Dynamic Facility Location with Stochastic Demand and Congestion

Masoud Madani

A Thesis
in
The Department
of
Mechanical, Industrial and Aerospace Engineering

Presented in Partial Fulfillment of the Requirements
for the Degree of Master of Applied Science (Industrial Engineering) at
Concordia University
Montreal, Quebec, Canada

June 2018

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis prepared

By: **Masoud Madani**
Entitled: **Dynamic Facility Location with Stochastic Demand and Congestion**

and submitted in partial fulfillment of the requirements for the degree of

Master of Applied Science (Industrial Engineering)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.
Singed by the final examining committee:

Dr. Brandon Gordon	Chair
Dr. Mingyuan Chen	Internal Examiner
Dr. Satyaveer Chauhan	External Examiner
Dr. Ivan Contreras	Supervisor
Dr. Navneet Vidyarthi	Co-supervisor

Approved by

Chair of Department or Graduate Program Director

2018

Dean of Faculty

Abstract

Dynamic Facility Location with Stochastic Demand and Congestion

Masoud Madani

In this thesis, we study a multi-periodic facility location problem with stochastic demand to determine the optimal location, capacity selection and demands allocation of facilities within distinct time periods, while, each facility contains a server with a limited capacity. It causes facilities to experience a period of congestion, when not all arriving demands can be served immediately. Customers that arrive in this period might await service in a queue. This thesis perspective incorporates customers waiting costs as part of the objective. In this case, facilities do not utilize whole of the established capacity to ensure a maximum waiting time of the allocated customers. Firstly, a mathematical model is presented for a dynamic facility location problem with stochastic demand and congestion. The problem is setup as a network of spatially distributed queues and formulated as a nonlinear mixed integer program (MINLP). To transform the nonlinear congestion function to a piecewise linear, a linearization method is adapted. This method adds a set of inequalities to the model. We show that lifting this set of inequalities, with keeping generality of the method, reduces CPU times up to 3.5 times, on average. Moreover, a decent heuristic is proposed to solve the problem. Computational experiments indicate that the heuristic results in less costly solutions than them obtained by CPLEX algorithms, in 58% of relatively-difficult test problems.

Keywords: *Facility Location (FL), Congestion, Linearization, Valid Inequalities, Lifting Inequalities, Branch-and-Cut*

Acknowledgements

I wish to express my deepest gratitude to my supervisors
Dr. Ivan Contreras
Dr. Navneet Vidyarthi
who have always supported me in my master study and research with patience, motivation and immense knowledge. Their guidance and kind support helped me in all the time of research and writing this thesis. I cannot imagine having better supervisors for my master thesis.

My thankfulness is also to
Carlos Armando Zetina
Moayad Tanash
for their valuable suggestive points in this work. It was a great honor and privilege to have friendship and knowledge-sharing with such generous lab-mates.

I am especially grateful to my parents
Mojtaba Madani
Fatemeh Nabizadeh
who have my constant appreciation. They have always been there to affectionately support and encourage me to realize my dreams. Your teachings go beyond any academic program and your endless devotion will always be inspiring for me.

I would also like to thank my forever love
Neda Molanorouzi
for her support and encouragement. I feel so fortunate that ending of this thesis coincides with emergence of your love and commencement of a new life with your company.

My last but not least appreciation is to all of my friends who helped and accompanied me in my master studies and also to the staff of the Department of Mechanical, Industrial and Aerospace Engineering of Concordia University.

Contents

1 Preliminaries

1.1 Introduction

1.1.1 Facility Location

The ubiquity of locational decision-making has led to a strong interest in location analysis and modeling within the operations research and management science communities. The long and voluminous history of location research results from several factors. First, location decisions are frequently made at all levels of human organization from individuals and households to firms, government agencies and even international agencies. Second, such decision are often strategic in nature. That is, they involve large sums of capital resources and their economic effects are long term. Third, they frequently impose economic externalities such as pollution, economic development, congestion, etc. For an introduction to basics of this topic, the reader referred to the texts by Drezner (1996), Hamacher and Drezner (2002), Daskin (2011) and Laporte, Nickel, and Gama (2015).

The mathematical science of facility locating has attracted much attention in discrete and continuous optimization over nearly last four decades. Facility location problems locate a set of new facilities (resources) to minimize the cost of satisfying some set of demands (of the customers) with respect to some set of constraints. In basic facility location problems, this cost is consist of two parameters; establishment cost of new facilities (also called location cost or fixed cost) and transportation cost from facilities to customers (also called allocation cost or variable cost). The basic components of location-allocation problems can be thought to consist of facilities, locations, and customers. The type of a facility is another important property, in the simplest case, all the facilities are supposed to be identical with respect to their size and the kind of service they offer. However, it is often necessary to locate facilities that differ from one another (Farahani and Hekmatfar, 2009). Investigators have focused on both algorithms and formulations in diverse settings in both the private sectors (e.g., industrial plants, banks, retail facilities, etc.) and the public sectors (e.g., hospitals, post stations, etc.). The study of location theory started formally in 1909 when Weber considered how to locate a single warehouse in order to minimize the total distance between the warehouse and several customers. After that, location theory was driven by a few applications. Location theory gained researchers' interest again in 1964 with a publication by Hakimi (1964), who wanted to locate switching centers in a communications network and police stations in a highway system. Facility location books are numerous. Francis, McGinnis, and White (1992) introduced some prevalent models such as single/multi facility location problems, quadratic assignment location problems (QAP) and covering problems. Mirchandani and Francis (1990) wrote about discrete location theory. The network based location theory book by Daskin (1995) focused on discrete location problems. Drezner (1995) represented some models and applications in location environments. Hamacher and Drezner (2002) published a book about the theory and applications of facility location.

The term "facility" is used in its broadest sense. That is, it is meant to include entities such as distribution centres (DCs), air and maritime ports, factories, warehouses, retail outlets, schools, hospitals, bus stops, subway stations, electronic switching centres, computer concentrators and terminals, rain gages, emergency warning sirens, and satellites, to name but a few that have been analyzed in the research literature. The term "location problem" refers to the modeling, formulation, and solution of a class of problems that can best be described as locating facilities in some given spaces. Deployment, positioning, and locating are frequently used as synonymous. There are differences between location and layout problems: the facilities in location problems are small relative to the space in which they are sited and the interaction among facilities may occur; but in layout problems, the facilities to be located are large relative to the space in which they are positioned, and the interaction among facilities is common.

1.1.2 Congestion, why it is important

Traditionally, logistics analysts have divided decision levels into strategic, tactical and operational (Miranda and Garrido, 2006). There are also three important decisions within a supply chain: facilities location decisions; inventory management decisions; and distribution decisions (Shen and Qi, 2007). For example, in a distribution network,we could mention location of Distribution centers(DCs) as a strategic decision, distribution decisions as a tactical decision and inventory service level as a tactical or operational decision. Often, for modeling purposes, these levels are considered separately, and this may conduce to make non-optimal decisions, since in reality there is interaction between the different levels (Miranda and Garrido, 2006). For example, most well-studied location models do not consider inventory costs, and shipment costs are estimated by direct shipping. Although one may argue that tactical inventory replenishment decisions and shipment schemes are not at the strategic level, and we should not consider them in the strategic planning phase, failure to take the related inventory and shipment costs into consideration when deciding the locations of facilities can lead to sub-optimality, since strategic location decisions have a big impact on inventory and shipment costs (Shen and Qi, 2007). In this end, in addition to the generic facility location setup, also other areas such as allocation, capacity acquisition, procurement, production, inventory and routing have to be considered (Cordeau, Pasin, and Solomon, 2006). As Klose and Drexl (2005) state, researchers have focused relatively early on the design of distribution systems but without considering the supply chain as a whole.

On the other hand, firms would like to consider cost and service levels simultaneously. Due to competitiveness of today's global business environment, one of the most critical considerations in distribution network design (DND) is lead time, because it strongly impacts the overall distribution cost and also the customers contentment. Actually, lead time is viewed as an important performance measure that represents the firm's commitment on customers satisfaction (Vidyarthi, Elhedhli, and Jewkes, 2009). Part of the planning processes in Supply chain management (SCM) aims at finding the best possible distribution network configuration. It is good to have many DCs, since this reduces the cost of transporting product to customers (or retailers) and will provide better service. Also, it is good to have few DCs, since this reduces the cost of holding inventory via pooling effects, and reduces the fixed costs associated with operating DCs via economies of scale (Erlebacher and Meller, 2000). On the basis of the above, facility location has become a major challenge for firms as they simultaneously try to reduce costs and improve customer service in today's increasingly competitive business environment (Daskin, Coullard, and Shen, 2002). Chopra and Meindl (2010) study network design strategist with various objectives ranging from low cost to high responsiveness. The goal of cost reduction is to provide motivation for centralization of facilities. On the other hand, the goal of customer responsiveness is to provide motivation for having goods (or service centers) as near to the final consumer as possible, with the least waiting time to receive the required goods (or service). Thus, there is a basic conflict between these objectives and facility location is a critical decision in finding an effective balance between them (Nozick and Turnquist, 2001).

As mentioned before, the element which specifies the responsibility and service level of a distribution network is lead time (Beamon, 1998). One body of previous work constituted by the papers of Berman, Larson, and Chiu (1985), Crainic and Laporte (1997), Owen and Daskin (1998), Jamil, Baveja, and Batta (1999), Eskigun (2002), Eskigun et al. (2005) and Sourirajan, Ozsen, and Uzsoy (2007) explicitly considers lead time in network design. In traditional models that support lead time reduction, customers demands are supposed deterministic and the main objective is minimizing fixed cost of facility location and variable transportation cost (see Dogan and Goetschalckx (1999), Vidal and Goetschalckx (2000), Teo and Shu (2004), Shen (2005), Amiri (2006), Elhedhli and Gzara (2008), and references therein). Min and Zhou (2002) suggest that future research should obviously consider interaction of logistics cost with lead time.

