AMERICAN UNIVERSITY OF BEIRUT

MULTICOMMODITY HUB-AND-SPOKE NETWORK DESIGN

by

SABINE ANTOINE KHALIL

A thesis
submitted in partial fulfillment of the requirements
for the degree of Master of Engineering Management
to the Engineering Management Program
of the Faculty of Engineering and Architecture
at the American University of Beirut

Beirut, Lebanon
May 2014

AMERICAN UNIVERSITY OF BEIRUT

MULTICOMMODITY HUB-AND-SPOKE NETWORK DESIGN

by

SABINE ANTOINE KHALIL

Approved by:

Dr. Joe Naoum-Sawaya, Assistant Professor — Advisor
Engineering Management

Dr. Bacel Maddah, Associate Professor — Committee Member
Engineering Management

Dr. Walid Nasr, Assistant Professor — Committee Member
Business Administration

Date of thesis defense: May 9^{th}, 2014

AMERICAN UNIVERSITY OF BEIRUT

THESIS, DISSERTATION, PROJECT RELEASE FORM

Student Name: KHALIL SABINE ANTOINE
Last First Middle

● Master's Thesis ○ Master's Project ○ Doctoral Dissertation

☒ I authorize the American University of Beirut to: (a) reproduce hard or electronic copies of my thesis, dissertation, or project; (b) include such copies in the archives and digital repositories of the University; and (c) make freely available such copies to third parties for research or educational purposes.

☐ I authorize the American University of Beirut, **three years after the date of submitting my thesis, dissertation, or project,** to: (a) reproduce hard or electronic copies of it; (b) include such copies in the archives and digital repositories of the University; and (c) make freely available such copies to third parties for research or educational purposes.

14-May-2014

Signature Date

This form is signed when submitting the thesis, dissertation, or project to the University Libraries

ACKNOWLEDGEMENTS

I would like to express my deepest appreciation to my advisor Prof. Joe Naoum-Sawaya for the useful remarks and full engagement throughout the process of my thesis. I would like to thank him for introducing me to this topic, motivating me and pushing me to learn new things. I wouldn't have been able to accomplish this work if it weren't for his kind patience, endless support and continuous guidance.
In addition, I would like to thank my committee members Prof. Bacel Maddah and Prof. Walid Nasr for their constructive comments, insightful suggestions and valuable advice.

AN ABSTRACT OF THE THESIS OF

Sabine Antoine Khalil for Master of Engineering Management
Major: Engineering Management

Title: Multicommodity hub-and-spoke network design

Hub-and-spoke is a network architecture with several applications in a number of industries. This type of networks allows the flow of information, people, and products from origin to destination points through central nodes known as hubs. The consolidation of flow at central locations leads to significant reduction in cost due to the economies of scale. Although in practice, several types of commodities flow over such networks, the majority of models in the literature consider single commodity type networks. Dealing with a single commodity is a limitation for such networks, as it is impossible to model multiple types of flow and it does not allow to model multiple source and destination pairs which is primordial in all sorts of fields. In this research, we present an optimization model for multicommodity hub-and-spoke network design. We present computational testing on a set of instances randomly generated with the Mulgen generator, in addition to a real case network inspired by the French rail network. The results show that small to medium-size networks can be solved within a reasonable computational time. Furthermore, since the major drawback of hub-and-spoke networks is congestion on hubs, we present an extension to the model that includes congestion costs, in order to mitigate the effect of congestion in the network. Since the resulting optimization problem is non-linear, we then present a cutting plane approach based on a piecewise linear approximation as a solution approach.

Keywords: Hub-and-spoke networks, multicommodity, congestion, integer programming, non-linear optimization.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION

A. Hub-and-spoke Networks

Over the last few decades, innovation in industries such as telecommunication, technology, logistics, and transportation has been at its best. Significant research has been devoted to improving the service level while decreasing the total costs. For instance, in industries such as logistics, postal deliveries, and airline companies, the transportation costs are the primordial focus. Thus, hub-and-spoke networks witnessed a lot of attention in these fields since they offer an efficient way to utilize resources at lower costs.

A hub-and-spoke network is a special kind of network examining the demand flows between different origins and destinations. The hub represents the central point, while the spokes are all the nodes surrounding the hub and linked to it. These networks aim to minimize the aggregate costs as well as to maximize the service level. Routing the flows through intermediate nodes, i.e. the hubs, before attaining their destinations is the main characteristic of a hub-and-spoke network (Figure I.1). The difficulty comes from identifying the nodes that should be designated as hubs and the implicated routing. After assigning the hubs, the spokes are linked to them to allow the flow of commodities in the network. As shown in Figure I.1, two types of hub-and-spoke networks exist. The first one (a) being a single allocation network where spokes can only be allocated to one hub and (b) being a multiple allocation network where a spoke might be allocated to multiple hubs.

The progress made in the hub-and-spoke networks has been acknowledged as an extremely primordial innovation in the industry. Hubs serve as centers that handle the

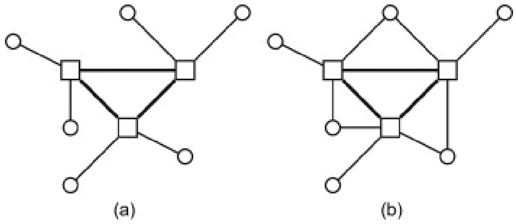

Figure I.1: (a) Single Allocation Network; (b) Multiple Allocations Network

complicated processes such as consolidating and sorting packages before transshipping them to their final destinations. Thus, instead of repeating the same procedures at every node, they are limited to one central node; the hub. Furthermore, it is assumed that inter-hub connections have a lower price per unit than hub-to-spoke connections due to the economies of scale. This reduction in the number of connections leads to not only a simplified network structure but also to a better performance and an improved use of the resources. The main advantage of implementing such models is the reduction of the aggregate costs; transportation costs, labor costs, holding costs, among others.

Due to the hub-and-spoke network's efficient features, those models are used widely in several industries. They are found mostly in transportation networks [32] as well as in telecommunication networks [28]. In addition, hub-and-spoke models are applied in urban traffic networks [35], trucking systems [43], postal networks [17], express delivery service networks [25], among other industries. The transportation industry, for instance, includes air, rail and road transportation.

In the air transportation industry, hubs typically correspond to central airports where flights are directed through and spokes are the paths taken by the airplanes out of the hub airport [6]. Figure I.2 is an example of a hub-and-spoke network implemented by Delta Airlines across North of America. As shown in this figure, Delta Airlines has several hubs such as Minneapolis, Detroit, Atlanta, Salt Lake City. In the rail industry, a hub corresponds to a central train station at major cities. Figure I.3 represents a rail map of Europe where, for example, Madrid, Paris, Frankfurt represent hubs. In the road transportation industry, such as bus transportation, hubs typically represent a central bus station at major cities. Moreover, they are exploited in the urban traffic networks where

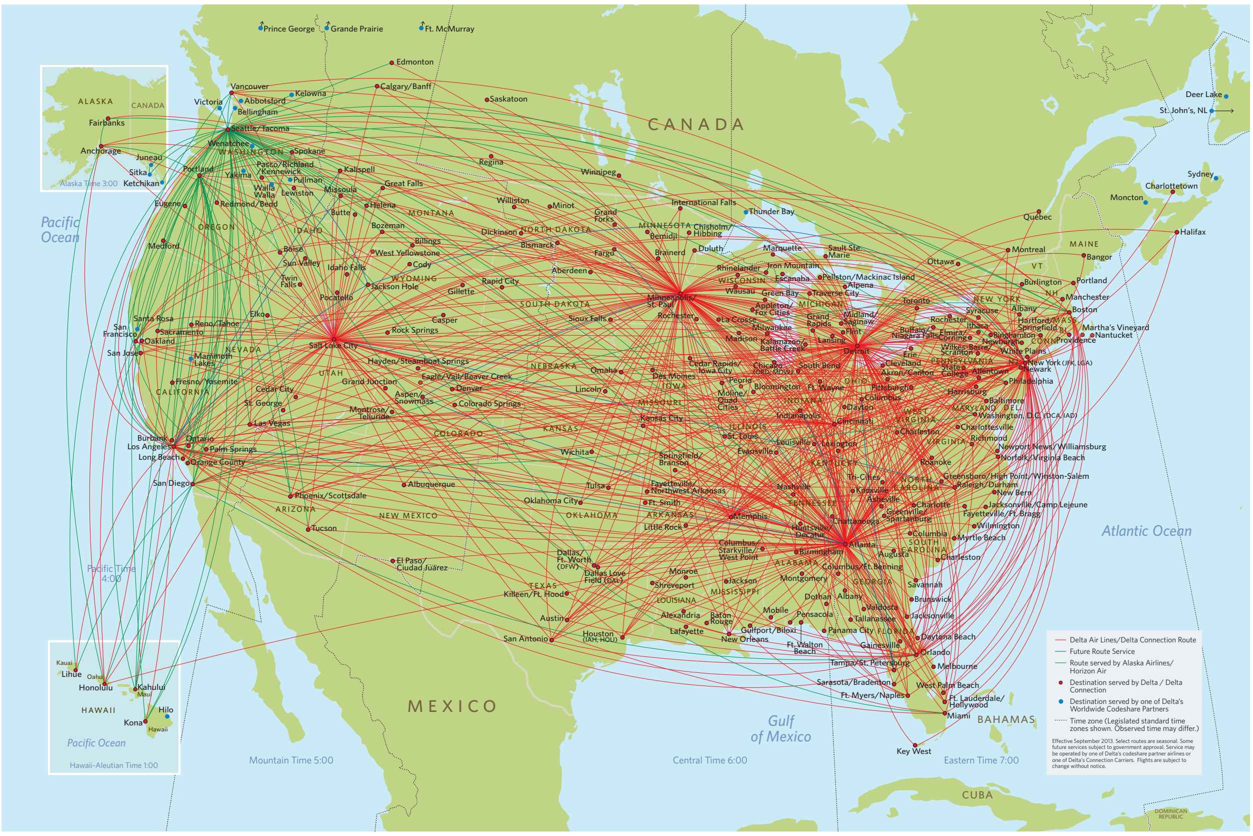

Figure I.2: Delta Airlines Route Map across North of America

the hub denotes a transit stop for several routes [35]. In the trucking networks and systems, the hub represents simply a center facility or a warehouse [43]. In the postal service networks, the hub is the central post office that will receive, sort and transmit mail [17]. In the express delivery service networks, the hub is a center for sorting and switching the operations [25]. Figure I.4 is an example of a map for the bus routes in North of America. In the telecommunication industry, the same concept of hub-and-spoke is applied but with different aims; it can be a central website or a central router, for example. Figure I.5 exemplifies the hub-and-spoke structure worldwide.

In addition to that, constructing hub-and-spoke networks is extremely costy. Table 1 illustrates typical costs for rail networks [13]. These costs are extremely variable. As shown in Table 1, they might vary from $9 million to $250 million per km. This brings us to the conclusion that it is critical to identify the optimal design of a hub-and-spoke

Figure I.3: Europe Rail Map

network, mainly due to the high costs and the long-term commitment. The models that are presented in this thesis are steps in that direction.

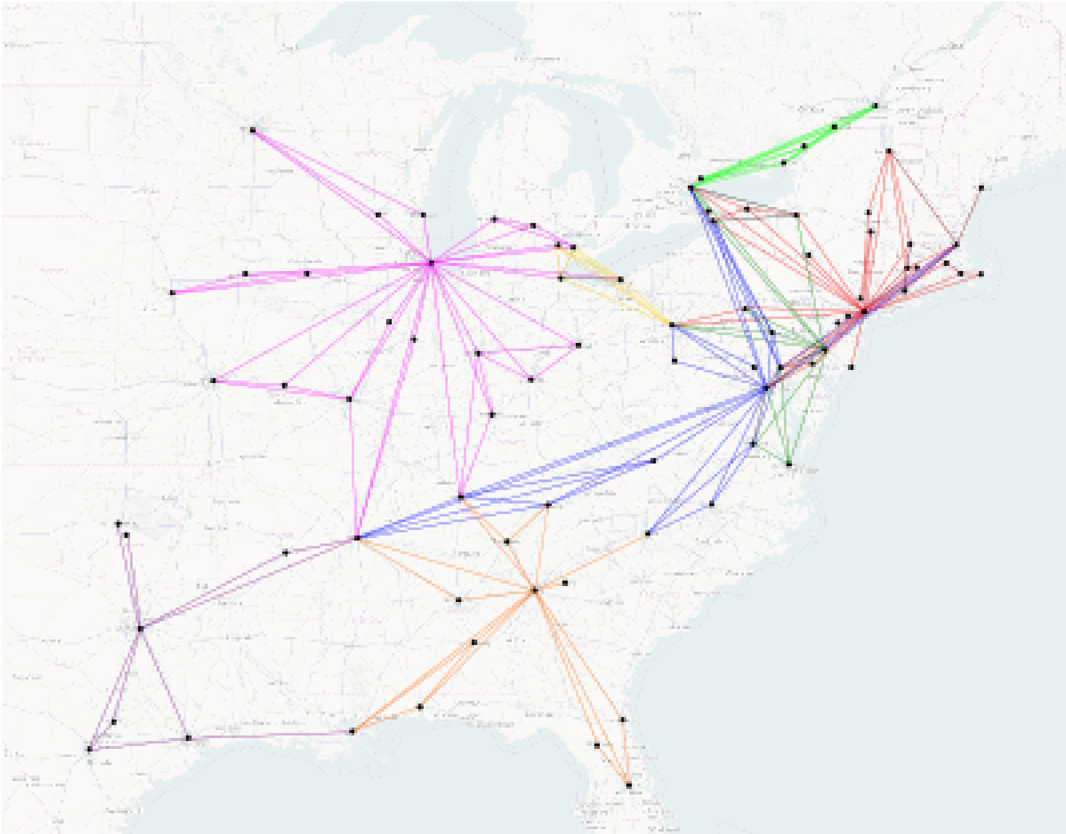

Figure I.4: Bus Map of North of America

Figure I.5: Worldwide Internet Network

Table 1: Railway Construction Costs

Railway	Date	Type of System	Cost per km	Distance
Madrid-Albacete	2010	High Speed Line	9.57 million	304 km
Seoul-Gimpo	2010	Airport Line	98.1 million	20.4 km
Yichang-Wanzhou	2011	Main Line	9.1 million	377 km
Haikou-Sanya	2010	High Speed Line	10 million	308 km
Copenhagen	2011	New Metro Line	247.5 million	16 km

B. Multicommodity Networks

Hub-and-spoke networks deal with the demand flow from origin to destination. The majority of research on hub-and-spoke networks deals with a single commodity type network, while in reality multiple commodities are transported, hence it becomes necessary to model multicommodity hub-and-spoke networks.

