (14)-(21), (37)-(39), (76)-(78)
[SK-MISOCP2] (3)-(5), (7), (9)-(10), (12), (14)-(21), (41)-(44), (76)-(78).
[SK-MISOCP3] (3)-(5), (7), (9)-(10), (12), (14)-(21), (41)-(43), (46), (76)-(78).
[SK-MISOCP4] (3)-(5), (7), (9)-(10), (12), (14)-(21), (41)-(43), (47), (49), (50), (76)-(78).
[SK-MISOCP5] (3)-(5), (7)-(10), (12), (14)-(21), (52)-(54), (56), (76)-(78).

7.2 Three-subscripted capacity allocation variable

$$\textbf{[SK-3s]} \quad \min \quad \sum_i \sum_j \sum_k \sum_m F_{ijkm} x_{ijkm} + \sum_k \sum_l Q_k^l t_{kk}^l + \theta E[N_k]$$

$$\text{s.t.} \quad (26) - (28), (30), (32), (34)$$

$$\sum_m x_{ijkm} = \sum_l t_{ik}^l \qquad \forall i, j, k \tag{79}$$

$$\sum_k x_{ijkm} = \sum_l t_{jm}^l \qquad \forall i, j, m \tag{80}$$

$$x_{ijkm} \in 0, 1 \qquad \forall i, j, k, m, l \tag{81}$$

[SK-MISOCP6] (26)-(28), (30), (32),(34), (58)-(60), (79)-(81).
[SK-MISOCP7] (26)-(28), (30), (32),(34), (62)-(64), (79)-(81).
[SK-MISOCP8] (26)-(28), (30), (32),(34),(66)-(70), (79)-(81).
[SK-MISOCP9] (26)-(28), (30), (32),(34), (71)-(73), (75), (79)-(81).

Table 8: Computation time for N=25

c	t0	Hub	Cap	Costs		CPU time (sec)										
						SK	SK	SK-MISOCP								
				FC, LC, CC	TC	OA-2s	OA-3s	1	2	3	4	5	6	7	8	9
							p=4, a=0.4									
0	20	1,4,12,17	2,3,2,3	695, 900, 109	1704	1299	1334	1727	2952	2334	1601	3647	2290	2344	*	1660
	50	1,4,12,17	3,3,2,3	795, 950, 219	1963	2531	3001	4203	7094	5006	4020	5525	4254	9471	*	7612
1	20	1,4,12,17	2,3,2,3	792, 900, 165	1857	3000	3531	3816	3771	4101	3002	3366	3457	4993	*	2294
	50	1,4,12,17	3,3,3,3	803, 1000, 247	2050	4474	7000	8092	7735	11477	10445	8077	*	*	*	*
2	20	1,4,12,17	3,3,3,3	803, 1000, 185	1988	4691	6531	6539	*	*	11061	2623	1988	*	*	*
	50	1,4,12,17	3,3,3,3	824, 1000, 420	2244	5855	9001	*	*	*	*	2856	2244	*	*	*

* represents no solution in the given time limit.
SK-OA-2s represents outer-approximation results for the two-subscripted model based on SK.

8 OA Method

8.1 EK-OA-2s: OA-based method for EK-2s

The auxillary variable L_{kl} and ρ_k which were defined as $L_{kl} = \rho_k y_{kl}$ and $\rho_k = \frac{s_k}{1+sk}$, imply

$$L_{kl} = \begin{cases} 0, & \text{if } y_{kl} = 0 \\ \frac{s_k}{1+s_k}, & \text{if } y_{kl} = 1. \end{cases} \tag{82}$$

Also, earlier results , $\sum_l L_{kl} = \rho_k \quad \forall k$, $L_{kl} \leq y_{kl} \quad \forall k, l$ and $\sum_i \sum_j W_{ij} z_{ik} = \sum_l \gamma_k^l L_{kl}$, remain. The function, $L_{kl} = \frac{s_k}{(1+s_k)}$ is a concave function which can be approximated with piecewise linear functions that are basically tangents to the function L_{kl} . The method chooses the minimum of these tangents at points $s_k^h \quad \forall h \in H, k \in N, l \in L$ to best approximate the function. The cuts are given by

$$\begin{aligned} L_{kl} &= \min_{h \in H} \left\{ \frac{1}{(1+s_k^h)^2} s_k + \left(\frac{s_k^h}{1+s_k^h} \right)^2 \right\} \\ \iff L_{kl} &\leq \frac{1}{(1+s_k^h)^2} s_k + \left(\frac{s_k^h}{1+s_k^h} \right)^2 \quad \forall k \in N, l \in L, h \in H \end{aligned} \tag{83}$$