1.1.3 Dynamic Facility Location

As Ballou (1968) states: "the effect of future time dimensions cannot be neglected in location analysis." In many situations, several parameters change over time. Thus, to adapt the configuration of facilities to these parameters, dynamic facility locations have been interest of researchers since the pioneering work of Manne (1961). Dynamic models incorporate time. Current, Ratick, and ReVelle (1998) define two categories of dynamic models: "implicitly" dynamic and "explicitly" dynamic. Implicitly dynamic models are "static" in the sense that all of the facilities are to be opened at one time and remain open over the planning horizon. They are dynamic because they recognize that problem parameters (e.g., demand) may vary over time and attempt to account for these changes in the facility location scheme generated. Examples of implicitly dynamic models include Mirchandani and Odoni (1979), Weaver and Church (1983), Drezner and Wesolowsky (1991) and Drezner (1995), which consider problems where demand and travel times change over time. Explicitly dynamic models are those designed for problems where the facilities will be opened (and possible closed) over time. Early examples of such problems include Roodman and Schwarz (1975), Wesolowsky and Truscott (1976), Campbell (1990) and Schilling (1980). As pointed out by Arabani and Farahani (2012), the notion of what dynamic means may differ when dealing with different areas of facility location. The decision to open and close facilities over time is related to changes in the problem parameters over time. Examples of parameters that might change include demand, travel time/cost, facility availability, fixed and variable costs, profit and the number of facilities to be opened. Owen and Daskin (1998) and Farahani, Abedian, and Sharahi (2009) review a survey on dynamic facility location problem (FLP). To approach these problems, multi-period location models have been proposed in the literature. In these models, the planning horizon is divided into several time periods. Most dynamic FLPs can be seen as multi-periodic extensions of classical location problems (Jena, Cordeau, and Gendron, 2015). Such a planning horizon leads to several achievements including appropriate timing of location decisions and adjustable anticipation of favorable/unfavorable fluctuations (Miller et al., 2007).

1.1.4 Facility Location with Congestion

Multi-period planning could also be combined with stochasticity. This is the situation when the probabilistic behavior of the uncertain parameters changes itself over time (Melo, Nickel, and Saldanha-Da-Gama, 2009). Snyder (2006) reviews the literature on stochastic and robust facility location models, where costs, demands, travel times and other inputs to the classical models might be highly uncertain. If these uncertain parameters includes the both of demands and service time of customers, while facilities are capacitated, this circumstance leads to congestion, where some of the arriving demands cannot be served immediately and must wait in queue for the service (Berman and Krass, 2001). Huang, Batta, and Nagi (2005) is one of the first to explicitly model the effect of congestion in location problems. Queuing aspects of the problem is considered by Larson (1974), Berman, Larson, and Chiu (1985), Marianov and ReVelle (1996), Arkat and Jafari (2016), Alijani et al. (2017), Ahmadi-Javid and Hoseinpour (2017) and Zaferanieh and Fathali (2017). Applications of these models range from public service facilities such as hospital, medical clinics and government offices, to private facilities such as retail stores or repair shops. Several of these applications are listed in table 1.1. One of the most significant application areas FLP with congestion is distribution network design of emergency service facilities, such as hospitals, fire stations, police stations or ambulances, where lack of immediate demands satisfactions could be disastrous (Marianov and ReVelle, 1996).

1.1.5 Impact of congestion on facility location decisions

Regarding congestion cost as an element of total cost rises new considerations as decision criteria. One of the common observations in congested networks is that service providers do not utilize all of the established capacity of distribution facilities. As mentioned before, in real world problems, many parameters such as customer demands and taken provider's time to satisfy each demand are uncertain, when the capacity of distribution facilities is limited. Consequently, it is possible that a customer (demand) arrives to a provider's facility when the facility is occupied by another customer (or issuing another demand). As a result, the customer (demand) must wait in a queue or would be lost for the system. Thus, the established capacity would be more than the predicted workload in each distribution facility. In other words, in a congestion network, distribution facilities consider a safety zone in determination of capacities to avoid corruption and improve the responsiveness of the network. As much as the uncertainty escalates, the ideal quantity of this safety zone increases. Therefore, in a congested distribution network design, the established potential of the network is higher than the aggregation of predicted demands, while, such a determination is counted fruitless or even counterproductive in traditional network design without concern of lead time.

Long waiting times in distribution facilities are counted as inefficiency of a network. Such a decision criteria brings new considerations to locational decisions. One of the most essential considerations, among others, is the utilization of distribution facilities which should be discriminated from the capacity established for the facility. Facility utilization strategies or workload allocation strategies are scrutinized in congested networks to arrange efficient distribution networks with concern of lead time. Usually, these strategies are driven by urgency level of the product/service being provided in the network. For example, high urgency level of the product/service leads to centralization of the proportions of facilities utilization (allocation workload / established capacity). It is observed that in distribution networks of highly urgent products/services, the proportions of utilization are relatively normally distributed, while, in it of less urgent products/services, diverse utilizations are more common among distribution facilities, i.e. several facilities might be substantially more congested than others.

In mathematical modeling perspective, usually, there is not any additional decision variable to incorporate congestion into the classic facility location problem. In most of the FLP with congestion addressed in the literature, as well as the classic FLP, decision variables are routing and flow variables which represent facility establishment in candidate locations and their allocation to demand zones, respectively. However, in FLP with congestion, these variables are determined with some additional criteria. Therefore, they deliver more inclusive assumptions. Actually, determining the fixed cost location and variable allocation decisions specifies the congestion, implicitly. Thus, there a strong correlation between traditional concepts of FLP and the new congestion-relative concepts such as proportion (or percent) of capacity utilization, allocated workload to a facility, congestion cost, average waiting time for each demand (customer) and so on. Moreover, in contrast with the mathematical model representing a classic FLP, the model which represents it with congestion is a nonlinear model. These factors make solving congested FLPs noticeably more challenging.

1.1.6 Contribution of the research

This research studies a strategical problem which is a dynamic facility location problem, while demand arrivals and general service time in facilities are under Poisson distribution. The objective of this problem seeks to simultaneously determine the location and capacity of facilities and allocate stochastic customers demands to facilities by minimizing the fixed cost of establishing facilities, equipping them with sufficient capacity and the variable cost of serving customers, in addition of congestion cost and transportation cost between demand zones and facility locations, for whole of the planning horizon divided into consecutive time periods. Considering congestion cost as an element of the total cost leads the problem to a mixed integer nonlinear program (MINLP). Having dealt with a nonlinear model, it is not possible to obtain the optimal solution by usual optimization solvers. Thus, as the same as several of previous researches on similar problems, a linearization method is introduced with controllable approximation gap. As a result, a lower bound (LB) and an upper bound (UB) are provided for the optimal solution value. Moreover, due to the complexity of the program and also, the massiveness of large-scaled networks, it is a challenge to estimate a tight bound in an appropriate time, even by state-of-are solvers. Although the fact that either traditional solution methods or default configuration of general solvers are able to provide generally-accepted bounds in reasonable times, the strategic nature of this decision motivates us to investigate modern solution methods to achieve tighter bounds for the optimal solution value of the problem.

The contributions of this research are in various aspects as follows;

- **Mathematical modeling:** Firstly, a mathematical model for the problem is presented. To the best of our knowledge, it is the first time that congestion consideration is modeled for a dynamic version of facility location problem. In addition, a linearization method is adapted for the model. Thus, this thesis is an explicit extension of the two following papers;
  - Vidyarthi, Elhedhli, and Jewkes, 2009, which studies a single-period facility location with stochastic demands and congestion
  - Jena, Cordeau, and Gendron, 2015, which presents a generalized modular formulation for dynamic facility location problems
  - Elhedhli, 2005, which introduces a piecewise linearization that is solved by a cutting plane method
- **Formulation tightening:** As the proposed model is nonlinear, a piecewise linearization method is adapted which transforms the MINLP model into MIP and approximates the original optimal solution value ϵ-optimally. One of the impacts of this method is that several new sets of constraints with inequality form are added to the model. In this research, one set of these inequalities is lifted to tighten the formulation of the MIP model. In other words, a new technique is presented that is an explicit contribution to the linearization method introduced by Elhedhli, 2005. As is shown in the numerical result, this lifting tightens the LP relaxation of the model tremendously and results in expedition in the LB estimation, so that the problem could be solved more than 3.5 times faster, on average. It means that the CPU times taken to solve the tightened model is averagely less than 28% of it for the traditional model. As a consequence, in a given time,

a proper LB could be obtained for larger networks (or more difficult problems in any aspect). Also, for a same problem, a higher LB could be earned within an equal time.

- **Heuristic solution methods:** At the end, a heuristic method is introduced to attain close-to-optimal solutions in considerably less time than exact methods. Comparison of the heuristically obtained solution with the bounds calculated exactly in medium/large-scaled problems demonstrates the quality of the heuristic. Thus, for very large-scaled problems, the heuristic methods could be employed to approximate the optimal solution with a reasonable time limitation, when the traditional solution are not able to provide even a feasible solution. In strategical problems, the quality of solution are superior to solving time. So, the main benefit of this heuristic is that it results in obtaining solutions with less cost in several test problems. As is indicated in the last chapter, in 58% of relatively difficult test problems, the best known solution is obtained by the heuristic.

1.2 Literature Review

A related branch of literature considers models in which the facilities may be unable to provide service due to facility disruptions (Bundschuh, Klabjan, and Thurston (2003); Berman, Krass, and Menezes (2007); Snyder and Daskin (2005); Zarrinpoor, Fallahnezhad, and Pishvaee (2018)) or link failure (Nel and Colbourn (1990); Eiselt, Gendreau, and Laporte (1992)). One of the new approaches to study FLP in distribution networks is considering traffic congestion. Examples of these models are addressed by Bai et al. (2011) and Jouzdani, Sadjadi, and Fathian (2013). In the following, several parts of the literature are introduced that mostly focus on the congestion impacted by facility location.