Multicommodity networks handle a set of K commodities that should be shipped from a source (origin) to a sink (destination). Each commodity has a particular quantity that must be shipped through the network to ultimately reach its final destination. The advantage of a multicommodity network is that it applies to real-life activities where multiple types of products flow through a certain network. In addition, routing multicommodities is not trivial since these commodities interact by competing for the actual arc capacity, when they flow on the same arcs [1].

Multicommodity networks are more practical than single-commodity networks. They are mostly exploited in telecommunication where multiple types of data need to be transferred through the network. Logistics companies implement multicommodity networks to be able to deliver all the different goods to their final destinations. In traffic engineering, the aim is to minimize the maximal linkage usage in the network to supply as much traffic as possible. In addition to that, modeling multiple sources and destinations of flow is equivalent to having multicommodities. This is primordial in airlines networks, rail networks or even bus networks.

However, the majority of research on multicommodity networks focuses on point-by-point networks. Thus, these networks do not exploit economies of scale as hub-and-spoke networks do. Building a model consisting of hub-and-spoke and multicommodity networks would allow us to benefit from the advantages of both networks.

The contributions of this thesis are:

- A hub-and-spoke network design model that takes into account multicommodity flows.
- A model that allows an arbitrary number of hubs between pairs of spokes.
- A model that explicitely accounts for the cost of congestion.
- A cutting plane solution approach for the presented model, based on a piecewise linear approximation.

The rest of this document is divided as follows. Chapter 2 introduces some literature review on hub-and-spoke networks as well as multicommodity networks. Chapter 3 presents a detailed mathematical formulation of the model. Chapter 4 illustrates the linearization of congestion before adding it to our model. Chapter 5 shows the computational results of our model. Finally, Chapter 6 concludes.

CHAPTER II

LITERATURE REVIEW

A. Hub-and-Spoke Extensions

Hub-and-spoke networks are extensively studied in the literature. Goldman [21] introduced the network hub location problem and O'kelly [38] was the first to formulate a hub-and-spoke mathematical model by studying passenger airline networks. O'Kelly [37] defined the hub-and-spoke network design to be "a complex mixture of locational analysis and spatial interaction theory." Hub capacities, hub facility locations, hub assignments, hub interlinkages as well as routes for origin-destination demand flows, are construed through a properly designed hub-and-spoke network. O'Kelly's [38] formulation represents the *single allocation p-hub median* problem which is the most fundamental model for hub-and-spoke networks. In a graph consisting of nodes and edges, each node sends and receives some data to/from the other nodes of the graph. The problem lies in choosing p of the given nodes to act as hubs with the remaining nodes connected to one of the chosen hubs.

Since then, research on hub-and-spoke networks has been divided into two categories. Some extended the existing models by including new features, while others designed methods in order to produce more efficient solutions. Campbell [7] improved the model presented by O'kelly [38] by allowing multiple allocations in the hub-and-spoke problem and studied the location and allocation to distribute a uniform demand with transshipments. Aykin [2] incorporated capacities to hubs. In addition, Kara and Tansel [24] presented the *p-hub central problem*, a model that aims to minimize the maximum hub origin/hub destination. Kara and Tansel [25] also focused on minimizing the arrival time of the items in the cargo in addition to developing a model that correctly calculates those arrival times. Other extensions include direct shipments [3] and modeling the non-linearity of economies of scale [39]. Liu et al. [31] created a new hybrid model for hub-and-spoke networks in truck delivery. This model constitutes a new method implemented along with the traditional hub-and-spoke model where vehicle routing will be performed through the two delivery modes. More recently, Campbell et al. [9] formulated new models based on hub arc location instead of hub facility location. It consists of creating new optimal solutions while paying attention to spatial patterns.

B. Hub-and-Spoke Solution Methods

Two different solution approaches were presented by O'kelly [38]. The first assigns the closest hub to every non-hub demand node while the second assigns the closest or the second closest hub to every demand node. Klincewicz [26] presented a new heuristic for the p-hub location problem that consists of having a primordial set of hubs, then automatically substituting other nodes for that set, based on local improvement measures. The exchange heuristic terminates once a local optima (i.e. an optimal solution having a neighboring set of candidate solutions) is found. Klincewicz [27] presented new heuristics for the same type of problems, based on tabu search as well as on a greedy randomized adaptive search procedure (GRASP). These heuristics differ from the previous ones since they examine several local optima in order to find better solutions. Concerning the p-hub median location problem, two greedy-interchange heuristics were formulated by Campbell [8]. The starting point of these heuristics is the solution of the multiple allocation p-hub median problem. The thought behind this idea is that multiple allocation problems are easier to solve due to the greater freedom in allocation allowed. A subgradient based Lagrangian heuristic for the single allocation p-hub problem was developed by Pirkul and Schilling [40]. This method yields solutions of very high quality, as well as a reasonable computational time. Liu et al. [31] discussed a new hybrid model for hub-and-spoke networks in the truck delivery sector. Their method consists of a two-step solution heuristic.

First, it treats the model as a pure hub-and-spoke network and as a node-to-node network, and solves the networks respectively. Then, the starting point for the hybrid network is considered to be a better solution than the previous step. Topcuoglu et al. [44] adapted a new solution method based on Genetic Algorithm (GA). Its purpose is reaching the optimal solution of the *uncapacitated single assignment hub location problem*. In this GA based heuristic, a chromosome string is constructed from a binary hub array and a binary hub-node assign array.

Elhedhli and Wu [16] presented a Lagrangian heuristic for the hub-and-spoke system with capacity selection and congestion. To solve this problem with this heuristic, Elhedhli and Wu [16] decomposed it into an easy subproblem and a more difficult non-linear subproblem. Gelareh and Nickel [18] presented the hub location model for public transport. They proposed to solve the problem with two different approaches; the first one is based on Benders decomposition for exact solutions for large size instances and the second one is based on a greedy neighborhood search to also solve large size instances.

C. Multicommodity Network Flow Applications

Multicommodity network flows are used widely, particularly in telecommunication and transportation networks. In the literature, one can find two types of networks; the linear multicommodity network flow and the non-linear multicommodity network flow. The first network has linear costs assigned to the flow of each commodity on each edge, while the second network consists of assigning a non-linear objective function to the edge flow. The linear multicommodity network flow problems are challenging because of their sizes, while the non-linear problems are difficult to deal with mainly due to the non-linearity of the objective function.

Many researchers used the multicommodity network flow in different applications. Clarke and Surkis [12] used a multicommodity network flow to model a solution for the school racial desegregation problem. White and Bomberault [45] described an application of the multicommodity network flow analysis in the empty freight car allocation industry. It consists of planning short-range car movements to allocate empty freight cars in railways. Moreover, Rao [41] applied a multicommodity network model to a warehousing problem with cash-liquidity constraints. Furthermore, Zangwill [46] worked with a backlogging and a multi-echelon model of a dynamic economic lot size production system and presented two models; a single product model with backlogging and a multi-echelon model. Both models aim to find a production schedule that minimizes inventory costs as well as the aggregate production. Zangwill [46] used dynamic programming to reach an optimal solution for both models. Bellmore et al. [5] dealt with a multicommodity minimal cost flow problem. Their aim was to achieve a prescribed schedule of deliveries with feasible flows corresponding to a feasible shipping schedule, by considering optimal routing for fuel tankers. Bellmore et al. [5] solved the multicommodity problem using a branch-and-bound enumerative scheme in conjunction with a decomposed linear program with network sub-problems. Geoffrion and Graves [19] described an algorithm based on Benders decomposition to solve the problem of multicommodity distribution system design. Moreover, Christofides and Beasley [11] developed a tree search algorithm and presented two lower bounds for this kind of problem while Beasley [4] developed an algorithm for solving large capacitated warehouse location problems. He presented a lower bound for the capacitated warehouse location problem based upon lagrangian relaxation of a mixed-integer formulation of the problem. Klincewicz and Luss [29] created a Lagrangian relaxation heuristic algorithm for the capacitated facility location problem with single-source constraints and without a branch-and-bound procedure [10]. Meanwhile, Guignard [23] presented an approach to strengthen the separable Lagrangian relaxation of simple plant location using Benders inequalities generated during a Lagrangian dual ascent procedure. Cao and Uebe [10] also dealt with container terminal problems. They presented an algorithm to solve the capacitated multicommodity p-median transportation problem. Their idea is based on re-stacking some old containers to provide the necessary space for the newly arriving containers. Hence, the aim of this algorithm is to minimize the aggregate costs, consisting of searching for and loading a new container, while maintaining some primordial conditions (such as available capacity in a storage row). Cao and Uebe [10] proposed a heuristic based on a branch-and-bound algorithm as well as Lagrangian relaxation combined with subgradient optimization for supplying lower bounds. As mentioned earlier, the non-linear multicommodity networks are challenging because of the non-linear objective function. Goffin et al. [20] proposed a decomposition method to solve such complex problems. This method relies on decomposing the master problem into several subproblems which are formulated as a short path algorithm and solved using Dijkstra's d-heap algorithm. Moreover, McBride and Mamer [34] worked with the undirected multicommodity flow problem that is mostly found while solving traffic-scheduling problems such as railroad and communication networks. They formulated this problem as a piecewise linear optimization problem where large instances were solved within an acceptable time. Larsson and Yuan [30] proposed a fast and convergent lower-bounding procedure that is based on an augmented Lagrangian reformulation of the multicommodity network flow. This algorithm is tested with instances having over 3,600 nodes, 14,000 edges and 80,000 commodities.

In the following chapter, we present a new cost minimization mixed integer problem for the multicommodity hub-and-spoke network design problem.

CHAPTER III

PROBLEM FORMULATION

A. Introduction

In this chapter, we present a mixed integer programming formulation for the multicommodity hub-and-spoke (MCHS) design problem.

B. Assumptions

As highlighted in the literature review, the majority of research is focused on hub-and-spoke models and on multicommodity models, separately. Thus, to benefit from the advantages of both models, we propose the "hub-and-spoke and multicommodity network flow" model. Three main assumptions are considered in our model. The first assumption states that using an inter-hub connection has a lower price per unit than using a spoke connection. Thus, it benefits from a discount factor ($0<\alpha<1$) [18]. With this discount factor α, it is cheaper to take the flow spoke-hub-spoke than spoke-spoke. We also assume that direct connections between spokes are not allowed. Finally, the triangle inequality (the sum of the lengths of any two sides of a given triangle is greater than the length of the third side) holds in the cost structure and costs are proportional to the distance [18].

C. Hub-and-Spoke and Multicommodity Network Flow

1. Parameters

The indices and parameters used in our model are the following:

N: set of all the nodes.

E: set of all the edges.

K: set of all the commodities.

(i, j): all the edges, $(i, j) \in E$.

q_{ij}^{c}: capacity of the cheap link between the nodes (i, j).

q_{ij}^{e}: capacity of the expensive link between the nodes (i, j).

$c_{ij}^{c,k}$: shipping cost on the cheap link between the nodes (i, j) for commodity k.

$c_{ij}^{e,k}$: shipping cost on the expensive link between the nodes (i, j) for commodity k.

f_{ij}^{c}: fixed cost incurred if the cheap link between nodes (i, j) is used in the network.

f_{ij}^{e}: fixed cost incurred if the expensive link between nodes (i, j) is used in the network.

h_i: fixed cost for establishing a hub facility at node i.

w_k: amount shipped from an origin node O to a destination node D.

2. Decision Variables

$u_{ij}^{c,k}$: amount of flow on edge (i, j) through the cheap link for commodity k.