In the OA-based method proposed by (48), for every $k - l$ pair, the non linear congestion term is approximated with tangents (cuts), given by (82), at points s_k^h . Here $h \in H$ denotes the indexing of these points. At every iteration a relaxed mixed integer linear problem is solved. The solution of which not only gives the lower bound but also supplies information for the next cut. Also, this solution is feasible for the main problem thus giving the upper bound. The algorithm terminates when both upper and lower bounds are ϵ (or less) away from each other where $\epsilon \geq 0$. (57) proposed perspective counterpart for (83) as

$$L_{kl} = \frac{s_k}{1 + s_k / y_k}$$

and the corresponding perspective cut at s_k^h as

$$L_{kl} \leq \frac{1}{(1+s_k^h)^2} s_k + \left(\frac{s_k^h}{1+s_k^h} \right)^2 y_{kl}, \quad \forall k, l \tag{84}$$

The formulation for the OA-2s is as follows:

$$\begin{aligned}
\textbf{[EK-OA-2s]} \quad \min \quad & \sum_i \sum_k C_{ik}(\chi O_i + \delta D_i) z_{ik} + \sum_i \sum_k \sum_m \alpha C_{km} x_{ikm} \\
& + \sum_k \sum_l Q_k^l y_{kl} + \theta/2 \sum_k \left\{ s_k + \rho_k + \sum_l c_{kl}^2 (V_{kl} - L_{kl}) \right\} \\
\text{s.t.} \quad & (3) - (21), (84)
\end{aligned}$$

Algorithm for the above discussed OA-based method is as follows:

1: Set $UB^{\tau-1} \to \infty$, $LB^{\tau-1} \to -\infty$, $\tau \to 0$

2: Choose initial cuts at points $s_k^{h_0}$ where $h_0 \in H$

3: **while** $\frac{UB^{\tau-1} - LB^{\tau-1}}{UB^{\tau-1}} \geq \epsilon$ **do**

4: Find LB^τ by solving $(OA{-}2s)^\tau$ and obtain the optimal solution $(x^\tau, z^\tau, y^\tau, \rho^\tau, s^\tau, L^\tau, V^\tau)$

5: Find UB^τ by substituting (x^τ, z^τ, y^τ) in the objective function as

$$\begin{aligned}
UB^\tau = \min \Bigg\{ & UB^{\tau-1}, \sum_i \sum_j \sum_k \sum_m F_{ijkm} x_{ijkm}^\tau + \sum_k \sum_l Q_k^l y_{kl}^\tau \\
& + 1/2 \Bigg\{ \Big(1 + \sum_l c_{kl}^2 y_{kl}^\tau\Big) \frac{\sum_i \sum_j W_{ij} z_{ik}^\tau}{\Big(\sum_l \gamma_k^l y_{kl}^\tau - \sum_i \sum_j W_{ij} z_{ik}^\tau\Big)} + \Big(1 - \sum_l c_{kl}^2 y_{kl}^\tau\Big) \frac{\sum_i \sum_j W_{ij} z_{ik}^\tau}{\sum_l \gamma_k^l y_{kl}^\tau} \Bigg\} \Bigg\}
\end{aligned}$$

6: Update new point $s_k^{h_{new}} = \frac{\sum_i \sum_j W_{ij} z_{ik}^\tau}{\sum_l \gamma_k^l y_{kl}^\tau - \sum_i \sum_j W_{ij} z_{ik}^\tau}$ with the current solution (x^τ, z^τ, y^τ).