1.2.1 Problem Objective

Facility location models with congestion are also classified by the goal of the problem. As an example, one of classes is consist of *Coverage* models, which aim to design a system providing sufficient service to customers. Typically, the objective of these models is to maximize the captured demands. As a consequence, they enforce each customers to travel to the closest available facility (Berman, Krass, and Wang, 2006). Among the very first of such models in the literature, Daskin (1982), ReVelle and Hogan (1989), Ball and Lin (1993), Baron, Berman, and Krass (2008), Berman and Drezner (2006), Moghadas and Kakhki (2011) and Marianov and Serra (1998) could be mentioned. It is shown that under quite general conditions, the optimal facility configuration is one that ensures that each facility sees (approximately) the same demand (Baron, Berman, and Krass, 2008). This important insight for coverage-type models motivated Baron et al. (2007), Berman et al. (2009) and Suzuki and Drezner (2009) to study "Equitable Location Problems", there is, a deterministic problem seeking to locate a set of facilities so that the attracted demand is distributed as evenly as possible. Amid case studies, Silva and Serra (2016) address an emergency services location problem with different queuing priorities.

1.2.2 Service Level

Another category of location models with congestion is *Service-Objective* models, which seek designing a system that optimizes customer service with limited

TABLE 1.1: Applications of balanced-objective models mentioned in the literature

Application	Reference
bank branches and automated banking machines	Seifbarghy and Mansouri (2016) Aboolian, Berman, and Drezner (2008) Pasandideh and Chambari (2010) Wang, Batta, and Rump (2002)
automobile emission testing stations	Castillo, Ingolfsson, and Sim (2009)
virtual call centres	Castillo, Ingolfsson, and Sim (2009)
web service providers' facilities	Aboolian, Sun, and Koehler (2009)
proxy/mirror servers in communication networks	Wang, Batta, and Rump (2004)
waterborne containerized imports	Jula and Leachman (2011) Leachman and Jula (2011)
distribution centres(DCs) in supply chains	Huang, Batta, and Nagi (2005) Vidyarthi and Jayaswal (2014) Vidyarthi, Elhedhli, and Jewkes (2009)
medical clinics and preventive health care facilities	Vidyarthi and Kuzgunkaya (2015) Zhang et al. (2010) Zhang, Berman, and Verter (2009) Zhang, Berman, and Verter (2012)

resources. Thus, in these models, available service capacity is specified through constraints, rather than through the objective function term. Among the paper which addressed such models, Drezner and Drezner (2011) and Hamaguchi and Nakade (2010) could be mentioned. Since service level is typically defined as the combination of travel and congestion cost, in these models, congestion is regarded in objective function. Due to the fact that the congestion term involved in objective function only measures the aggregate congestion, several authors (see Boffey, Galvão, and Marianov (2010), Marianov and Serra (2011), Marianov, Boffey, and Galvão (2009) and Wang, Batta, and Rump (2002)) impose service level constraints to ensure that congestion is controlled by each facility.

1.2.3 Balance Orientation

Balanced-Objective models are presented in modern approaches in this field, in sake of a social optimum in the designed distribution network, there is, the costs of service facility and the corresponding capacity establishment are regarded in the objective function, as well as travel and congestion costs that charge customers. Such models are pointed in Castillo, Ingolfsson, and Sim (2009), Elhedhli and Hu (2005), Elhedhli (2006), Kim (2013), Marianov and Ríos (2000), Vidyarthi and Jayaswal (2014) and Jayaswal and Vidyarthi (2017). In these models, customers accept the directed assignments to optimize social welfare, even if this lead to assignments that are suboptimal from individual customers' point of view. Aboolian, Berman, and Drezner (2008) and Abouee-Mehrizi et al. (2011) introduce models which incorporate customers response to the issue. Pasandideh and Chambari (2010) propose a bi-objective model to approach the balanced-oriented facility location problems within queuing framework. Rabieyan and Seifbarghy (2010) formulate profitability in FLP with congestion, while just a subset of stochastic demands is satisfied and the objective is maximizing the total benefit. Wang, Batta, and Rump (2004) present three models for FLP with congestion with different perspectives, that of (i) the service provider (wishing to limit costs of setup and operating servers), (ii) the customers (wishing to limit costs of accessing and waiting for service), and (iii) both the service provider and the customers combined. In all cases, a minimum level of service quality is ensured by imposing an upper bound on the server utilization rate at a service facility. Seifbarghy and Mansouri (2016) consider the quality of the service provided in server facilities experiencing M/M/1 queuing policy, in addition of the cost and time. Fischetti, Ljubić, and Sinnl (2016) assume the customer allocation cost to be a linear or separable convex quadratic function. Hajipour et al. (2016) propose a multi-objective FLP with congestion using classical queuing systems. They consider three objective functions aiming at: (1) minimizing the sum of aggregate travel and waiting times; (2) minimizing the cost of establishing the facilities; and (3) minimizing the maximum idle probability of the facilities. The problem is formulated as a multi-objective non-linear integer mathematical programming model. Tavakkoli-Moghaddam et al. (2017) consider situations in which immobile service facilities are congested by a stochastic demand following M/M/m/k queues.

According to the nature of the problem under study in this thesis, the presented model is classified as *balanced-objective*. Thus, a detailed literature review of presented solutions for these categories of congested location problems is demonstrated in the following. Furthermore, table 1.1 is provided to indicate several papers with case study which involve this category of congestion models.

1.2.4 Solution Methods

Immobile facility location problems with congestion regarding the both providers' and customers' cost in the objective function (balanced-objective) are approached via two typical models. The first one is addressed by Castillo, Ingolfsson, and Sim (2009), who assume an M/M/1 queuing system and the facilities and use the average number of customers in the system. It leads to a Mixed Integer Programming (MIP) problem with a single concave term in the objective. Shen (2005) proposes a Lagrangian Relaxation method to solve it, while, Aboolian, Berman, and Krass (2007) presented a piecewise linear approximation. Hijazi, Bonami, and Ouorou (2013) show that this approximation is possible by either inner or outer linearization or both of them simultaneously.

The second approach to obtain exact solutions of balanced-objective location problems with congestion is based on Elhedhli (2006) who considers the expected queue length of facilities as a decision variable. Kim (2013) presents column generation heuristics to solve this class of models, while Vidyarthi and Jayaswal (2014) introduce an efficient solution, where the problem is set up as a network of independent M/G/1 queues, whose locations, capacities and service zones could be determined to ϵ-optimality using a constraint generation method. Also, Wang, Batta, and Rump (2002) provides various solution methods to find exact or heuristic solutions. Table 1.2 indicates the addressed solution methods used for balance-orientation models in recent years.

However, most of the studies in this field have been on static models. Typically, explicitly dynamic models extend the basic, static models with the addition of temporal subscripts to the facility location and assignment variables and constraints linking these variables over time (Drezner and Hamacher, 2001). Jena, Cordeau, and Gendron (2015) introduce a unifying model that generalizes existing formulations for several dynamic facility location problems and provides stronger linear programming relaxations than the specialized formulations.

TABLE 1.2: Solution methods used for balanced-objective models mentioned in the literature

Reference	**Exact Methods (ϵ-optimally)**					**(Meta)Heuristic Methods**			
	Lagrangean Relaxation	Cutting Planes	Branch-and-Bound	Benders Decomposition	Column Generation	Genetic Algorithm	Simulated Annealing	Other Evolutionary Algorithms	Heuristics
Tavakkoli-Moghaddam et al. (2017)								•	
Jayaswal and Vidyarthi (2017)		•							
Fischetti, Ljubić, and Sinnl (2016)				•					
Seifbarghy and Mansouri (2016)									•
Hajipour et al. (2016)								•	
Vidyarthi and Kuzgunkaya (2015)		•							
Kim (2013)					•				
Pasandideh and Chambari (2010)						•			
Rabieyan and Seifbarghy (2010)						•	•		•
Castillo, Ingolfsson, and Sim (2009)	•								
Elhedhli (2006)		•							•
Elhedhli and Hu (2005)	•								
Wang, Batta, and Rump (2004)			•					•	•
Marianov and Ríos (2000)		•							

To solve a facility location problem by branch-and-cut efficiently, Contreras, Tanash, and Vidyarthi (2016) describes a family of problem-specific valid inequalities, while more general forms are introduced by Ortega and Wolsey (2003) and Marchand et al. (2002). Bodur and Luedtke (2016) extends a family of these inequalities to empower the branch-and-cut method. Moreover, Fischetti, Ljubić, and Sinnl (2016) introduce a Benders decomposition method without separability to solve capacitated facility location problems. This method is also applicable when congestion is regarded in the problem. Hajipour and Pasandideh (2011), Hajipour and Pasandideh (2012), Hajipour, Khodakarami, and Tavana (2014) and Hajipour, Farahani, and Fattahi (2016) propose various metaheuristics such as Particle Swarm Optimization (PSO) and Vibration Damping Optimization (VDO) algorithm to solve congested location problems.

1.3 Thesis Structure

This manuscript is organized as follows. Chapter 2 presents the formal definition, modeling assumption, mathematical formulation, linearization and tightening methods including valid inequalities and lifting a set of constraints. Chapter 3 introduces several solution methods consisted of exact methods and heuristics. Finally, chapter 4 reports the results of computational experiments.

2 Mathematical Modeling

In this chapter, a mathematical model is presented for the problem, where each facility is modeled as an M/G/1 queue. The model is nonlinear, thus, an approximation is used to linearize it. This Linearization adds some constraints to the model. It also provides a lower bound and an upper bound for the optimal solution value. Then, the LP relaxation of the model is tightened in order to accelerate MIP solvers executions to solve the linearized model. The proposed model has some interesting properties explained in the following of the chapter. At the end, it is shown that the linearization introduced in the literature could be implemented more efficiently.

2.1 Problem Definition

This research studies a dynamic facility location problem, while demand arrivals and general service time in facilities are under probabilistic distribution. The objective of this problem seeks to simultaneously determine the location and capacity of facilities and allocate stochastic customers demands to facilities by minimizing the fixed cost of establishing facilities, equipping them with sufficient capacity and the variable cost of serving customers, in addition of congestion cost and transportation cost between demand zones and facility locations, for whole of the planning horizon divided into consecutive time periods. As a key assumption of a congested location problem, demands are stochastic, typically assumed to be a Poisson process, or, more generally a renewal process. In each time period, once the demand for a product is realized at the customers' end, the order is allocated to facilities, which operate in a build-to-order setting. Facilities maintain inventory of multiple components and facilitate the assembly and shipment of a variety of finished products without carrying expensive finished-goods inventory and without incurring long lead times. In build-to-order systems, customer order triggers the final assembly of finished product from components, hence the total lead time consists of assembly lead time and delivery lead time. Also, it is assumed that facilities contain resources (often called "servers") that have limited capacity and total lead time (service time) is stochastic. Having such assumptions combined with stochasticity of customers demands, facilities may experience a period of congestion, where not all arriving demands can be served immediately. Customers that arrive in this period might await service in a queue. This behavior results in having queues in facilities (Vidyarthi, Elhedhli, and Jewkes, 2009).