$u_{ij}^{e,k}$: amount of flow on edge (i, j) through the expensive link for commodity k.

$$y_{ij}^{c} = \begin{cases} 1 & \text{if the cheap link is used between nodes } (i, j). \\ 0 & \text{otherwise.} \end{cases}$$

$$y_{ij}^{e} = \begin{cases} 1 & \text{if the expensive link is used between nodes } (i, j). \\ 0 & \text{otherwise.} \end{cases}$$

$$z_i = \begin{cases} 1 & \text{if node } i \text{ is a hub.} \\ 0 & \text{otherwise.} \end{cases}$$

3. Problem Formulation

The problem formulation is as follows:

$$\min \sum_{k\in K}\sum_{(i,j)\in E}(c_{ij}^{c,k}u_{ij}^{c,k}+c_{ij}^{e,k}u_{ij}^{e,k})+\sum_{(i,j)\in E}(f_{i,j}^{c}y_{i,j}^{c}+f_{i,j}^{e}y_{i,j}^{e})+\sum_{i\in N}h_i z_i \tag{1}$$

$$\text{s.t.} \sum_{j\in N^+}(u_{ij}^{c,k}+u_{ij}^{e,k})-\sum_{j\in N^-}(u_{ij}^{c,k}+u_{ij}^{e,k}) = \begin{cases} w^k & \text{if } i=O(k) \\ -w^k & \text{if } i=D(k) \\ 0 & \text{otherwise} \end{cases} \quad \forall i\in N,\ \forall k\in K, \tag{2}$$

$$\sum_{k\in K}u_{ij}^{c,k}\le q_{ij}^{c}y_{ij}^{c},\ \forall(i,j)\in E \tag{3}$$

$$\sum_{k\in K}u_{ij}^{e,k}\le q_{ij}^{e}y_{ij}^{e},\ \forall(i,j)\in E \tag{4}$$

$$y_{ij}^{c}\le\frac{z_i+z_j}{2},\ \forall(i,j)\in E \tag{5}$$

$$y_{ij}^{c}+y_{ij}^{e}\le(z_i+z_j),\ \forall(i,j)\in E \tag{6}$$

$$\sum_{k\in K}\sum_{i\in N}(u_{ij}^{c,k}+u_{ij}^{e,k})\le C_j,\quad \forall j\in N \tag{7}$$

$$\sum_{i\in N}(u_{ij}^{c,k}+u_{ij}^{e,k})-w_{kj}\le Mz_j,\quad \forall j\in N,\forall k\in K \tag{8}$$

$$u_{ij}^{c,k},u_{ij}^{e,k}\ge 0,\ \forall(i,j)\in E \tag{9}$$

$$y_{ij}^{c},y_{ij}^{e}\in\{0,1\},\ \forall(i,j)\in E \tag{10}$$

$$z_i\in\{0,1\},\ \forall i\in N \tag{11}$$

The objective function minimizes the total cost of the network. It has three components.

- The first component $\sum_{k\in K}\sum_{(i,j)\in E}(c_{ij}^{c,k}u_{ij}^{c,k}+c_{ij}^{e,k}u_{ij}^{e,k})$ represents the cost of shipping through both links, the cheap and expensive ones.

- The second component $\sum_{(i,j)\in E}(f_{i,j}^{c}y_{i,j}^{c}+f_{i,j}^{e}y_{i,j}^{e})$ is the fixed cost incurred if using the links on the network.

- The final component $\sum_{i\in N}h_i z_i$ stands for the fixed cost for establishing a hub facility.

The following constraints are imposed:

- Constraint (2) deals with the conservation of flows between the origin node and the destination node. If the amount of flow w_k is positive, then the flow is entering the node. However, if the amount of flow w_k is negative, then the flow is leaving the node.

- Constraints (3) and (4) ensure that the capacities of the cheap and expensive links, respectively, are not exceeded.

- Constraint (5) allows the cheap link to be used between two nodes, if and only if, both nodes are hubs.

- Constraint (6) ensures that to be able to send flow from one node to the other, one of them should be a hub.

- Constraint (7) makes sure that the total flow through the cheap and expensive links into hub j does not exceed the hub capacity C_j.

- Constraint (8) forces the assumption that in-transit flow is not allowed through spoke nodes.

- Constraint (9) guarantees the positivity of the variables u while constraints (10) and (11) are binary conditions on variables y and z, respectively.

CHAPTER IV

CONGESTION

A. Congestion Effects

Congestion is a very serious issue that delays daily activities. Traffic congestion is the most common form of congestion that occurs in our daily lives. The effects that traffic congestion has on businesses today is noticeable in more than just travel time and travel costs. They affect the size of the business market, vehicle and personal deployment, business relocation, delivery scheduling and intermodal connections. Longer hours spent on the road affect the retailers and distributors, the latter being obliged to keep a high amount of inventory in their warehouses in order to avoid lost sales. Furthermore, congestion delays can notably reduce the ability to manage the flow and inventory required in businesses that have rapid inventory turnover. This results in potential lost sales in addition to the costs of managing inventory and receiving emergency, unplanned, or late night/early morning shipments. The effect of congestion is clearly visible on intermodal connections such as train stations, airports and marine ports. In these stations, traffic congestion isn't the only problem; other types such as loading/unloading congestion are also problematic. First off, in traffic congestion, the increased traffic is causing the trains to miss their scheduled time of arrival to a station which counts as wasted operation cost and causes companies to reschedule their operations. In order to avoid this issue in train stations, some companies use a broadened schedule time, but this doesn't solve the extra cost generated by vehicles and drivers. As for airports, firms are considering diverting their air freights to more distant ones which are less crowded. For their part, some marine ports are trying to extend their working hours, and providing reservation systems to even out arrivals. On the other hand, due to the limited size of the intermodal stations and number of cranes available, delays happen in loading and unloading materials.

In a survey conducted by "The Associated General Contractors of America" [36] studying 1,200 construction contractors, 53% said that traffic congestion caused frequent to significant impact on their operations. Up to 40% said that it caused occasional impacts, while 64% reported that congestion is causing a loss of at least one day of productivity per worker annually.

Therefore, congestion, having numerous effects on production and productivity, is an extremely serious problem that must be handled carefully.

B. Congestion in the Literature

As stated earlier, hubs are the nodes that get affected by congestion [33]. Grove and O'Kelly [22] were the first to introduce the congestion problem in a hub-and-spoke network. They showed how the amount of flow at hubs leads to scheduled delays for airlines. Kara and Tansel [25] shed light on the scheduled delays at hub airports. The aim of their paper was to recognize the major factors affecting congestion in such networks. One method to avoid congestion was introduced by Aykin [2] with his capacitated hub location mode. He proposed to add capacity constraints when designing the network which would regulate congestion. However, the method of Aykin [2] was not too efficient since it did not show the concession between hub capacity and hub flow. In the research of Serra et al. [42], the hub-and-spoke network was modeled as an M/D/c queuing network, where the capacity constraints are set based on the probability of the waiting customers. In addition, Elhedhli and Hu [15] were the first to deal with congestion cost by introducing it in the objective function of the designed network.
In this thesis, we adopt an identical approach to Elhedhli and Hu [15] to model congestion in the proposed multicommodity hub-and-spoke network design problem.

C. Modeling Congestion

1. Congestion Function

The objective (1) is to minimize the transportation costs while ignoring all other costs. Hence, in such a model, cheaper hubs are overused while the more expensive ones are neglected. This increase in the demand flow of some hubs leads to a dramatic congestion which causes a significant increase in operational costs. Elhedhli and Hu [15] presented the economical effect of congestion as a power-law function of the hub flow u and as the difference between the capacity of the hub node and the flow at that hub node. It is represented by the fractional function $g(u)$:

$$g(u) = \frac{u}{C - u} \tag{12}$$

Where C is the capacity of the node.

Figure IV.1 represents a plot of the congestion rate at one node, denoted $g(u)$, versus the flow u at that node and a determined capacity C of 100.

2. Congestion Costs

Considering a cost of congestion T_0, the total congestion cost $g(u_j)$ for all hubs j is the following:

$$T_0 \sum_{j \in N} \left(\frac{\sum_{k \in K} \sum_{i \in N} u_{ij}^{k,c} + u_{ij}^{k,e}}{C_j - \sum_{k \in K} \sum_{i \in N} u_{ij}^{k,c} + u_{ij}^{k,e}} \right) \tag{13}$$

Where C_j represents the maximum capacity allowed at hub j.

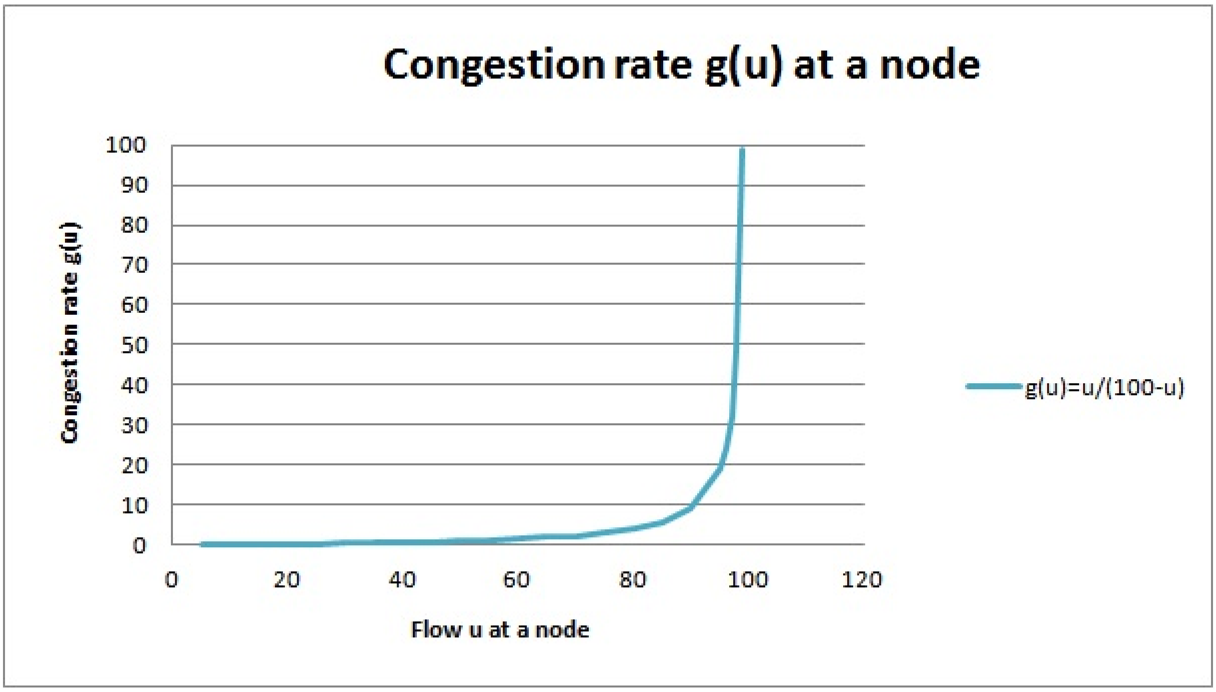

Figure IV.1: Congestion flow rate graph

Similar to Elhedhli and Hu [15], we propose adding the above equation to the objective function (1) to include the cost of congestion in the proposed model.

The new objective function will become:

$$\min \sum_{k\in K}\sum_{(i,j)\in E}(c_{ij}^{c,k}u_{ij}^{c,k}+c_{ij}^{e,k}u_{ij}^{e,k})+\sum_{(i,j)\in E}(f_{i,j}^{c}y_{i,j}^{c}+f_{i,j}^{e}y_{i,j}^{e})+\sum_{i\in N}h_i z_i$$
$$+T_0\sum_{j\in N}\left(\frac{\sum_{k\in K}\sum_{i\in N}u_{ij}^{k,c}+u_{ij}^{k,e}}{C_j-\sum_{k\in K}\sum_{i\in N}u_{ij}^{k,c}+u_{ij}^{k,e}}\right) \tag{14}$$

The new objective function is not linear. Hence next we propose an approach to solve the resulting non-linear problem and finally evaluate the impact of congestion on hub-and-spoke network design.

D. Linearization of Congestion

As stated earlier, the new objective function is non-linear due to the added congestion term. However, solving non-linear problems is typically difficult in optimization solvers. Hence, the typical approach to tackle such problems relies on linearization techniques.
In this section, we will discuss how to transform a non-linear problem into a linear one using a piecewise linear approximation approach.

The non-linear congestion term is the following:

$$f(u) = \frac{\sum_{k\in K}\sum_{i\in N}(u_{ij}^{c,k} + u_{ij}^{e,k})}{C_j - \sum_{k\in K}\sum_{i\in N}(u_{ij}^{c,k} + u_{ij}^{e,k})} \quad \forall j \in N \tag{15}$$

As shown in Figure IV.2, the continuous and concave function $f(u)$ can be approximated arbitrarily closely with a finite set of piecewise linear functions that are tangent to $f(u)$. It can be represented by the following equation:

$$f(u) \approx \max_{\forall \overline{u}\in R}(f'(u)(u-\overline{u}) + f(\overline{u})) \tag{16}$$

Hence, taking equation (16) as a reference, our congestion function (15) for every node j can be approximated by the following equation:

$$\begin{aligned} f_j(u) \approx \max_{\forall \overline{u}\in R}(&\sum_{i\in N} \frac{C_j}{(C_j - \sum_{k\in K}(\overline{u}_{ij}^{c,k} + \overline{u}_{ij}^{e,k}))^2}(\sum_{k\in K}(u_{ij}^{c,k} + u_{ij}^{e,k}) - \sum_{k\in K}(\overline{u}_{ij}^{c,k} + \overline{u}_{ij}^{e,k})) \\ &+ \sum_{i\in N} \frac{(\overline{u}_{ij}^{c,k} + \overline{u}_{ij}^{e,k})}{C_j - \sum_{k\in K}(\overline{u}_{ij}^{c,k} + \overline{u}_{ij}^{e,k})}) \end{aligned} \tag{17}$$