7: Generate new cut: $L_{kl} \leq \frac{1}{(1+s_k^{h_{new}})^2} s_k + \left(\frac{s_k^{h_{new}}}{1+s_k^{h_{new}}} \right)^2 y_{kl}, \quad \forall k, l$

8: Add new cut: $H^{\tau+1} \to H^\tau + h_{new}$

9: $\tau \to \tau + 1$

10: **end while**

8.2 EK-OA-3s: OA-based method for EK-3s

For formulations with t_{ik}^l, we had objective function as

$$\begin{aligned}
\min \quad & \Big(\sum_i \sum_k C_{ik}(\chi O_i + \delta D_i) \sum_l t_{ik}^l\Big) + \sum_i \sum_k \sum_m \alpha C_{km} x_{ikm} + \sum_k \sum_l Q_k^l t_{kk}^l \\
& + \theta \sum_k \sum_l 1/2 \Bigg\{ \Big(1 + c_{kl}^2\Big) \frac{\sum_i \sum_j W_{ij} t_{ik}^l}{\Big(\gamma_k^l - \sum_i \sum_j W_{ij} t_{ik}^l\Big)} + \Big(1 - c_{kl}^2\Big) \frac{\sum_i \sum_j W_{ij} t_{ik}^l}{\gamma_k^l} \Bigg\}
\end{aligned}$$

We introduce variable ρ_k and s_{kl} such that

$$\frac{\sum_i \sum_j W_{ij} t_{ik}^l}{\gamma_k^l} \leq \rho_{kl} \qquad \forall k, l \tag{85}$$

$$\frac{\sum_i \sum_j W_{ij} t_{ik}^l}{\gamma_k^l - \sum_i \sum_j W_{ij} t_{ik}^l} \leq s_{kl} \quad \forall k, l \tag{86}$$

ρ_{kl} and s_{kl} are related as

$$\rho_{kl} = \frac{s_{kl}}{1 + s_{kl}} \quad \forall k, l$$

which is non-linear concave function and can be approximated using tangent hyperplanes as discussed in the previous section. For perspective reformulation, we introduce the following constraint

$$\rho_{kl} \leq t^l_{kk} \tag{87}$$

to the following equivalent perspective form

$$\rho_{kl} = \frac{s_{kl}}{1 + (s_{kl}/t^l_{kk})} \tag{88}$$

By following logic similar to (83), we have perspective cuts as

$$\rho_{kl} \leq \frac{1}{(1 + s^h_{kl})^2} s_{kl} + \frac{(s^h_{kl})^2}{(1 + s^h_{kl})^2} t^l_{kk}, \quad \forall k, l \tag{89}$$

The overall formulation for the approximation method is as follows:

$$\textbf{[EK-OA-3s]} \quad \min \sum_i \sum_k C_{ik}(\chi O_i + \delta D_i) \sum_l t^l_{ik}) + \sum_i \sum_k \sum_m \alpha C_{km} x_{ikm} + \sum_k \sum_l Q^l_k t^l_{kk}$$
$$+ \theta \sum_k \sum_l 1/2 \left\{ \left(1 + c^2_{kl}\right) s_{kl} + \left(1 - c^2_{kl}\right) \rho_{kl} \right\}$$
$$\text{s.t.} \quad (26) - (34), (87), (89)$$

Algorithm for the above discussed OA-based method is as follows:

1: Set $UB^{\tau-1} \to \infty$, $LB^{\tau-1} \to -\infty$, $\tau \to 0$

2: Choose initial cuts at points $s^{h_0}_k$ where $h_0 \in H$

3: **while** $\frac{UB^{\tau-1} - LB^{\tau-1}}{UB^{\tau-1}} \geq \epsilon$ **do**

4: Find LB^τ by solving $(OA - 3s)^\tau$ and obtain the optimal solution $(x^\tau, t^\tau, \rho^\tau, s^\tau)$

5: Find UB^τ by substituting (x^τ, t^τ) in the objective function as

$$UB^\tau = \min \left\{ UB^{\tau-1}, \sum_i \sum_k C_{ik}(\chi O_i + \delta D_i) \sum_l t^{l(\tau)}_{ik}) + \sum_i \sum_k \sum_m \alpha C_{km} x^\tau_{ikm} + \sum_k \sum_l Q^l_k t^{l(\tau)}_{kk} \right.$$
$$\left. + 1/2 \left\{ \left(1 + c^2_{kl}\right) \frac{\sum_i \sum_j W_{ij} t^{l(\tau)}_{ik}}{\left(\gamma^l_k - \sum_i \sum_j W_{ij} t^{l(\tau)}_{ik}\right)} + \left(1 - c^2_{kl}\right) \frac{\sum_i \sum_j W_{ij} t^{l(\tau)}_{ik}}{\gamma^l_k} \right\} \right.$$