As a matter of fact, order processing lead time at facilities and consequently, waiting times in queues are highly dependent on the capacity of facilities and the allocated workload which are difficult to change (on a short term basis) once the facility is established. It is the consideration which differentiates this problem with a simple facility location problem, where facilities are homogeneous and whole of their capacity is allocated. Regarding queues in the established facilities stimulates the network to select a sufficient level of capacity for each facility. Furthermore, this consideration prohibits facilities from utilizing whole of the established capacity, because, if a demand arrives to a facility which whole of its capacity is occupied by other customers, this demand might be in queue for an infinite time.

Focus of this research is on immobile facilities, where customers-facility interactions happen as the result of customers traveling to facilities to seek service. Moreover, studied time horizon is divided into several equal time periods, where establishment of facilities or any modification of their capacities or the allocated workload are possible only at the beginning of each time period.

2.2 Notation

Decision Variables

$y_{j\acute{k}kt} = \begin{cases} 1\ ;\ \text{if facility } j \text{ holds capacity } k \text{ in period } t \text{ while it has been } \acute{k} \text{ in the previous period} \\ 0\ ;\ \text{otherwise} \end{cases}$

x_{ijt}: proportion of demand of customer i allocation to facility j in time period t

x_{ijkt}: proportion of demand of customer i allocation to facility j holding capacity level k in time period t

z_{jkt}: proportion of utilization of facility j holding capacity level k in time period t

w_{jkt}: average waiting time for each customer allocated to facility j holding capacity level k in time period t

Parameters

$f_{j\acute{k}kt}$: fixed cost for transition of facility j from capacity level $\acute{k}$ to k at the beginning of time period t

c_{ijt}: allocation cost of customer i to facility j in time period t

p_{jkt}: processing cost of facility j with capacity level k in time period t

λ_{it}: demand of customer i in time period t

μ_{jkt}: capacity of facility j equipped with capacity level k in time period t

h_{jt}: holding cost of work-in-process inventory per unit for facility j in time period t

$C^2_{s_{jt}}$: squared coefficient of variance of service times in facility j in time period t

M: a big value, where $\frac{M-1}{M}$ indicates the maximum allowed proportion of utilization in each facility

$\Lambda_{jt} = \sum_{i \in I} \lambda_{it}\ x_{ijt} = \sum_{i \in I} \sum\limits_{k \in K} \lambda_{it}\ x_{ijkt}$: total amount of ordered demands at facility j in time period t

$\mu_{jt} = \sum_{\acute{k} \in K} \sum_{k \in K} \mu_{jkt}\ y_{j\acute{k}kt}$: capacity of facility j in time period t

$\rho_{jt} = \begin{cases} \Lambda_{jt}/\mu_{jt}\ ;\ \text{if facility is open} \\ 0 \qquad\quad ;\ \text{if facility is closed} \end{cases}$: proportion of utilization of facility j in time period t

$E[S_{jt}] = \begin{cases} \frac{1}{\mu_{jt}}\ ;\ \text{if facility is open} \\ 0 \quad\ ;\ \text{if facility is closed} \end{cases}$: mean of service times in facility j in period t

$E[S_{jkt}] = \begin{cases} \frac{1}{\mu_{jkt}}\ ;\ \text{if } k > 0 \\ 0 \quad\ ;\ \text{if } k = 0 \end{cases}$: mean of service times in facility j with capacity level k in period t

$E[W_{jt}]$: mean of sojourn time of a customer allocated to facility j in period t

$E[WIP_{jt}]$: mean of delay/waiting time of a customer allocated to facility j in period t
R_{jt}: average waiting time for each customer allocated to facility j in time period t

2.3 Formulation

We denote by $I = \{1,\ 2,\ 3, ...,\ Number\ of\ Customers\}$ the set of customer demand points and by $J = \{1,\ 2,\ 3, ...,\ Number\ of\ Potential\ Locations\}$ the set of potential facility locations. Also, $T = \{1,\ 2,\ 3, ...,\ Number\ of\ Time\ Periods\}$ stands for the set of time periods with assumption throughout that the end of period t corresponds the beginning of period $t + 1$, while the set of possible capacity levels for each facility is denoted by $K = \{0,\ 1,\ 2, ...,\ Number\ of\ Capacity\ Levels\}$. Clearly, capacity level 0 interprets the closeness of the facility. So, $\mid K \mid = Number\ of\ Capacity\ Levels + 1$.

The demand of customer i in period t is denoted by λ_{it}. The cost to transport one unit from facility j to customer i in period t is c_{ijt}. This term is typically a cost function for allocation costs, based on the distance between customer i and facility j. To allocate a product to a customer, some operations and handling must be executed in the corresponding facility. Thus, moreover of allocation, transportation costs include other handling factors such as assembling or packaging that comprehended in a parameter called processing cost denoted by p_{jkt} for each unit in facility j equipped by capacity level k in period t. The cost matrix $f_{j\acute{k}kt}$ describes the combined cost to change the capacity level of facility j from $\acute{k}$ to k at the beginning of time period t and operating the facility at capacity level k throughout the period. Section 4.1.1 clarifies how this matrix is built. Furthermore, it is assumed that k^j is the existing capacity level of facility j at the beginning the studied horizon.

To formulate the problem, we use binary decision variables $y_{j\acute{k}kt}$ equal to 1 if facility j is equipped with capacity level k in time period t while it has held capacity level $\acute{k}$ in the previous time period. The allocation variables x_{ijkt} are positive continuous decision variables that denote the proportion of the demand of customer i which is served by facility j equipped with capacity level k in period t.

Having assumed that the service times at each facility in each time period are independent and identically distributed random variables that follow a general distribution, each facility can be modeled as an M/G/1 queue. Let $E[S_{jkt}]$, $V[S_{jkt}]$, and $E[S^2_{jkt}]$ denote the mean, variance, and second moment of service times at facility j with capacity level $k > 0$ in time period t, respectively. The mean *service rate* at facility j with capacity level $k > 0$ in time period t is denoted by μ_{jkt}, where $\mu_{jkt} = 1/E[S_{jkt}]$. Clearly, $\mu_{j0t} = E[S_{j0t}] = 0$. If Λ_{jt} denotes the total amount of orders at facility j in period t, then, if facility j is open in period t, its mean service rate is given by $\mu_{jt} = \sum_{\acute{k}\in K} \sum_{k\in K} \mu_{jkt}\ y_{j\acute{k}kt}$

and its mean utilization is $\rho_{jt} = \Lambda_{jt}/\mu_{jt} = \sum_{i\in I}\sum_{k\in K}\lambda_{it}\, x_{ijkt}/\sum_{\acute{k}\in K}\sum_{k\in K}\mu_{jkt}\, y_{j\acute{k}kt}$. Otherwise, $\mu_{jt} = \rho_{jt} = 0$. Thus, the distribution network can be modeled as a network of several independent M/G/1 queues in which the facilities are treated as servers with service rates proportional to their capacity levels, where the capacity levels are discrete (Gross and Harris, 1998). In this context, service rate reflects the amount of orders a facility can process and ship in a given time period. We also assume that there is abundant supply of raw materials/components and their inventory holding costs are insignificant. Under steady state conditions ($\Lambda_{jt} < \mu_{jt}$) and first-come first-serve queuing discipline, the *mean sojourn time* (waiting time in queue + service time) of an order at facility j in period t is given by the *Pollaczek-Khintchine* formula (Gross and Harris, 1998):

$$E[W_{jt}] = \frac{\Lambda_{jt}E[S_{jt}^2]}{2(1-\rho_{jt})} + E[S_{jt}]$$

The average amount of orders in waiting queue at facility j in period t is obtained by $E[WIP_{jt}] = \Lambda_{jt}\ E[W_{jt}]$. In addition, $C_{s_{jt}}^2 = V[S_{jt}]/E[S_{jt}]^2$ and $E[S_{jt}^2] = V[S_j] + E[S_j]^2 = (1 + C_{s_{jt}}^2)E[S_{jt}]^2$. Hence $E[WIP_{jt}]$ can be written as follows;

$$E[WIP_{jt}] = \Lambda_{jt}\ E[W_{jt}] = \frac{\Lambda_{jt}^2\ E[S_{jt}^2]}{2(1-\rho_{jt})} + \Lambda_{jt}\ E[S_{jt}] = \left(\frac{1 + C_{s_{jt}}^2}{2}\right)\frac{\Lambda_{jt}^2}{\mu_{jt}(\mu_{jt}-\Lambda_{jt})} + \frac{\Lambda_{jt}}{\mu_{jt}}$$

$$= \left(\frac{1 + \sum\limits_{\acute{k}\in K}\sum\limits_{k\in K} C_{s_{jkt}}^2 y_{j\acute{k}kt}}{2}\right)\frac{(\sum\limits_{i\in I}\sum\limits_{k\in K}\lambda_{it}x_{ijkt})^2}{\sum\limits_{\acute{k}\in K}\sum\limits_{k\in K}\mu_{jkt}y_{j\acute{k}kt}(\sum\limits_{\acute{k}\in K}\sum\limits_{k\in K}\mu_{jkt}y_{j\acute{k}kt} - \sum\limits_{i\in I}\sum\limits_{k\in K}\lambda_{it}x_{ijkt})}$$
$$+ \frac{\sum\limits_{i\in I}\sum\limits_{k\in K}\lambda_{it}x_{ijkt}}{\sum\limits_{\acute{k}\in K}\sum\limits_{k\in K}\mu_{jkt}y_{j\acute{k}kt}} \tag{2.1}$$