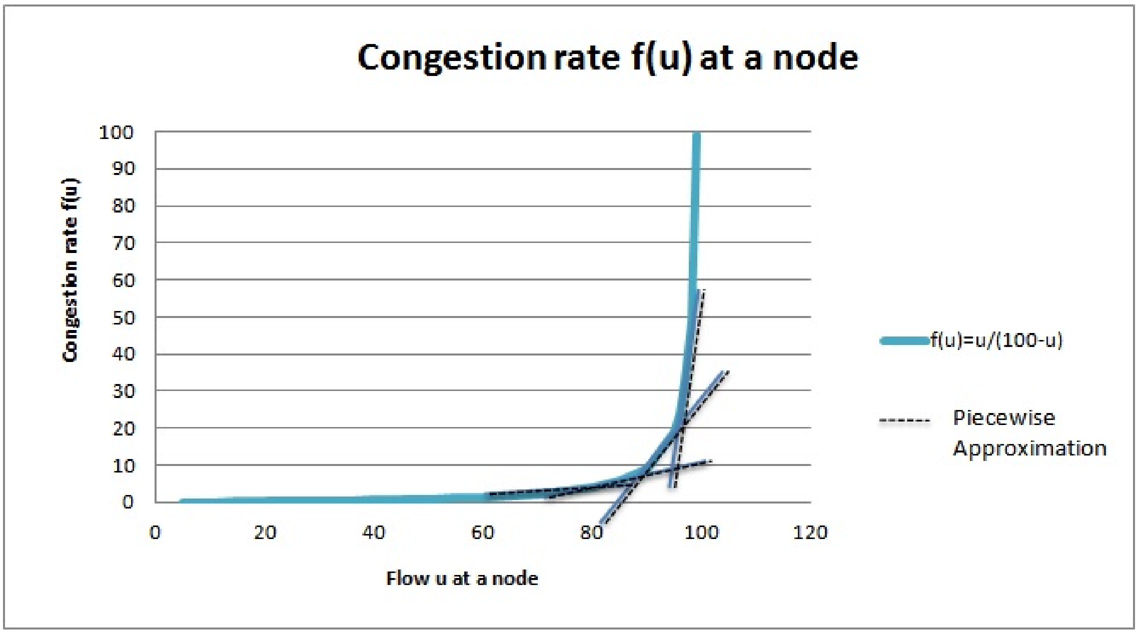

Figure IV.2: Linear Approximation of an example

E. Model Formulation

Adding the congestion costs to problem (1)-(11), the resulting model is as follows:

$$\min \sum_{k\in K}\sum_{(i,j)\in E}(c_{ij}^{c,k}u_{ij}^{c,k}+c_{ij}^{e,k}u_{ij}^{e,k})+\sum_{(i,j)\in E}(f_{i,j}^{c}y_{i,j}^{c}+f_{i,j}^{e}y_{i,j}^{e})+\sum_{i\in N}h_i z_i+T_0\sum_{j\in N}\theta_j \tag{18}$$

$$\text{s.t.} \sum_{j\in N^+}(u_{ij}^{c,k}+u_{ij}^{e,k})-\sum_{j\in N^-}(u_{ij}^{c,k}+u_{ij}^{e,k})=\begin{cases} w^k & \text{if } i=O(k)\\ -w^k & \text{if } i=D(k)\\ 0 & \text{otherwise}\end{cases} \quad \forall i\in N,\ \forall k\in K, \tag{19}$$

$$\sum_{k\in K}u_{ij}^{c,k}\le q_{ij}^{c}y_{ij}^{c},\ \forall(i,j)\in E \tag{20}$$

$$\sum_{k\in K}u_{ij}^{e,k}\le q_{ij}^{e}y_{ij}^{e},\ \forall(i,j)\in E \tag{21}$$

$$y_{ij}^{c}\le\frac{z_i+z_j}{2},\ \forall(i,j)\in E \tag{22}$$

$$y_{ij}^{c}+y_{ij}^{e}\le(z_i+z_j),\ \forall(i,j)\in E \tag{23}$$

$$\begin{aligned}\theta_j\approx\max_{\forall\bar{u}\in R}\Big(&\sum_{i\in N}\frac{C_j}{(C_j-\sum_{k\in K}(\bar{u}_{ij}^{c,k}+\bar{u}_{ij}^{e,k}))^2}\sum_{k\in K}(u_{ij}^{c,k}+u_{ij}^{e,k})+\sum_{i\in N}\frac{(\bar{u}_{ij}^{c,k}+\bar{u}_{ij}^{e,k})}{C_j-\sum_{k\in K}(\bar{u}_{ij}^{c,k}+\bar{u}_{ij}^{e,k})}\\ &-\sum_{i\in N}\frac{C_j}{(C_j-\sum_{k\in K}(\bar{u}_{ij}^{c,k}+\bar{u}_{ij}^{e,k}))^2}\sum_{k\in K}(\bar{u}_{ij}^{c,k}+\bar{u}_{ij}^{e,k})\Big),\ \forall j\in N\end{aligned} \tag{24}$$

$$\sum_{k\in K}\sum_{i\in N}(u_{ij}^{c,k}+u_{ij}^{e,k})\le C_j,\ \forall j\in N \tag{25}$$

$$\sum_{i\in N}(u_{ij}^{c,k}+u_{ij}^{e,k})-w_{kj}\le Mz_j,\ \forall j\in N,\forall k\in K \tag{26}$$

$$u_{ij}^{c,k},u_{ij}^{e,k}\ge 0,\ \forall(i,j)\in E \tag{27}$$

$$y_{ij}^{c},y_{ij}^{e}\in\{0,1\},\ \forall(i,j)\in E \tag{28}$$

$$z_i\in\{0,1\},\ \forall i\in N \tag{29}$$

$$\theta_j\ge 0,\ \forall j\in N \tag{30}$$

Since the objective function is a minimization, constraint (24) can be replaced by an inequality, of the form

$$\theta_j \geq \sum_{i \in N} \frac{C_j}{(C_j - \sum_{k \in K}(\overline{u}_{ij}^{c,k} + \overline{u}_{ij}^{e,k}))^2} \sum_{k \in K}(u_{ij}^{c,k} + u_{ij}^{e,k}) + \sum_{i \in N} \frac{(\overline{u}_{ij}^{c,k} + \overline{u}_{ij}^{e,k})}{C_j - \sum_{k \in K}(\overline{u}_{ij}^{c,k} + \overline{u}_{ij}^{e,k})}$$
$$- \sum_{i \in N} \frac{C_j}{(C_j - \sum_{k \in K}(\overline{u}_{ij}^{c,k} + \overline{u}_{ij}^{e,k}))^2} \sum_{k \in K}(\overline{u}_{ij}^{c,k} + \overline{u}_{ij}^{e,k}), \ \forall \overline{u} \in R, \forall j \in N \tag{31}$$

Hence, the resulting model is:

$$\min \sum_{k\in K}\sum_{(i,j)\in E}(c_{ij}^{c,k}u_{ij}^{c,k}+c_{ij}^{e,k}u_{ij}^{e,k})+\sum_{(i,j)\in E}(f_{i,j}^{c}y_{i,j}^{c}+f_{i,j}^{e}y_{i,j}^{e})+\sum_{i\in N}h_i z_i+T_0\sum_j \theta_j \tag{32}$$

$$\text{s.t.} \sum_{j\in N^+}(u_{ij}^{c,k}+u_{ij}^{e,k})-\sum_{j\in N^-}(u_{ij}^{c,k}+u_{ij}^{e,k})=\begin{cases} w^k & \text{if } i=O(k)\\ -w^k & \text{if } i=D(k)\\ 0 & \text{otherwise}\end{cases} \quad \forall i\in N,\ \forall k\in K, \tag{33}$$

$$\sum_{k\in K}u_{ij}^{c,k}\le q_{ij}^{c}y_{ij}^{c},\ \forall(i,j)\in E \tag{34}$$

$$\sum_{k\in K}u_{ij}^{e,k}\le q_{ij}^{e}y_{ij}^{e},\ \forall(i,j)\in E \tag{35}$$

$$y_{ij}^{c}\le\frac{z_i+z_j}{2},\ \forall(i,j)\in E \tag{36}$$

$$y_{ij}^{c}+y_{ij}^{e}\le(z_i+z_j),\ \forall(i,j)\in E \tag{37}$$

$$\theta_j\ge\sum_{i\in N}\frac{C_j}{(C_j-\sum_{k\in K}(\overline{u}_{ij}^{c,k}+\overline{u}_{ij}^{e,k}))^2}\sum_{k\in K}(u_{ij}^{c,k}+u_{ij}^{e,k})+\sum_{i\in N}\frac{(\overline{u}_{ij}^{c,k}+\overline{u}_{ij}^{e,k})}{C_j-\sum_{k\in K}(\overline{u}_{ij}^{c,k}+\overline{u}_{ij}^{e,k})}$$
$$-\sum_{i\in N}\frac{C_j}{(C_j-\sum_{k\in K}(\overline{u}_{ij}^{c,k}+\overline{u}_{ij}^{e,k}))^2}\sum_{k\in K}(\overline{u}_{ij}^{c,k}+\overline{u}_{ij}^{e,k}),\ \forall\overline{u}\in R,\forall j\in N \tag{38}$$

$$\sum_{k\in K}\sum_{i\in N}(u_{ij}^{c,k}+u_{ij}^{e,k})\le C_j,\ \forall j\in N \tag{39}$$

$$\sum_{i\in N}(u_{ij}^{c,k}+u_{ij}^{e,k})-w_{kj}\le Mz_j,\ \forall j\in N,\forall k\in K \tag{40}$$

$$u_{ij}^{c,k},u_{ij}^{e,k}\ge 0,\ \forall(i,j)\in E \tag{41}$$

$$y_{ij}^{c},y_{ij}^{e}\in\{0,1\},\ \forall(i,j)\in E \tag{42}$$

$$z_i\in\{0,1\},\ \forall i\in N \tag{43}$$

$$\theta_j\ge 0,\ \forall j\in N \tag{44}$$

F. Cutting Plane Approach

The set of points $\overline{u} \in R$, at which $f(u)$ is approximated, is not known beforehand. However, they are generated with a cutting plane approach. The cutting plane approach aims at solving the problem iteratively, where each iteration generates a solution that is not necessarily optimal.

Since the set R is not known, this approach starts with an empty set $\emptyset$, where constraint (38) is ignored. Thus, the optimal solution for the relaxed problem, which is a minimization, results in a lower bound $[LB]$. Given this solution, the actual congestion is computed according to equation (14) and hence a feasible solution is then obtained which provides an upper bound $[UB]$. Furthermore, given the optimal solution of the relaxation, a tangential approximation is obtained and a new constraint of the form (31) is added to the problem which is then resolved. Optimality is reached once the $[LB]$ and the $[UB]$ are within a desired gap ε. Here, the two bounds act as a stopping criteria for this iterative algorithm.

CHAPTER V

COMPUTATIONAL RESULTS

In this chapter, we propose three different experiments to illustrate the performance of our proposed model.

A. Small networks to illustrate the difference in optimal routes when congestion is included.

B. Randomly generated instances by the *Mulgen* generator [14] to illustrate the computational performance of our model.

C. A real network based on an example inspired by the French railroad map.

CPLEX 12 is used as an optimization solver with default settings and a CPU time limit of 10,000 seconds is set.

A. Sample Networks

In this section, we present two sample examples. The first network consists of 10 nodes, 20 edges and 2 commodities. Figure V.1 shows the network topology in addition to the flow of the two commodities. Figure V.2 represents the optimal routing throughout the 10 nodes while taking into account the congestion costs in the network. However, Figure V.3 represents the optimal routing through the 10 nodes but without taking into account the congestion costs (congestion costs = 0). In Figure V.2, nodes 1, 3, 4, and 8 are chosen as the hubs while nodes 2, 5, 7, and 10 are the spokes. However, in Figure V.3, nodes 3, 4, and 8 are chosen as hubs, while nodes 1, 2, 5, 7, and 10 are the spokes.

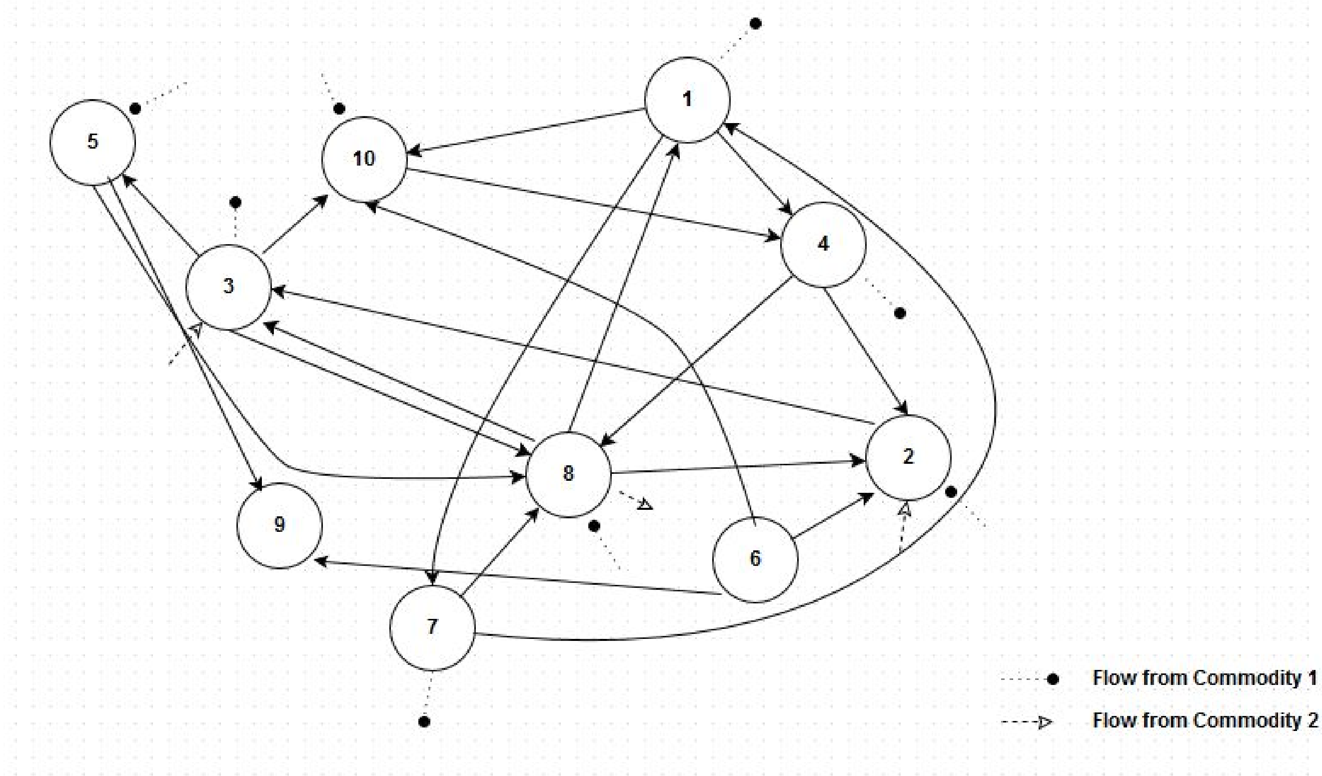

Figure V.1: Network of 10 nodes, 20 edges, 2 commodities

Table 2 as well as Figure V.4 show the congestion on the 10 nodes for both figures. When congestion costs are taken into account, nodes are less congested. For example, as shown in Table 2, node 3 had a congestion of 1.049. When congestion costs were added to the objective function, congestion on node 3 dropped to 0.63. The same effect is observed on node 4 where it dropped from 0.497 to 0.26 and on node 8 where it dropped from 1.109 to 0.67.