6: Update new point $s^{h_{new}}_k = \frac{\sum_i \sum_j W_{ij} t^{l(\tau)}_{ik}}{\sum_l \gamma^l_k - \sum_i \sum_j W_{ij} t^{l(\tau)}_{ik}}$ with the current solution (x^τ, t^τ).

7: Generate new cut: $\rho_{kl} \leq \frac{1}{(1+s^{h_{new}}_{kl})^2} s_{kl} + \frac{(s^{h_{new}}_{kl})^2}{(1+s^{h_{new}}_{kl})^2} t^l_{kk}, \quad \forall k, l$

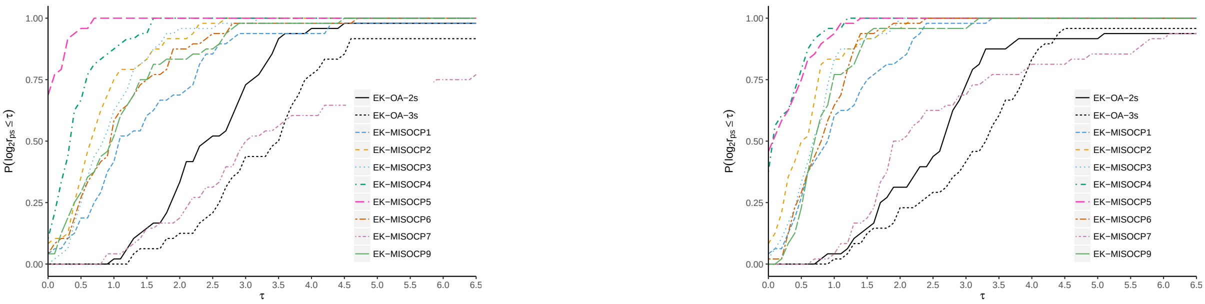
8: Add new cut: $H^{\tau+1} \to H^\tau + h_{new}$

9: $\tau \to \tau + 1$

10: **end while**

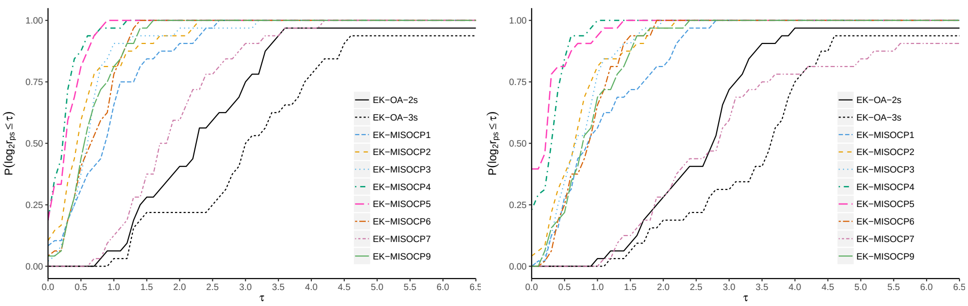
9 Performance Profile for coefficient of variation (c), and unit congestion cost θ.

Figure 2: Performance Profile of EK-MISOCPs and EK-OA-based method for θ= 20, 50

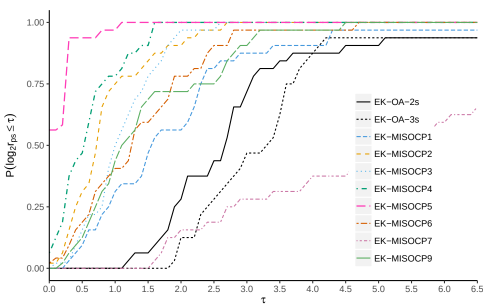


(a) t=20 **(b) t=50**

Figure 3: Performance Profile of EK-MISOCPs and EK-OA-based method for c= 0, 1, 2



(a) c=0 **(b) c=1**



(c) c=2