If h_{jt} denotes the holding cost of work-in-process inventory per unit during the time that a customer's order is in process (and/or the loss of goodwill due to delay in order-to-delivery lead time because of congestion) for facility j in period t, then the *total congestion cost* can be expressed as a product of h_{jt} and total expected WIP in the system, $E[WIP_{jt}]$. Given the fixed cost of opening facilities and equipping them with adequate capacity and a variable cost of processing and transportation of finished product from facilities to customers, the system-wide total expected cost can be expressed as;

$$Z(\mathbf{x},\mathbf{y}) = \sum_{j\in J}\sum_{\acute{k}\in K}\sum_{k\in K}\sum_{t\in T} f_{j\acute{k}kt}\, y_{j\acute{k}kt} + \sum_{i\in I}\sum_{j\in J}\sum_{k\in K}\sum_{t\in T}(p_{jkt} + c_{ijt})\lambda_{it}\, x_{ijkt}$$
$$+ \sum_{j\in J}\sum_{t\in T} h_{jt}\ E[WIP_{jt}(\mathbf{x},\mathbf{y})]$$

Given a set of customers with stochastic demand and a set of potential facility locations with multiple capacity levels, the model formulated below simultaneously determines the location of facilities, their capacity levels, and the allocation of customer demands to facilities in order to minimize the sum of fixed location and capacity acquisition cost, transportation cost and congestion cost. The nonlinear MIP (MINLP) formulation is as follows;

minimize Z(x, y)=

$$\sum_{j\in J}\sum_{\acute{k}\in K}\sum_{k\in K}\sum_{t\in T} f_{j\acute{k}kt}\, y_{j\acute{k}kt} + \sum_{i\in I}\sum_{j\in J}\sum_{k\in K}\sum_{t\in T}(p_{jkt} + c_{ijt})\lambda_{it}\, x_{ijkt} + \sum_{j\in J}\sum_{t\in T} h_{jt}\, E[WIP_{jt}(\mathbf{x},\mathbf{y})] \tag{2.2}$$

subject to:

$$\sum_{i\in I}\sum_{k\in K}\lambda_{it}\, x_{ijkt} \le \sum_{\acute{k}\in K}\sum_{k\in K}\mu_{jkt}\, y_{j\acute{k}kt} \quad ; \quad \forall \quad j\in J,\ t\in T \tag{2.3}$$

$$\sum_{k\in K} y_{jk^j k1} = 1 \quad ; \quad \forall \quad j\in J \tag{2.4}$$

$$\sum_{\acute{k}\in K} y_{j\acute{k}k(t-1)} = \sum_{\acute{k}\in K} y_{jk\acute{k}t} \quad ; \quad \forall \quad j\in J,\ k\in K,\ t\in T\backslash\{1\} \tag{2.5}$$

$$\sum_{j\in J}\sum_{k\in K} x_{ijkt} = 1 \quad ; \quad \forall \quad i\in I,\ t\in T \tag{2.6}$$

$$y_{j\acute{k}kt} \in \{0,1\} \quad ; \quad \forall \quad j\in J,\ \acute{k}\in K,\ k\in K,\ t\in T \tag{2.7}$$

$$0 \le x_{ijkt} \le 1 \quad ; \quad \forall \quad i\in I,\ j\in J,\ k\in K,\ t\in T \tag{2.8}$$

Regarding the objective function 2.2, the first, second and the third expressions stand for location, transportation and congestion costs, respectively. As mentioned before, location cost includes fixed location and capacity level acquisition costs, when transportation cost is consist of processing and allocation costs. Constraints (2.3) ensure that the total allocated demand is less than the capacity in each facility in each time period, whereas constraints 2.4 state that each facility must have a capacity at the beginning of the studied horizon. Clearly, the facility holds the artificial capacity level 0 if and only if it is closed throughout the corresponding period. Constraints 2.5 link the capacity levels changes in consecutive time periods. Combination of 2.4 and 2.5 guarantees that exactly one capacity level is selected at a facility in each period. Constraints 2.6 ensure that the demand of each customer is completely satisfied. Constraints 2.7 and 2.8 are binary and non-negativity restrictions on the location and capacity selection variables and allocation variables, receptively.

Removing the congestion cost expression from the objective function 2.2, the formulation would be called *Generalized Modular Capacities (GMC)* formulation which provides stronger LP relaxation than it in other special cases of existing models for a multi-period facility location problem with capacity expansion and reduction or temporary facility closing and reopening (Jena, Cordeau, and Gendron, 2015).

2.3.1 Tightening the Nonlinear MIP Model

Having considered formula of (2.1), it could be ascertained that inequalities 2.3 are never biding. Otherwise, the value of the objective function leads to infinity as the denominator of the first term of 2.1 turns to zero. In other words, having the congestion cost expression in Z(x, y) guarantees that capacity constraints are never violated. Furthermore, it is assumed that steady state conditions of a queuing system are met, which means $\sum_{i\in I}\sum_{k\in K} \lambda_{it}\ x_{ijkt} = \Lambda_{jt} \leq \mu_{jt} = \sum_{\acute{k}\in K}\sum_{k\in K} \mu_{jkt}\ y_{j\acute{k}kt}$. As a result, constraints of the nonlinear MIP model could be written as an uncapacitated FLP. In this case, the LP relaxation of the model would be tighter (Gendron and Crainic, 1994). Therefore, inequalities

$$x_{ijkt} \leq \sum_{\acute{k}\in K}\sum_{k\in K\setminus\{0\}} y_{j\acute{k}kt} \qquad ; \qquad \forall \quad i\in I,\ j\in J,\ k\in K,\ t\in T \tag{2.9}$$

could be substituted with constraints 2.3 to have a tighter LP relaxed nonlinear MIP model with identical MIP feasible area of solutions.

By imposing binary restrictions on x_{ijkt}, the formulation can handle single sourcing requirements that would restrict the assignment of entire demand of a customer to one and only one facility. Although the model presented here explicitly considers just one product (or one family of products), it can be easily extended to handle multiple products (or families of products) by adding an index to the decision variables x_{ijkt} for the different products and modifying the corresponding constraints accordingly. The approximation and solution methods presented in the following sections can also be easily modified to handle the extended model (Vidyarthi, Elhedhli, and Jewkes, 2009).

2.4 Linearization

The nonlinearity in the presented model arises due to the third expression of 2.2 for the total expected WIP at the facilities, $E[WIP_j]$, which is a function of the decision variables corresponding to location and capacity selection ($y_{j\acute{k}kt}$) and allocation (x_{ijkt}).

Thus, a solution procedure is developed based on a simple transformation and piecewise linearization of the nonlinear congestion cost function. Linearization of the graph illustrated in figure 2.1 is introduced by Elhedhli, 2005 and Vidyarthi, Elhedhli, and Jewkes, 2009 contribute it by rewriting the formulation. In this research, a similar linearization method is used as is explained in the following sections.

2.4.1 Auxiliary Variables

Having defined $\rho_{jt} \in [0,1)$ as variables which interpret the proportion of utilization of facility j in period t, it is possible to define nonnegative variables $R_{jt} = \frac{\rho_{jt}}{1-\rho_{jt}}$ that interpret the average of waiting time for each customer allocated to facility j in period t. So, we have;

$$\rho_{jt} = \frac{\Lambda_{jt}}{\mu_{jt}} \qquad \Longrightarrow \qquad R_{jt} = \frac{\rho_{jt}}{1-\rho_{jt}} = \frac{\Lambda_{jt}}{\mu_{jt} - \Lambda_{jt}}$$

Then, to linearize $E[WIP_{jt}]$, the expression 2.1 is rewritten as follows:

$$E[WIP_{jt}] = \left(\frac{1+C^2_{s_{jt}}}{2}\right) \frac{\Lambda^2_{jt}}{\mu_{jt}(\mu_{jt} - \Lambda_{jt})} + \frac{\Lambda_{jt}}{\mu_{jt}}$$
$$= \frac{1}{2}\left\{\left(1+C^2_{s_{jt}}\right)\frac{\Lambda_{jt}}{\mu_{jt}-\Lambda_{jt}} + \left(1-C^2_{s_{jt}}\right)\frac{\Lambda_{jt}}{\mu_{jt}}\right\} = \frac{1}{2}\left\{\left(1+C^2_{s_{jt}}\right)R_{jt} + \left(1-C^2_{s_{jt}}\right)\rho_{jt}\right\}$$

We define nonnegative auxiliary decision variables z_{jkt} and w_{jkt} such that

$$\rho_{jt} = \sum_{\acute{k}\in K} z_{jkt}\, y_{j\acute{k}kt} \qquad \text{and} \qquad R_{jt} = \sum_{\acute{k}\in K} w_{jkt}\, y_{j\acute{k}kt}$$

Since there is only one capacity level k with $\sum_{\acute{k}\in K} y_{j\acute{k}kt} = 1$ while $\sum_{\acute{k}\in K} y_{j\acute{k}\hat{k}t} = 0$ for all other capacity levels $\hat{k} \neq k$, the expression $\rho_{jt} = \sum_{\acute{k}\in K} z_{jkt}\, y_{j\acute{k}kt}$ can be ensured by adding the following;

$$\begin{cases} z_{jkt} \leq \sum_{\acute{k}\in K} y_{j\acute{k}kt} \\ z_{j0t} = 0 \end{cases} \quad ; \quad \forall \quad j \in J,\ k \in K\backslash\{0\},\ t \in T \tag{2.10}$$

while the interpretation of z_{jkt} is the proportion of utilization of facility j with capacity level k in period t. As a consequence, 2.10 prohibits utilization of closed facilities.
Similarly, adding the expression $R_{jt} = \sum_{\acute{k}\in K} w_{jkt}\, y_{j\acute{k}kt}$ could be ensured by adding

$$\begin{cases} w_{jkt} \leq M\sum_{\acute{k}\in K} y_{j\acute{k}kt} \\ w_{j0t} = 0 \end{cases} \quad ; \quad \forall \quad j \in J,\ k \in K\backslash\{0\},\ t \in T \tag{2.11}$$

where the interpretation of w_{jkt} is the average waiting time for customers allocated to facility j with capacity level k in period t and M is the *Big-M*. Consequently, 2.11 prohibits having queue in closed facilities. More details about the *Big-M* and its impact on the model is explained in section 4.5.