Figure V.5 represents a larger network with 10 nodes, 46 edges and 3 commodities. Figure V.6 shows the optimal routing solution for this network with congestion costs. Nodes 1, 2, 6, 7, 8, and 10 represent the hubs for the network while nodes 3, 4, 5, and 9 are the spokes. On the other hand, Figure V.7 represents the optimal route when congestion costs are 0. Nodes 1, 6, 7, 8, 9, and 10 are the hubs whereas nodes 3, 4, and 5 are the spokes for this network.

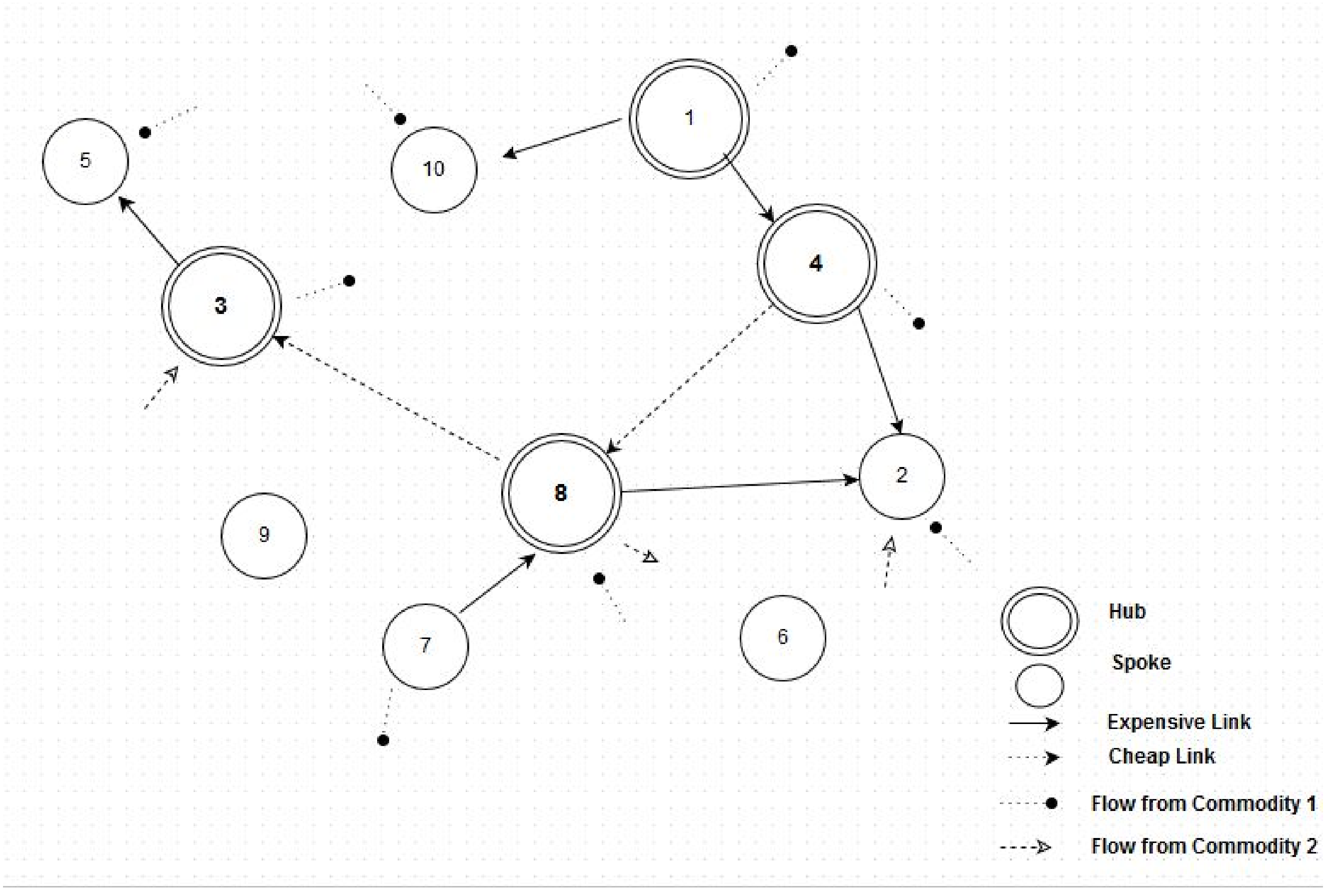

Figure V.2: Solution of the network in Figure V.1 with congestion costs

Table 3 and Figure V.8 illustrate the difference between congestion on all nodes with and without congestion costs taken into account. As mentioned earlier, congestion costs affect nodes by decreasing their congestion. For example, congestion on node 6 dropped from 0.845 to 0.58, on node 10 it dropped from 3.25 to 3.10 and on node 7 from 0.966 to 0.952.

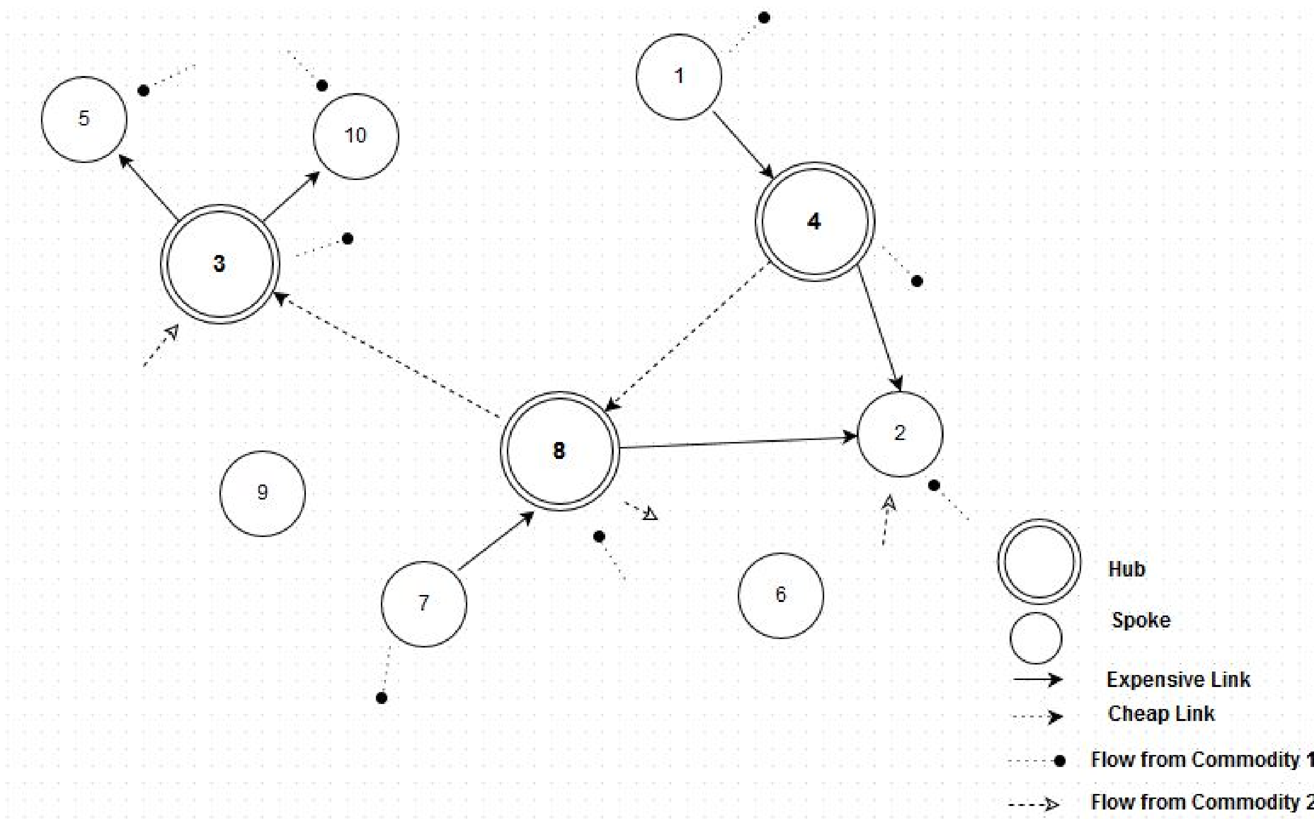

Figure V.3: Solution of the network in Figure V.1 without congestion costs

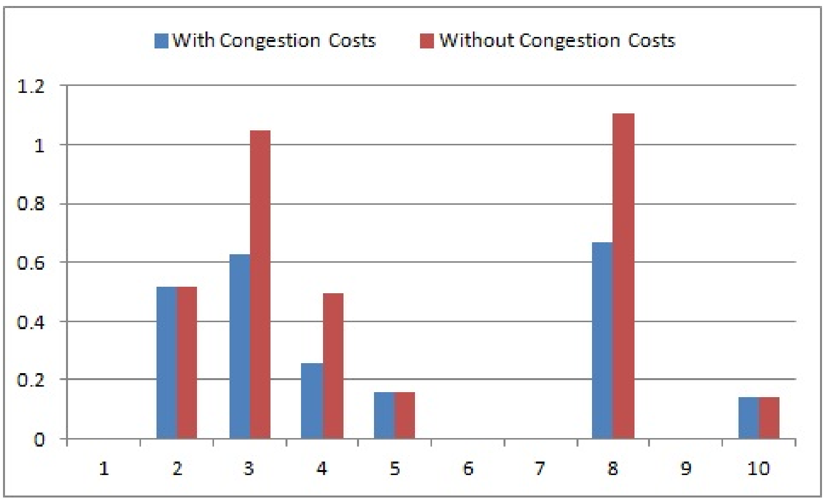

Figure V.4: Comparison of congested nodes for the 10 nodes, 20 edges and 2 commodities network

Table 2: Comparison of congested nodes for the 10 nodes, 20 edges and 2 commodities network

Nodes	With Congestion Costs	Without Congestion Costs
1	0	0
2	0.52	0.52
3	0.63	1.049
4	0.26	0.497
5	0.16	0.16
6	0	0
7	0	0
8	0.67	1.109
9	0	0
10	0.14	0.142

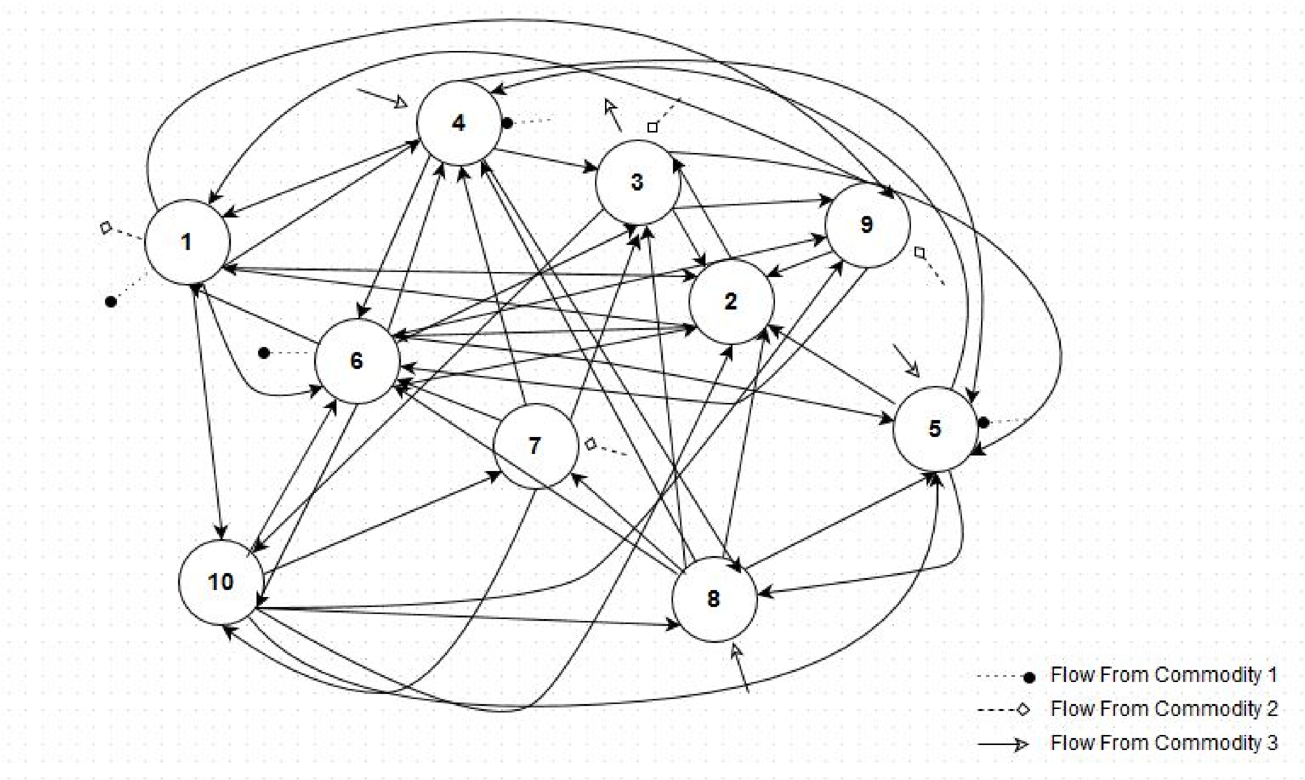

Figure V.5: Network of 10 nodes, 46 edges, 3 commodities

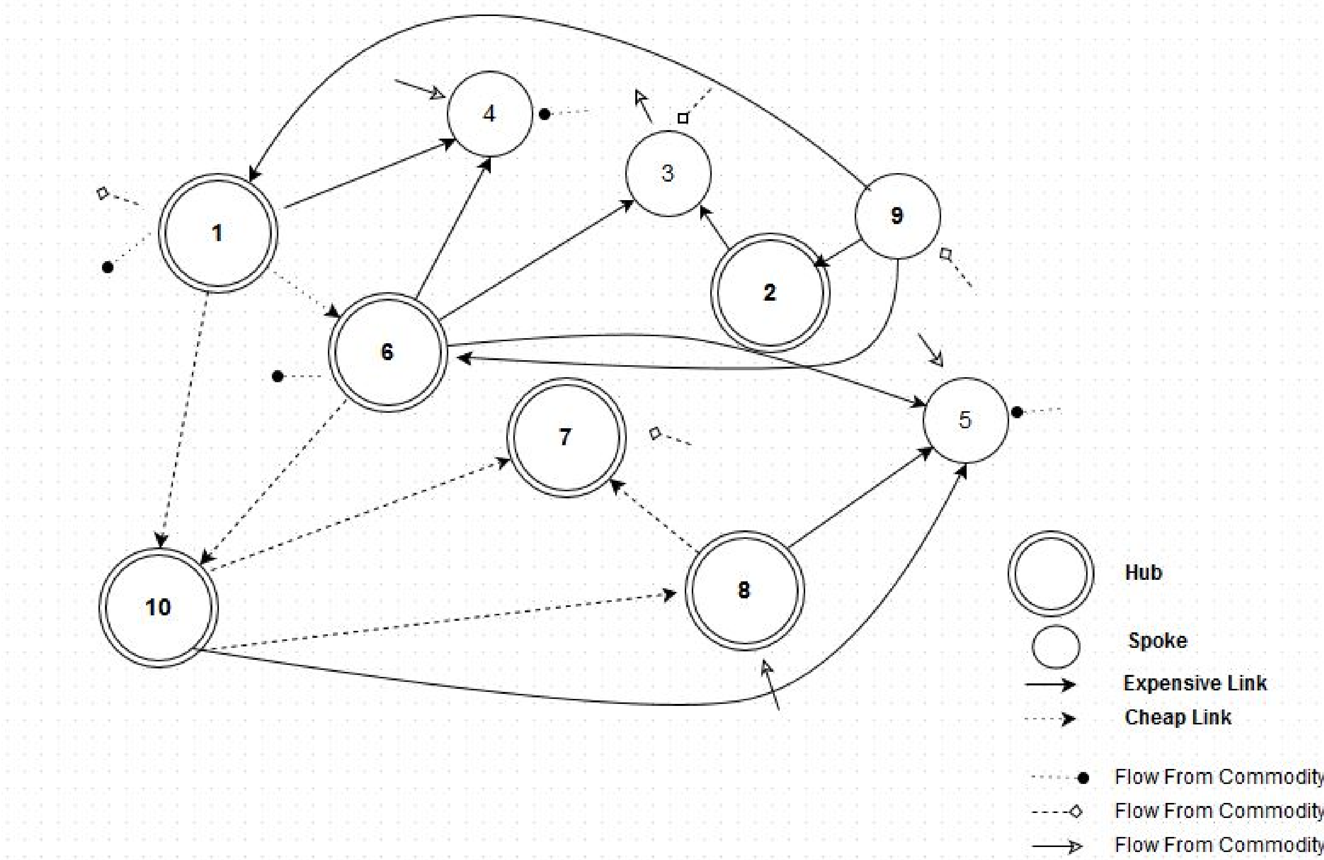

Figure V.6: Solution of the network in Figure V.5 with congestion costs

Table 3: Comparison of congested nodes for the 10 nodes, 46 edges and 3 commodities network

Nodes	With Congestion Costs	Without Congestion Costs
1	0.149	0.149
2	0.096	0
3	0.206	0.206
4	0.213	0.213
5	0.492	0.49
6	0.58	0.845
7	0.952	0.966
8	0.402	0.407
9	0	0.007
10	3.10	3.25

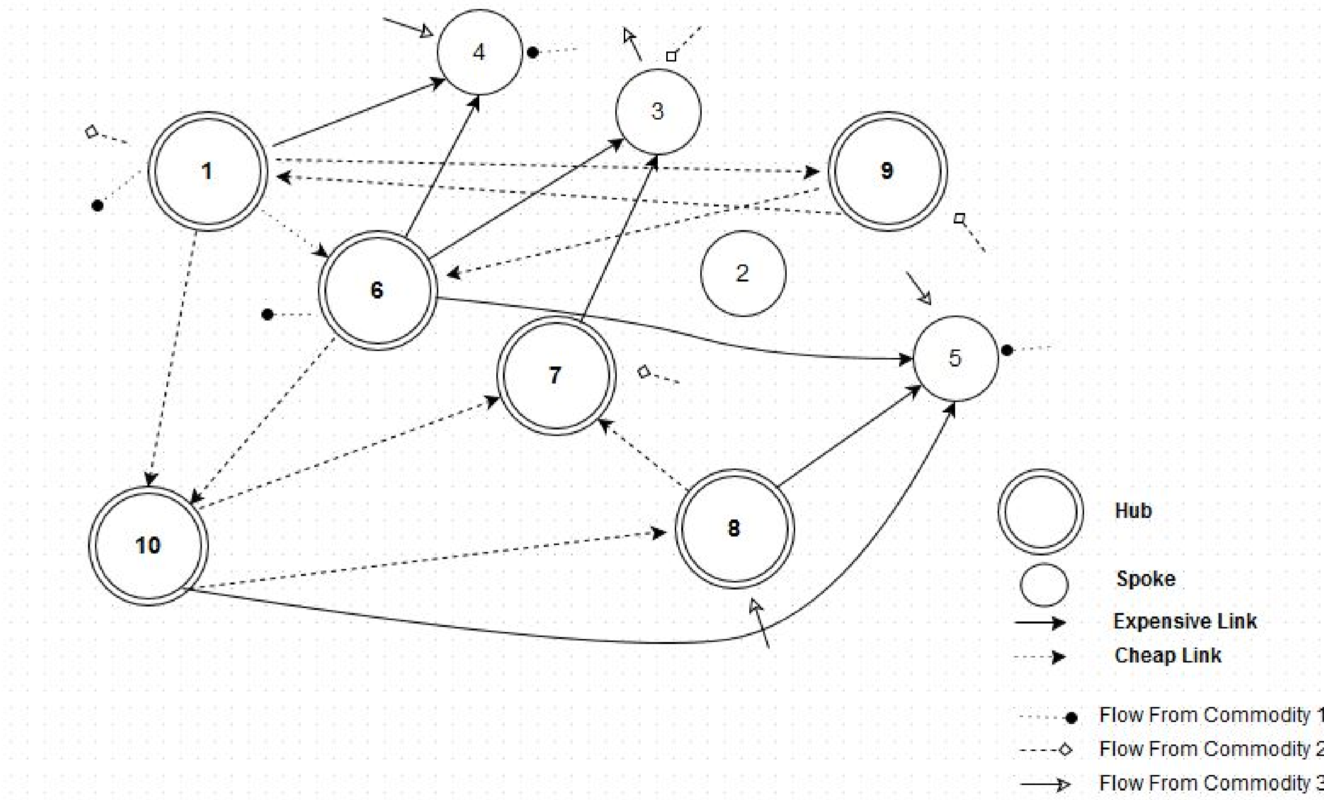

Figure V.7: Solution of the network of Figure V.5 without congestion costs

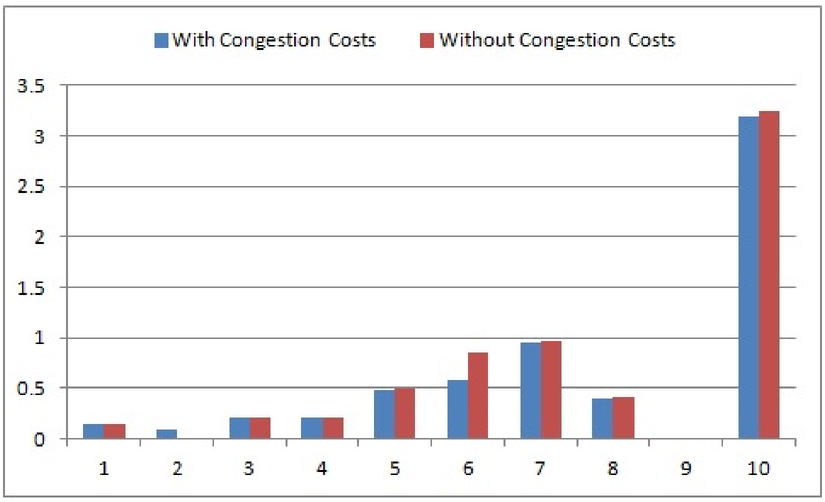

Figure V.8: Comparison of congested nodes for the 10 nodes, 46 edges and 3 commodities network

B. Computational Performance

We tested the proposed MCHS network using instances that are generated randomly with the *Mulgen* generator [14]. Several sets of instances were created with varying numbers of nodes (30, 40, and 50), edge density (10% to 90%), and commodities (3, 5, 7, and 10). The number of edges for our directed graph is as follows:

$$E = N * (N-1) \tag{45}$$

Where E is the number of edges and N is the number of nodes.

We note that for each network size, 5 instances were generated randomly and average results were reported.

Table 4 represents the results of all the instances generated with 30 nodes and varying % density and commodities. It shows the difference between taking congestion costs into account in the objective function and disregarding them. It is important to notice that the % gap is not zero for all the instances which means that not all of them reached optimality within our time limit (10,000 seconds).

Table 4 states that the CPU time increases as the number of commodities increases. In addition, as shown in Table 4 and Figures V.9, V.10, V.11, and V.12, solving networks with 30 nodes while taking congestion costs into account demands more computational time than disregarding these costs. For example, for 30 nodes, 90% density and 7 commodities, the % gap differs from 0% to 7.87% while the average CPU time differs from 6,261.9 seconds to 10,000 seconds.

Table 5 illustrates the change in CPU time and % gap, as the % density and the number of commodities vary for 40-node networks, when including or disregarding congestion costs. Similarly, Table 5 and Figures V.13, V.14, V.15, and V.16 show that it is harder to solve networks when congestion costs are included in the objective function. When they

Table 4: Results of networks with 30 nodes, varying % density and commodities

Nodes	% density	Commodities	% gap without Congestion Costs	% gap with Congestion Costs	Avg CPU time without Congestion Costs	Avg CPU time with Congestion Costs
30	20%	3	0	0	18.02	32.52
30	30%	3	0	0	64.28	132.25
30	40%	3	0	0	123.86	170.82
30	50%	3	0	0	139.64	248.88
30	60%	3	0	0	164.99	207.09
30	70%	3	0	0	97.76	282.85
30	80%	3	0	0	252.04	368.07
30	90%	3	0	0	173.18	483.55
30	20%	5	0	0	78.01	145.72
30	30%	5	0	0	141.77	245.55
30	40%	5	0	0	575.75	1066.55
30	50%	5	0	0	1115.58	1686.44
30	60%	5	0	0	1259.69	2787.48
30	70%	5	0	0	698.02	1027.92
30	80%	5	0	0	935.92	1802.11
30	90%	5	0	0	1449.87	3413.53
30	20%	7	0	0	241.39	418.71
30	30%	7	0	0	806.35	1714.26
30	40%	7	0	0	2163.14	4592.29
30	50%	7	0	0	4462.19	8980.27
30	60%	7	0	0	5124.83	9399.67
30	70%	7	0	0	4849.37	8782.68
30	80%	7	0	0	3240.47	6613.76
30	90%	7	0	7	6261.90	10000.3
30	20%	10	0	0	1780.47	3753.36
30	30%	10	0	6	7025.61	10000.3
30	40%	10	0	0	8470.75	9857.89
30	50%	10	0	12	8158.02	10000.3
30	60%	10	15	18	10000.1	10000.4
30	70%	10	17	22	10000	10000.2
30	80%	10	21	26	10000.1	10000.7
30	90%	10	0	24	8755.92	10000.4

are included, networks with 10 commodities and 20%-90% density do not reach optimality within the time limits of 10,000 seconds (Table 5). However, when these costs are disregarded, networks with 10 commodities and 30%-90% density do not reach optimality.This is also illustrated in Figure V.16 where the curves reach saturation.