As a result, the congestion cost expression can be rewritten as

$$\sum_{j\in J}\sum_{t\in T}\frac{h_{jt}}{2}\left(\sum_{k\in K}\left((1+C^2_{s_{jt}})\,w_{jkt}+(1-C^2_{s_{jt}})\,z_{jkt}\right)\right) \tag{2.12}$$

as the congestion cost of the problem.

Moreover, regarding $\rho_{jt}=\frac{\Lambda_{jt}}{\mu_{jt}}$, we have

$$\rho_{jt}=\frac{\Lambda_{jt}}{\mu_{jt}}=\frac{\sum\limits_{i\in I}\sum\limits_{k\in K}\lambda_{it}\,x_{ijkt}}{\sum\limits_{\acute{k}\in K}\sum\limits_{k\in K}\mu_{jkt}\,y_{j\acute{k}kt}}\Rightarrow\sum_{i\in I}\sum_{k\in K}\lambda_{it}\,x_{ijkt}=\sum_{\acute{k}\in K}\sum_{k\in K}\mu_{jkt}\,\rho_{jt}\,y_{j\acute{k}kt}=\sum_{k\in K}\mu_{jkt}\,z_{jkt}$$

Having $x_{ijt}=\sum\limits_{k\in K}x_{ijkt}$ results in

$$\sum_{i\in I}\lambda_{it}\,x_{ijt}=\sum_{k\in K}\mu_{jkt}\,z_{jkt} \tag{2.13}$$

As $z_{jkt}\leq\sum\limits_{\acute{k}\in K}y_{j\acute{k}kt}$, constraints 2.3 can be dominated and replaced by 2.13 which ensure that aggregation of demands allocated to facility j is exactly equal with workload of that, in time period t.

Consequently, the transportation cost (the second expression of 2.2) can be rewritten as

$$\begin{aligned}\sum_{i\in I}\sum_{j\in J}\sum_{k\in K}\sum_{t\in T}(p_{jkt}+c_{ijt})\lambda_{it}x_{ijkt}&=\sum_{j\in J}\sum_{t\in T}\left(\sum_{i\in I}\sum_{k\in K}p_{jkt}\lambda_{it}x_{ijkt}+\sum_{i\in I}\sum_{k\in K}c_{ijt}\lambda_{it}x_{ijkt}\right)\\&=\sum_{j\in J}\sum_{t\in T}\left(\sum_{i\in I}p_{jkt}\lambda_{it}x_{ijt}+\sum_{i\in I}c_{ijt}\lambda_{it}x_{ijt}\right)\\&=\sum_{j\in J}\sum_{t\in T}\left(\sum_{k\in K}p_{jkt}\mu_{jkt}z_{jkt}+\sum_{i\in I}c_{ijt}\lambda_{it}x_{ijt}\right)\end{aligned} \tag{2.14}$$

that implies transportation costs that are aggregation of processing and allocation costs.

Furthermore, constraints 2.6 can be replaced by

$$\sum_{j \in J} x_{ijt} = 1 \qquad ; \qquad \forall \quad i \in I,\ t \in T \tag{2.15}$$

with the same interpretation.

2.4.2 Piecewise Linearization

As z_{jkt} and w_{jkt} are independent decision variables, expression $w_{jkt} = \frac{z_{jkt}}{1-z_{jkt}}$ would be disregarded while $R_{jt} = \frac{\rho_{jt}}{1-\rho_{jt}}$. Hence, in addition to having auxiliary variables to transform ρ_{jt} and R_{jt}, we have to propose a set of constraints to ensure $w_{jkt} = \frac{z_{jkt}}{1-z_{jkt}}$ at least by an approximation. In this end, we linearize the relation of ρ_{jt} and R_{jt} as explained below. Firstly, as $R_{jt} = \frac{\rho_{jt}}{1-\rho_{jt}}$, we easily obtain the function $\rho_{jt}(R_{jt}) = \frac{R_{jt}}{1-R_{jt}}$.

Proposition 2.1: *The function $\rho_{jt}(R_{jt}) = \frac{R_{jt}}{1+R_{jt}}$ is twice differentiable, continuous, non-decreasing, and concave function of $R_{jt} \in [0, \infty]$.*

Proof:
Differentiating ρ_{jt} with respect to R_{jt}, it is possible to get the first derivative $\frac{\delta \rho_{jt}}{\delta R_{jt}} = \frac{1}{(1+R_{jt})^2} > 0$, and the second derivative $\frac{\delta^2 \rho_{jt}}{\delta R_{jt}^2} = \frac{-2}{(1+R_{jt})^3} < 0$, which proves that the function is concave in R_{jt}.

Let the domain H of the auxiliary variable R_{jt} be a set of indices of points $\{R^h\}_{h \in H}$, at which the function $\rho_{jt}(R_{jt}) = R_{jt}/(1+R_{jt})$ can be approximated arbitrary closely by a set of piecewise linear functions that are tangent to ρ_{jt}. This implies that the function $\rho_{jt}(R_{jt}) = R_{jt}/(1+R_{jt})$ can be expressed as the finite minimum of linearizations of ρ_{jt} at a given set of point $\{R^h\}_{h \in H}$ as follows:

$$\rho_{jt} = \min_{h \in H} \left\{ \frac{1}{(1+R^h)^2} R_{jt} + \frac{(R^h)^2}{(1+R^h)^2} \right\}$$

This can be expressed as the following constraints (Elhedhli, 2005):

$$\rho_{jt} \leq \frac{R_{jt} + (R^h)^2}{(1+R^h)^2} \qquad ; \qquad \forall\, j \in J,\ h \in H,\ t \in T$$

that imply:

$$\sum_{k \in K} z_{jkt} \leq \frac{(\sum_{k \in K} w_{jkt}) + (R^h)^2}{(1+R^h)^2} \qquad ; \qquad \forall\, j \in J,\ h \in H,\ t \in T \tag{2.16}$$

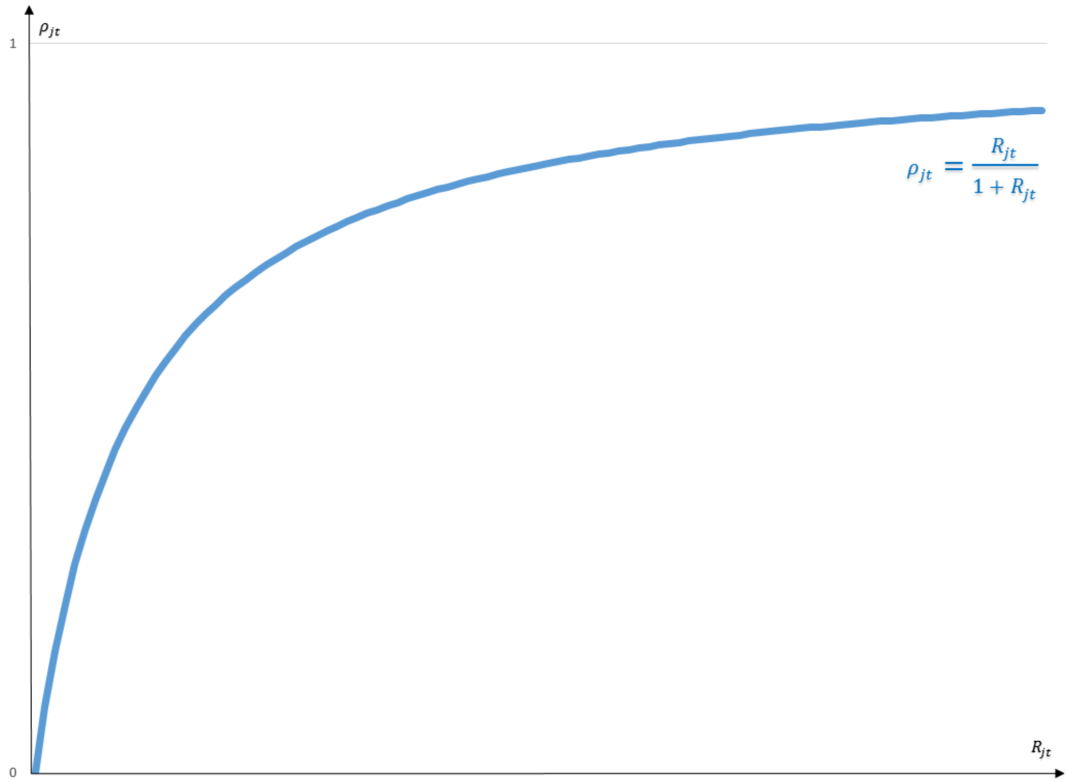


FIGURE 2.1: Nonlinear graph of $\rho_{jt} - R_{jt}$

The number of these constraints are based on the accuracy of linearization. More detail is provided in section 2.4.3.