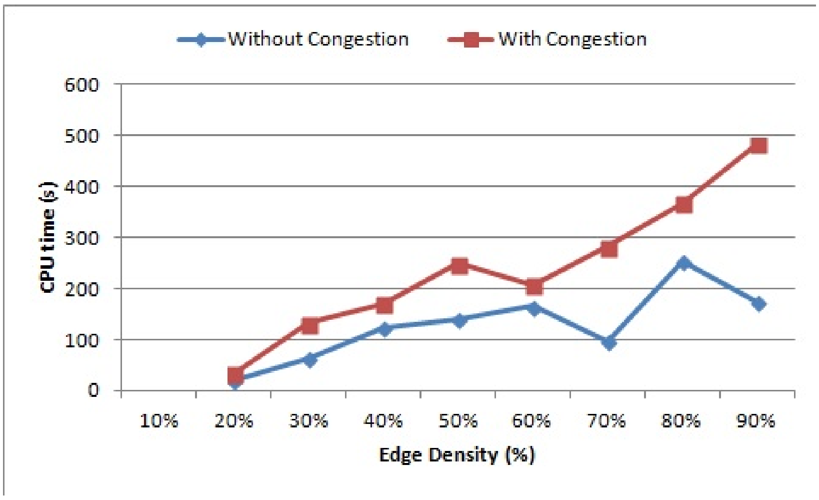

Figure V.9: CPU time vs. Edge Density for 30 nodes and 3 commodities networks

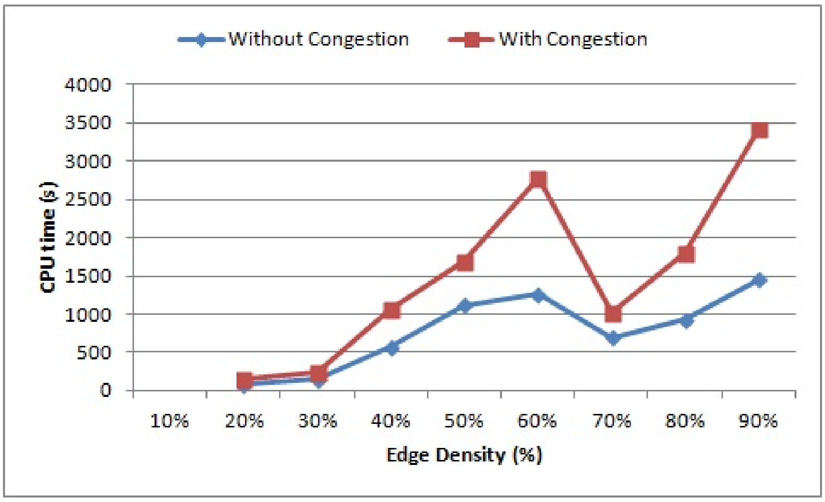

Figure V.10: CPU time vs. Edge Density for 30 nodes and 5 commodities networks

Table 6 shows the relationship between congestion costs and CPU time for 50-node networks with varying edge % density and commodities. It is important to note first that for 50-node networks, 5 commodities is the maximum number of commodities reaching optimality within our time limit. Table 6 as well as Figures V.17 and V.18 prove that dealing with congestion costs needs more computational time. Figure V.18 shows that the curve where congestion costs are included reaches saturation from 30 to 80% density.

The computational results show that the Multicommodity Hub-and-Spoke network design is computationally very challenging even for the state of the optimization solver.

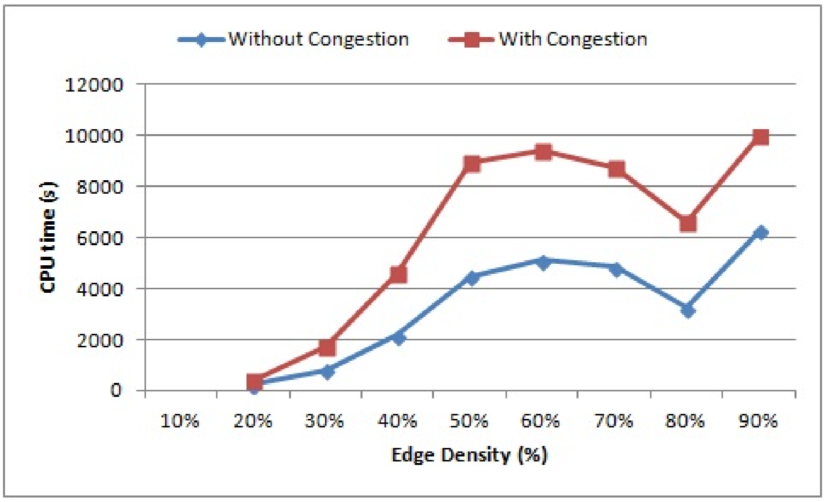

Figure V.11: CPU time vs. Edge Density for 30 nodes and 7 commodities networks

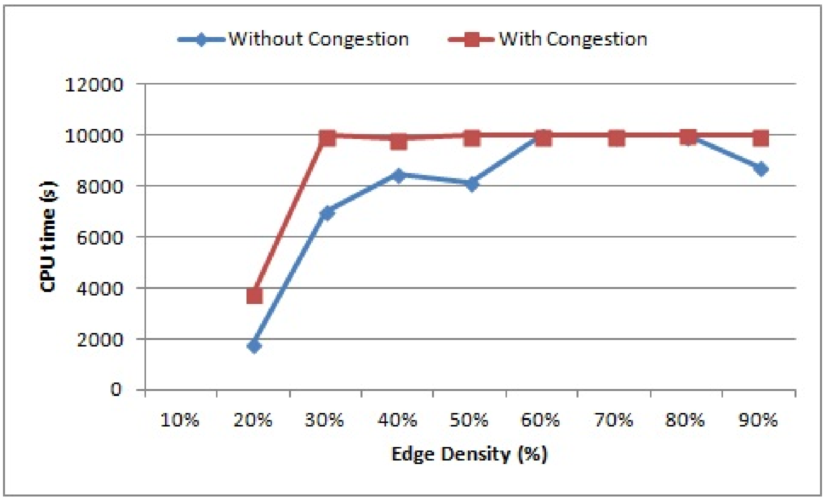

Figure V.12: CPU time vs. Edge Density for 30 nodes and 10 commodities networks

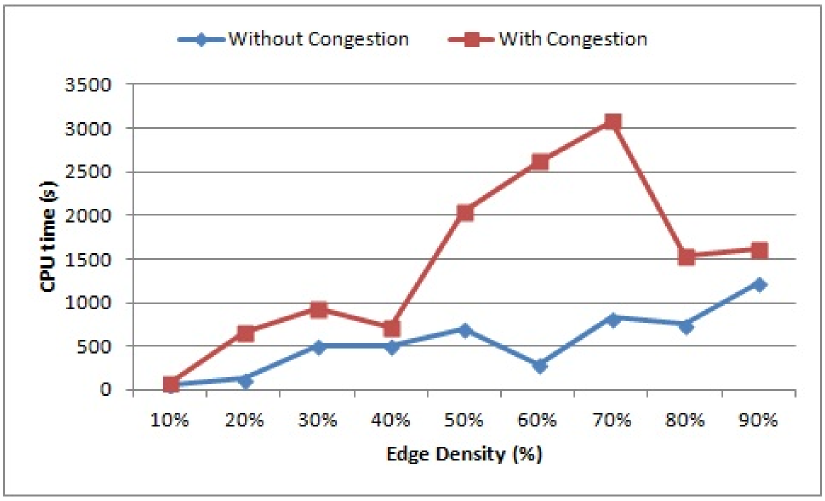

Figure V.13: CPU time vs. Edge Density for 40 nodes and 3 commodities networks

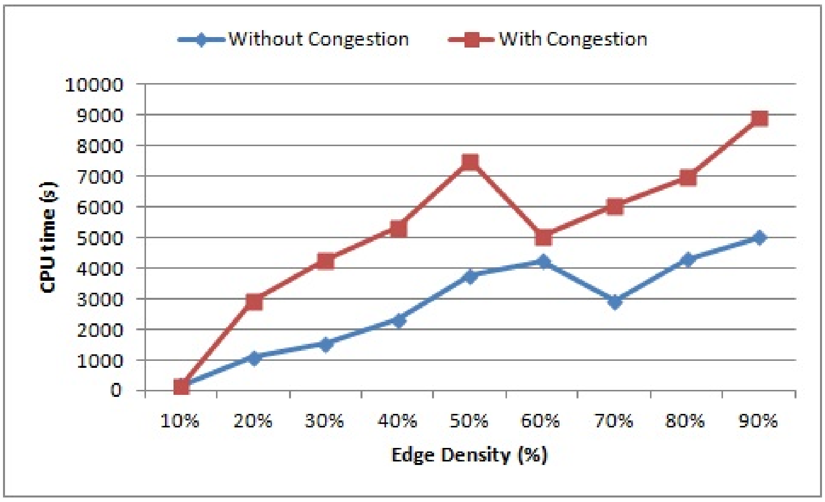

Figure V.14: CPU time vs. Edge Density for 40 nodes and 5 commodities networks

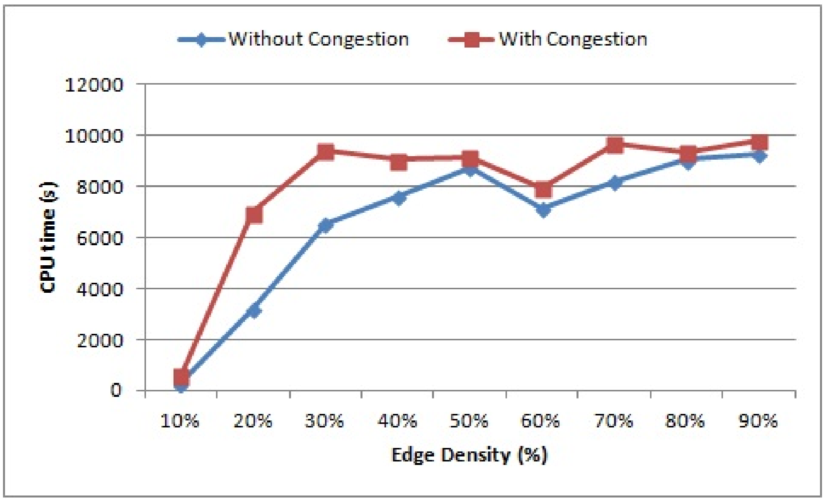

Figure V.15: CPU time vs. Edge Density for 40 nodes and 7 commodities networks

Table 5: Results of networks with 40 nodes, varying % density and commodities

Nodes	**% density**	**Commodities**	**% gap without Congestion Costs**	**% gap with Congestion Costs**	**Avg CPU time without Congestion Costs**	**Avg CPU time with Congestion Costs**
40	10%	3	0	0	62.77	82.21
40	20%	3	0	0	126.29	666.06
40	30%	3	0	0	512.10	933.16
40	40%	3	0	0	512.12	724.99
40	50%	3	0	0	702.02	2040.40
40	60%	3	0	0	295.63	2619.85
40	70%	3	0	0	830.71	3087.54
40	80%	3	0	0	750.98	1544.63
40	90%	3	0	0	1222.57	1608.79
40	10%	5	0	0	184.37	186.66
40	20%	5	0	0	1098.95	2927.92
40	30%	5	0	0	1531.75	4301.87
40	40%	5	0	0	2314.7	5334.37
40	50%	5	0	0	3764.04	7490.91
40	60%	5	0	0	4238.89	5076.33
40	70%	5	0	0	2948.35	6045.43
40	80%	5	0	0	4302.55	6984.12
40	90%	5	0	0	5016.87	8932.11
40	10%	7	0	0	300	587
40	20%	7	0	0	3186.38	6983.08
40	30%	7	0	0	6563.74	9411.85
40	40%	7	0	0	7600	9060.45
40	50%	7	0	0	8751.71	9181.05
40	60%	7	0	0	7123.57	7929.16
40	70%	7	0	0	8200	9701.26
40	80%	7	0	0	9060.98	9071.98
40	90%	7	0	0	9263.32	9400.9
40	10%	10	0	0	979.64	1131.25
40	20%	10	0	5	9340.91	10000.4
40	30%	10	15	8	10003.8	10008.3
40	40%	10	28	21	10000.5	10005.3
40	50%	10	33	33	10003.3	10006.3
40	60%	10	30	39	10000.0	10000.3
40	70%	10	28	37	10002.6	10005.8
40	80%	10	39	45	10000.5	10000.6
40	90%	10	32	47	10000.1	10000.2