Therefore, by replacing the congestion cost expression and 2.3 with 2.12 and 2.13, respectively, and adding constraints 2.16 the proposed nonlinear MIP (MINLP) model turns to a piecewise linear MIP model illustrated in the next page;

Linearized Model:

minimize $Z_H(x^H, y^H, z^H, w^H)$=

$$\sum_{j\in J}\sum_{\acute{k}\in K}\sum_{k\in K}\sum_{t\in T} f_{j\acute{k}kt}\, y_{j\acute{k}kt} \qquad (\textit{the 1st part of } 2.2)$$

$$+\sum_{j\in J}\sum_{t\in t}\left(\sum_{k\in K} p_{jkt}\,\mu_{jkt}\,z_{jkt} + \sum_{i\in I} c_{ijt}\,\lambda_{it}\,x_{ijt}\right) \qquad (2.14)$$

$$+\sum_{j\in J}\sum_{t\in T}\frac{h_{jt}}{2}\left(\sum_{k\in K}\left((1+C^2_{s_{jt}})\,w_{jkt} + (1-C^2_{s_{jt}})\,z_{jkt}\right)\right) \qquad (2.12)$$

subject to:

$$\sum_{i\in I}\lambda_{it}\,x_{ijt} = \sum_{k\in K}\mu_{jkt}\,z_{jkt} \quad ; \quad \forall\ j\in J,\ t\in T \qquad (2.13)$$

$$\sum_{k\in K} y_{jk^jk1} = 1 \quad ; \quad \forall\ j\in J \qquad (2.4)$$

$$\sum_{\acute{k}\in K} y_{j\acute{k}k(t-1)} = \sum_{\acute{k}\in K} y_{jk\acute{k}t} \quad ; \quad \forall\ j\in J,\ k\in K,\ t\in T\backslash\{1\} \qquad (2.5)$$

$$\sum_{j\in J} x_{ijt} = 1 \quad ; \quad \forall\ i\in I,\ t\in T \qquad (2.15)$$

$$\sum_{k\in K} z_{jkt} \le \frac{(\sum_{k\in K} w_{jkt}) + (R^h)^2}{(1+R^h)^2} \quad ; \quad \forall\ j\in J,\ h\in H,\ t\in T \qquad (2.16)$$

$$\begin{array}{l} z_{jkt} \le \sum_{\acute{k}\in K} y_{j\acute{k}kt} \\ z_{j0t} = 0 \end{array} \quad ; \quad \forall\ j\in J,\ k\in K\backslash\{0\},\ t\in T \qquad (2.10)$$

$$\begin{array}{l} w_{jkt} \le M\sum_{\acute{k}\in K} y_{j\acute{k}kt} \\ w_{j0t} = 0 \end{array} \quad ; \quad \forall\ j\in J,\ k\in K\backslash\{0\},\ t\in T \qquad (2.11)$$

$$y_{j\acute{k}kt} \in \{0,1\} \quad ; \quad \forall\ j\in J,\ \acute{k}\in K,\ k\in K,\ t\in T \qquad (2.7)$$

$$0 \le x_{ijt} \le 1 \quad ; \quad \forall\ i\in I,\ j\in J,\ t\in T \qquad (2.17)$$

$$0 \le z_{jkt} \le 1 \quad ; \quad \forall\ j\in J,\ k\in K,\ t\in T \qquad (2.18)$$

$$0 \le w_{jkt} \quad ; \quad \forall\ j\in J,\ k\in K,\ t\in T \qquad (2.19)$$

The interpretations of all the expression in the linear model are already explained, in except of 2.17, 2.18 and 2.19 that are nonnegativity constraints for the allocation, utilization and auxiliary variables, respectively.

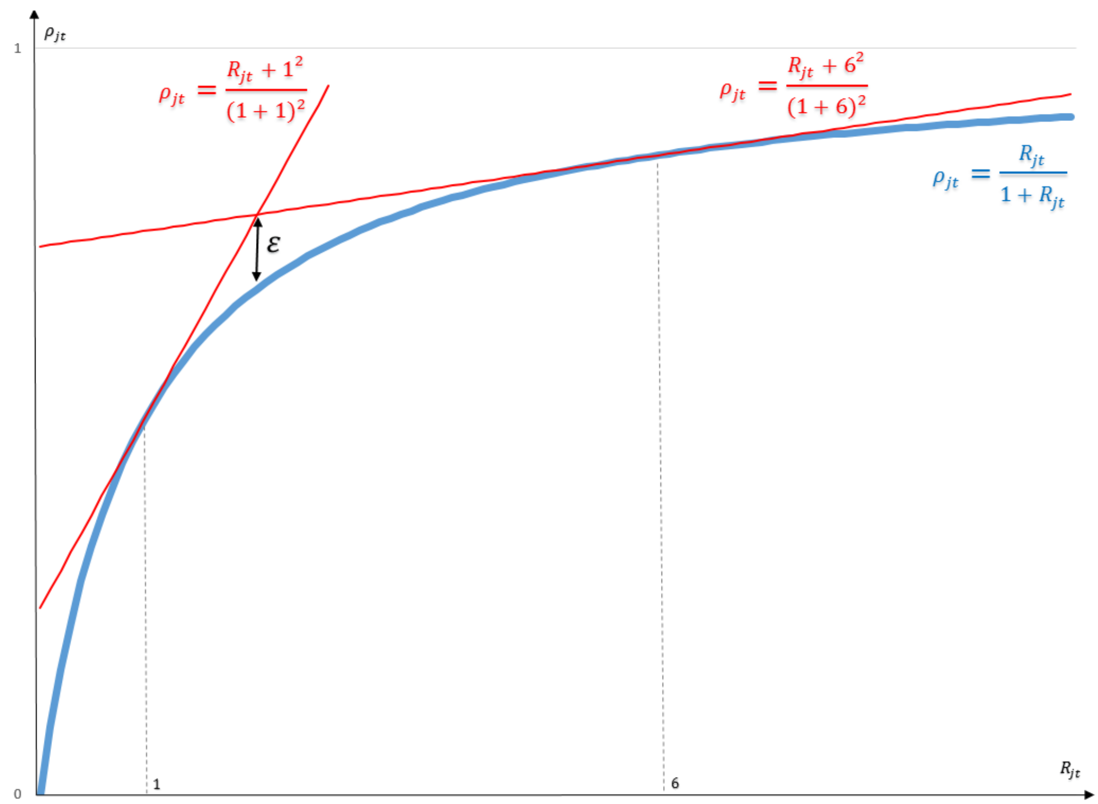


FIGURE 2.2: A sample of several added constraints for linearization

2.4.3 Approximation Accuracy

A priori set of points $\{R^h\}_{h\in H}$ could be generated for the function $\rho_{jt}(R_{jt})$ in such a way that the piecewise linear approximation $\widehat{\rho_{jt}}(R_{jt})$ satisfies

$$0 \leq \widehat{\rho_{jt}}(R_{jt}) - \rho_{jt}(R_{jt}) \leq \varepsilon \tag{2.20}$$

for all $R_{jt} \geq 0$ and $\varepsilon > 0$ (Elhedhli, 2005) that is called *outer linearization*. To know more about this linearization method, it is referred to the appendix A or the original paper, Elhedhli, 2005.

The most critical point is that by setting the value of ε, the numerical value of $\mid H \mid$ is determined, while, the number of constraints 2.16 could be obtained by $\mid J \mid * \mid H \mid * \mid T \mid$. Clearly, small values of the acceptable linearization gap lead to adding high numbers of linearization constraints to the MIP model, and vice-versa.

For example

$$\begin{cases} \mid H \mid = 3 & \text{if} \quad \varepsilon = 10^{-1} \\ \mid H \mid = 10 & \text{if} \quad \varepsilon = 10^{-2} \\ \mid H \mid = 31 & \text{if} \quad \varepsilon = 10^{-3} \\ \mid H \mid = 100 & \text{if} \quad \varepsilon = 10^{-4} \\ \mid H \mid = 316 & \text{if} \quad \varepsilon = 10^{-5} \\ \mid H \mid = 1000 & \text{if} \quad \varepsilon = 10^{-6} \end{cases}$$

Proposition 2.2: *For every subset of points $\{R^h\}_{h\in H}$, $Z_H^*(x^H, y^H, z^H, w^H)$ and $Z(x^H, y^H)$ are a lower bound and an upper bound to Z*(x,y), respectively, where Z*(x,y) is the optimal objective function value of the problem.*

Proof:
For an infinite set of points in H, the feasible region of the linearized model is same as that of the nonlinear problem. Therefore, for a finite set of points in H, the piecewise linearized model is the relaxation of the nonlinear MIP model. As a subsequence, a lower bound on the optimal objective function value is provided by $Z_H^*(x^H, y^H, z^H, w^H)$, where

$$LB^Z = Z_H^*(x^H, y^H, z^H, w^H) = \tag{2.21}$$
$$\sum_{j\in J}\sum_{\acute{k}\in K}\sum_{k\in K}\sum_{t\in T} f_{j\acute{k}kt}\, y^*_{j\acute{k}kt} + \sum_{j\in J}\sum_{t\in t}\left(\sum_{k\in K} p_{jkt}\,\mu_{jkt}\, z^*_{jkt} + \sum_{i\in I} c_{ijt}\,\lambda_{it}\, x^*_{ijt}\right)$$
$$+\sum_{j\in J}\sum_{t\in T}\frac{h_{jt}}{2}\left(\sum_{k\in K}\left((1 + C^2_{s_{jt}})\, w^*_{jkt} + (1 - C^2_{s_{jt}})\, z^*_{jkt}\right)\right)$$

Furthermore, the optimal solution of the linearized model $((x^H)^*, (y^H)^*, (z^H)^*, (w^H)^*)$ is always feasible to the nonlinear one, because, it satisfies all the constraints 2.3-2.8 which are common to both the models. A feasible solution of the nonlinear (original) problem provides an upper bound on its optimal objective function value. Hence, we can get an upper bound on the optimal objective function value by obtaining x^*_{ijkt} values from $x^*_{ijt} = \sum_{k\in K} x^*_{ijkt}$ regarding expressions $x^*_{ijkt} \leq \sum_{\acute{k}\in K} y^*_{j\acute{k}kt}$ and then

$$UB^Z = Z((x^H)^*, (y^H)^*) = \tag{2.22}$$
$$\sum_{j\in J}\sum_{\acute{k}\in K}\sum_{k\in K}\sum_{t\in T} f_{j\acute{k}kt}\, y^*_{j\acute{k}kt} + \sum_{i\in I}\sum_{j\in J}\sum_{k\in K}\sum_{t\in T}(p_{jkt}+c_{ijt})\lambda_{it}\, x^*_{ijkt}$$
$$+\sum_{j\in J}\sum_{t\in T}\frac{h_{jt}}{2}\left\{\left(1+C^2_{s_{jt}}\right)\frac{\sum\limits_{i\in I}\sum\limits_{k\in K}\lambda_{it}x^*_{ijkt}}{\sum\limits_{\acute{k}\in K}\sum\limits_{k\in K}\mu_{jkt}y^*_{j\acute{k}kt}-\sum\limits_{i\in I}\sum\limits_{k\in K}\lambda_{it}x^*_{ijkt}}+\left(1-C^2_{s_{jt}}\right)\frac{\sum\limits_{i\in I}\sum\limits_{k\in K}\lambda_{it}x^*_{ijkt}}{\sum\limits_{\acute{k}\in K}\sum\limits_{k\in K}\mu_{jkt}y^*_{j\acute{k}kt}}\right\}$$

As is clear, the values of location and transportation costs in 2.21 are equal with them in 2.22. However, the value of congestion cost in 2.21 is less than or equal to it in 2.22.