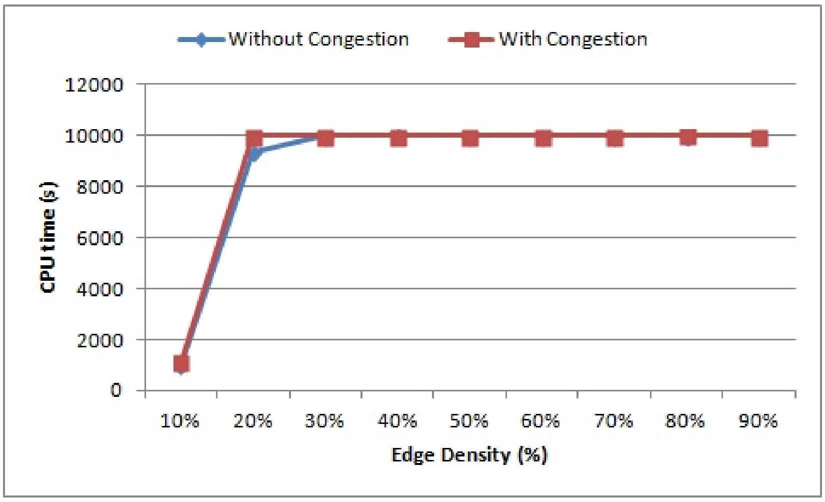

Figure V.16: CPU time vs. Edge Density for 40 nodes and 10 commodities networks

Table 6: Results of networks with 50 nodes, varying % density and commodities

Nodes	% density	Commodities	% gap without Congestion Costs	% gap with Congestion Costs	Avg CPU time without Congestion Costs	Avg CPU time with Congestion Costs
50	10%	3	0	0	162.35	1500
50	20%	3	0	0	631.99	1750.97
50	30%	3	0	0	1842.24	4549.98
50	40%	3	0	0	4371.92	5055.56
50	50%	3	0	0	2220.50	5629.54
50	60%	3	0	0	1709.36	5722.62
50	70%	3	0	0	1959.88	7736.54
50	80%	3	0	6	3473.86	10003.4
50	10%	5	0	0	487.48	1988.46
50	20%	5	0	0	2877.21	8404.80
50	30%	5	0	5	5759.06	10001.3
50	40%	5	0	9	7388.33	10008.2
50	50%	5	0	6	8459.41	10000.4
50	60%	5	0	10	7610.30	10004.1
50	70%	5	0	7	7893.57	10001.1
50	80%	5	0	28	8505.48	10001.4

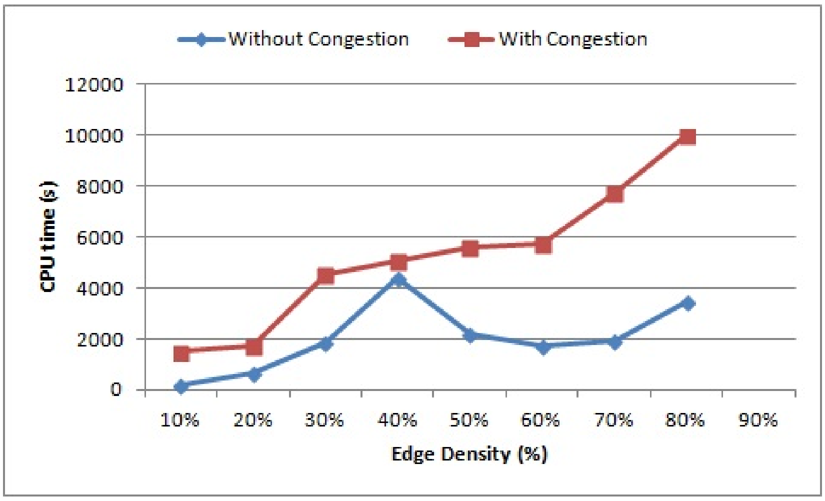

Figure V.17: CPU time vs. Edge Density for 50 nodes and 3 commodities networks

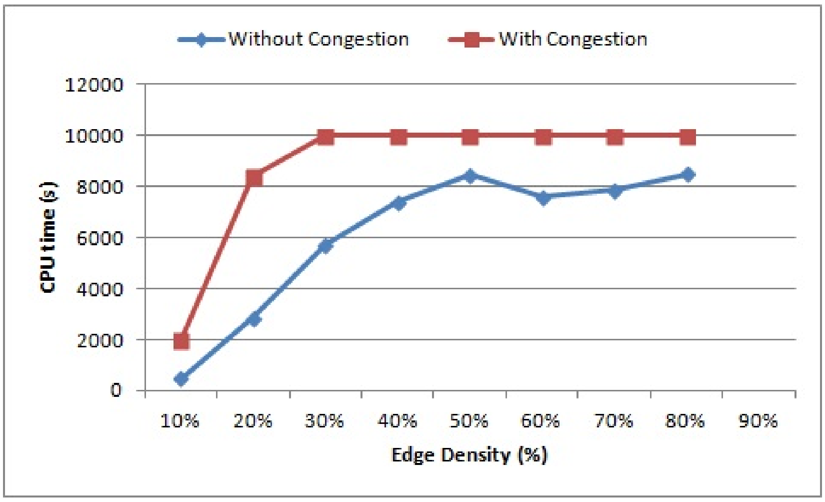

Figure V.18: CPU time vs. Edge Density for 50 nodes and 5 commodities networks

C. French Rail Network

In this section we consider an example based on the French rail network. Figure V.19 illustrates the existing French railroad map. We constructed an example for this rail network, by adding potential routes between cities. The newly constructed network is shown in Figure V.20. This figure has 88 nodes (i.e. 88 cities) and 394 edges.

To model the flow of traffic across the network, we consider two origins of flows; one in the North (Paris) and one in the South (Montpellier). Even if more accurate designs might be obtained when more origins of flow are considered, however, we were bound by the computational time. This lead to the consideration of only two origins of flow.

This network was run with our proposed model while including congestion costs. Figure V.21 shows the optimal route for the French network for people whose origins are Paris or Montpellier.

We obtained different routes than Figure V.19. The main differences are shown, for example, when going from Paris to Tours. In Figure V.19, this is possible through two different routes: Paris-Orleans-Tours or through a direct route (Paris-Tours). However, according to the optimal solution obtained by our proposed model, there is only one optimal route for people going from Paris to Tours and it is by passing by Le Mans. Another example would be the route from Paris to Dole. Figure V.19 states that the route is Paris-Dijon-Dole. However, Figure V.21 shows that the only way to get to Dole is through this particular route: Paris-Reims-Saint Diziers-Dijon-Dole. This reaffirms the main aim of hub-and-spoke networks which is getting the optimal route at the lowest costs, however, not through direct routes. To further prove our point, let's take for example people leaving Montpellier and headed to Clermont Ferrand. Figure V.19 would take them to Nimes, to Ales and finally to Clermont Ferrand. However, our built network would take them through the following route: Montpellier-Nimes-Arles-Avignon-Orange-Montelimar-Valence-Grenoble-Lyon-Lapalisse-Clermont Ferrand. This is a longer route, however, it is surely a cheaper one. In addition, this real-life application proves that dealing with multiple origin-destination pairs needs the modelling of multicommodities, where every origin-destination pair is considered a commodity.

Figure V.19: France Railroad Map

Figure V.20: Modified France Railroad Map

Figure V.21: Solution of the France Railroad Map

CHAPTER VI

CONCLUSION

Hub-and-spoke networks present a lot of advantages, mainly related to the consolidation of flow resulting in cost reduction due to economies of scale. However, having to deal with one type of products does not mesh well with real-life applications. Thus, it is primordial to seek a model that accomplishes the flow of several commodities in a hub-and-spoke network. This thesis aimed at proposing such a model.

We developed a mixed integer problem combining multicommodity within the hub-and-spoke network. Knowing that congestion plays a primordial role in hub-and-spoke networks, we presented an extension to the model that includes congestion costs to lessen their effect on the network. We then conducted computational testing using instances generated randomly with the *Mulgen* generator. In addition, we presented a real case network inspired by the French rail network. It was noticeable that including congestion costs in the model made it more challenging to solve. It required a longer computational time. Nevertheless, with 3 hours of CPU time, a maximum of 50 nodes and 5 commodities can be solved to optimality.

BIBLIOGRAPHY

[1] A. A. Assad. Multicommodity network flows; a survey. *Networks*, 8(1):37–91, 1978.

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

[3] T. Aykin. The hub location and routing problem. *European Journal of Operational Research*, 83(1):200–219, 1995.

[4] J. E. Beasley. An algorithm for solving large capacitated warehouse location problems. *European Journal of Operational Research*, 33(3):314–325, 1988.

[5] M. Bellmore, G. Bennington, and S. Lubore. A multivehicle tanker scheduling problem. *Transportation Science*, 5(1):36–47, 1971.

[6] K. Bonsor. How airlines work. June 2001. (http://science.howstuffworks.com/transport/flight/modern/airline3.htm).

[7] J. F. Campbell. Location and allocation for distribution systems with transshipments and transportion economies of scale. *Annals of Operations Research*, 40(1):77–99, 1992.

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

[9] J. F. Campbell, A. T. Ernst, and M. Krishnamoorthy. Hub arc location problems: Part ii-formulations and optimal algorithms. *Management Science*, 51(10):1556–1571, 2005.

[10] B. Cao and G. Uebe. An algorithm for solving capacitated multicommodity p-median transportation problems. *Journal of the Operational Research Society*, pages 259–269, 1993.

[11] N. Christofides and J. E. Beasley. A tree search algorithm for the p-median problem. *European Journal of Operational Research*, 10(2):196–204, 1982.

[12] S. Clarke and J. Surkis. An operations research approach to racial desegregation of school systems. *Socio-Economic Planning Sciences*, 1(3):259–272, 1968.

[13] P. Connor. Railway technical finance. October 2013. (www.railway-technical.com/finance.shtml).

[14] T. G. Crainic, A. Frangioni, and B. Gendron. Bundle-based relaxation methods for multicommodity capacitated fixed charge network design. *Discrete Applied Mathematics*, 112(1):73–99, 2001.

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

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

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

[18] S. Gelareh and S. Nickel. Hub location problems in transportation networks. *Transportation Research Part E: Logistics and Transportation Review*, 47(6):1092–1111, 2011.

[19] A. M. Geoffrion and G. W. Graves. Multicommodity distribution system design by benders decomposition. *Management science*, 20(5):822–844, 1974.

[20] J.-L. Goffin, J. Gondzio, R. Sarkissian, and J.-P. Vial. Solving nonlinear multicommodity flow problems by the analytic center cutting plane method. *Mathematical Programming*, 76(1):131–154, 1997.

[21] A. Goldman. Optimal locations for centers in a network. *Transportation Science*, 3 (4):352–360, 1969.

[22] P. G. Grove and M. E. O'Kelly. Hub networks and simulated schedule delay. *Papers in Regional Science*, 59(1):103–119, 1986.

[23] M. Guignard. A lagrangean dual ascent algorithm for simple plant location problems. *European Journal of Operational Research*, 35(2):193–200, 1988.

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

[25] B. Y. Kara and B. Ç. Tansel. The latest arrival hub location problem. *Management Science*, 47(10):1408–1420, 2001.

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

[27] J. G. Klincewicz. Avoiding local optima in the p-hub location problem using tabu search and grasp. *Annals of Operations Research*, 40(1):283–302, 1992.

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

[29] J. G. Klincewicz and H. Luss. A lagrangian relaxation heuristic for capacitated facility location with single-source constraints. *Journal of the Operational Research Society*, 1986.

[30] T. Larsson and D. Yuan. An augmented lagrangian algorithm for large scale multicommodity routing. *Computational Optimization and Applications*, 27(2):187–215, 2004.

[31] J. Liu, C.-L. Li, and C.-Y. Chan. Mixed truck delivery systems with both hub-and-spoke and direct shipment. *Transportation Research Part E: Logistics and Transportation Review*, 39(4):325–339, 2003.

[32] J. C. Martín and C. Román. Analyzing competition for hub location in intercontinental aviation markets. *Transportation Research Part E: Logistics and Transportation Review*, 40(2):135–150, 2004.

[33] C. Mayer and T. Sinai. Network effects, congestion externalities, and air traffic delays: Or why all delays are not evil. Technical report, National Bureau of Economic Research, 2002.

[34] R. D. McBride and J. W. Mamer. Solving the undirected multicommodity flow problem using a shortest path-based pricing algorithm. *Networks*, 38(4):181–188, 2001.

[35] S. Nickel, A. Schöbel, and T. Sonneborn. Hub location problems in urban traffic networks. *Mathematical methods on optimization in transportation systems, Kluwer Academic Publishers, Dordrecht, The Netherlands*, pages 95–107, 2001.

[36] A. G. C. of America. Agc national traffic survey part one: Measuring the impact of highway congestion on the construction industry. 2010. (www.agc.org/galleries/news/National20Congestion20Survey.pdf).

[37] M. E. O'Kelly. The location of interacting hub facilities. *Transportation Science*, 20(2):92–106, 1986.

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

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

[40] H. Pirkul and D. A. Schilling. An efficient procedure for designing single allocation hub and spoke systems. *Management Science*, 44:235–242, 1998.

[41] M. Rao. *A Column Generation Scheme for the Deterministic Multi-commodity Warehousing Model with Cash-liquidity Constraints*. Center for Operations Research & Econometrics, 1969.

[42] D. Serra, V. Marianov, et al. Location models for airline hubs behaving as m/d/c queues. 2013.

[43] G. D. Taylor, T. Meinert, R. Killian, and G. Whicker. Development and analysis of alternative dispatching methods in truckload trucking. *Transportation Research Part E: Logistics and Transportation Review*, 35(3):191–205, 1999.

[44] H. Topcuoglu, F. Corut, M. Ermis, and G. Yilmaz. Solving the uncapacitated hub location problem using genetic algorithms. *Computers & Operations Research*, 32 (4):967–984, 2005.

[45] W. W. White and A. M. Bomberault. A network algorithm for empty freight car allocation. *IBM Systems Journal*, 8(2):147–169, 1969.

[46] W. I. Zangwill. A backlogging model and a multi-echelon model of a dynamic economic lot size production system-a network approach. *Management Science*, 15 (9):506–527, 1969.