2.5 The Linearized Model Properties

2.5.1 The Artificial Capacity Level

In the presented formulation, capacity level changes are represented by the $y_{j\acute{k}kt}$ variables. For each facility, this transition from one capacity to another can be represented in a graph, where each node represents a capacity level and each arc a capacity transition where the arc cost is $f_{j\acute{k}kt}$.

This interpretation provides some special characters for the location and capacity selection variables. As an example, it leads to $\mid J \mid$ independent shortest path problem by relaxing demand and linearization constraints (2.15 and 2.16) and adding $\mid K \mid$ hypothetical arcs converged to a same hypothetical node in each graph (figure 2.4). Having known that there are various methods to solve a shortest path problem, we have different options to solve a relaxation of the formulation. In other words, having the artificial capacity level 0, that interprets the closeness of the facility, makes such a property for location and capacity selection decisions which is beneficial in Lagrangian Relaxation, Danzig-Wolfe decomposition, Column Generation, etc.

As a result, we are enabled to have various solution methods to take advantage of this property.

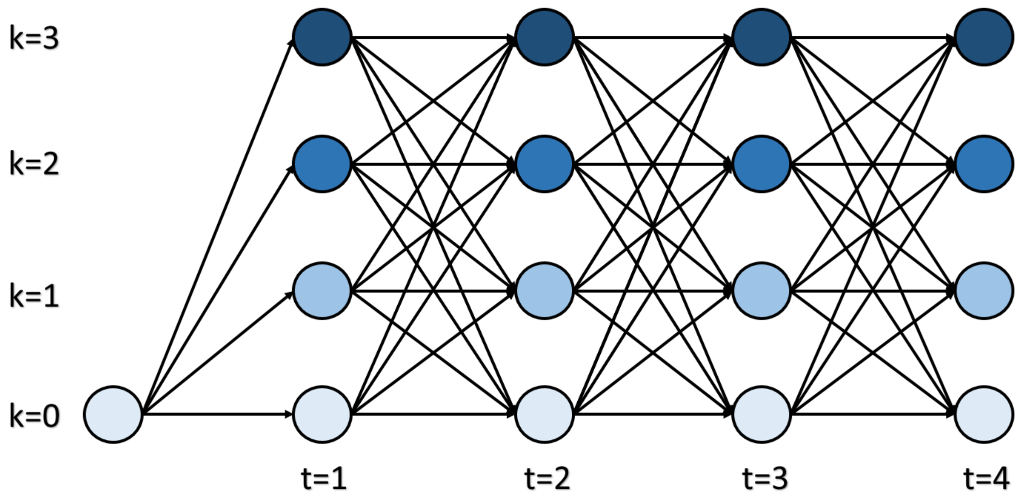


FIGURE 2.3: The graph representing the Location and Capacity Selection variables ($y_{j\acute{k}kt}$) for each facility $j \in J$

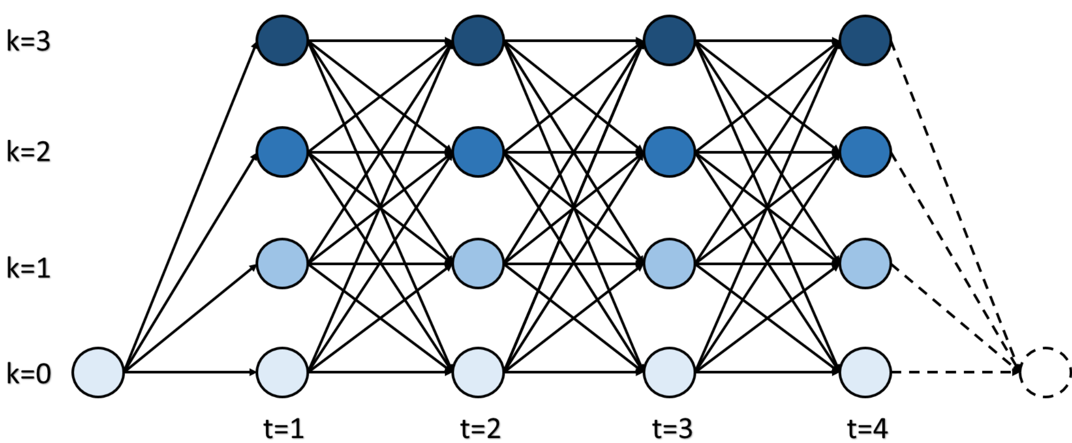


FIGURE 2.4: The graph representing the independent shortest path problems for each facility when constraints 2.15 and 2.16 are relaxed. Value of each are is $f_{j\acute{k}kt}$

2.5.2 Reduction of Allocation Variables Index

In the nonlinear model, allocation variables are x_{ijkt} that are involved in the transportation cost (the second part of 2.2) while $i \in I,\ j \in J,\ k \in K$ and $t \in T$. Due to the fact that the transportation cost includes the both processing and allocation cost and procession costs are not equal in different levels of capacity, allocation variables must specify the level of capacity of facilities as well as the corresponding customers and periods. Beside that, as allocation variables are decision variables, it is not possibles to consider them as $x_{ijt} = \sum_{k \in K} x_{ijkt}$.

Nevertheless, in the piecewise linearized model, expression 2.13 enables us to calculate processing cost by utilization decision variables (z_{jkt}). As a consequence, necessity of demonstrating the corresponding level of capacity is waived for allocation variables. Hence, the index k is removed from allocation variables. Such a reduction in the index of these variables significantly accelerates MIP solver performances to solve the model.

2.5.3 Opening and Re-opening of Facilities

As this research considers a strategic problem with generality in the benchmark, the distinction between opening and reopening a facility is ignored. Nevertheless, opening a facility in an uncivilized location could be more costly than reopening a temporarily closed facility in that location, because the establishment includes activities such as installment of water/gas tubes, electricity, construction and so on. Thus, in the literature, some of the existing models differentiate opening with reopening a facility (such as Dias, Captivo, and Clímaco, 2006).

To adapt our formulation to incorporate this contrast, we have to adapt the location and capacity selection variables as $y_{j\acute{k}kt} = Y_{j\acute{k}kt} + R_{jkt}$ where $Y_{j\acute{k}kt} \in \{0,1\}$ has a same concept as it of $y_{j\acute{k}kt}$ and $R_{jkt} \in \{0,1\}$ stands for reopening a temporarily closed facility j with capacity level k at the beginning of time period t. Also, the location cost would be calculated as $\sum_{j \in J} \sum_{\acute{k} \in K} \sum_{k \in K} \sum_{t \in T} (f_{j\acute{k}kt}\, Y_{j\acute{k}kt} + \hat{f}_{jkt}\, R_{jkt})$ while $\hat{f}_{jkt}$ is the cost of reopening and the following sets of constraints must be added to the model;

$$\sum_{\acute{k} \in K} y_{j\acute{k}0(t-1)} = \sum_{k \in K} Y_{j0kt} + \sum_{k \in K \backslash \{0\}} R_{jkt} \qquad ; \qquad \forall \quad j \in J,\ t \in T \backslash \{1\}$$

that ensure reopening happens only for closed facilities, plus;

$$\begin{aligned} & R_{jkt} \leq \sum_{\acute{k} \in K} \sum_{k \in K \backslash \{0\}} \sum_{\hat{t}=1}^{t-1} y_{j\acute{k}k\hat{t}} \qquad ; \qquad \forall \quad j \in J,\ k \in K,\ t \in T \backslash \{1\} \\ & R_{jk0} = 0 \end{aligned}$$

that state reopening happens if the facility has been open before.

2.6 Tightening Inequalities for the Linearized Model

2.6.1 Strong Inequalities

As is proved by Gendron and Crainic (1994), LP relaxation of an *Uncapacitated Network Design Problem* model is tighter than *Capacitated* form. Consequently, as *Facility Location Problem* is a special case of a *Network Design Problem*, the theory is valid and could be used to tighten the formulation. Thus, having adapted constraints 2.13 with disregarding capacity limitations, we can have constraints such

$$x_{ijt} \leq \sum_{\acute{k} \in K} \sum_{k \in K \setminus \{0\}} y_{j\acute{k}kt} \quad ; \quad \forall \quad i \in I,\ j \in J,\ t \in T \tag{2.23}$$

that imply customers are allocated only to open facilities.

Despite of redundancy for MIP model when constraints 2.13 are existed, having 2.23 could considerably tighten the LP relaxation the formulation and lead to a significant expedition in probing Branch-and-Cut.

In the literature, this set of valid inequalities are mentioned is *Strong Inequalities (SI)* (Jena, Cordeau, and Gendron, 2015). In this formulation, having decision variables such as z_{jkt} enables us to add two similar sets of inequalities as;

$$x_{ijt} \leq \sum_{k \in K} \lceil z_{jkt} \rceil \quad ; \quad \forall \quad i \in I,\ j \in J,\ t \in T \tag{2.24}$$

with a same interpretation as 2.23, and,

$$\sum_{k \in K \setminus \{0\}} y_{j\acute{k}kt} \leq \sum_{k \in K} \lceil z_{jkt} \rceil \quad ; \quad \forall \quad j \in J,\ \acute{k} \in K,\ t \in T \tag{2.25}$$

that means open facilities are enforced to have utilization.

In this paper, *SI* are always added to the model as constraints. However, it is possible to add them in a Branch-and-Cut manner only when they are violated in the solution of LP relaxation. The result of adding SI to the formulation is demonstrated in section 4.2.

2.6.2 Aggregated Demands Constraints

In addition of SI, there is another set of valid inequalities provided for GMC formulation of FLPs as is written below;

$$\sum_{j\in J}\sum_{\acute{k}\in K}\sum_{k\in K}\mu_{jkt}\,y_{j\acute{k}kt}\geq\sum_{i\in I}\lambda_{it}\qquad;\qquad\forall\quad t\in T\tag{2.26}$$

that implies in each time period, established capacity must be greater than or equal to aggregation of demands of all customers.

In the literature, this set of inequalities are called *Aggregated Demands Constraints (ADC)* (Jena, Cordeau, and Gendron, 2015). In this formulation, having z_{jkt} as decision variables moreover of inequalities 2.10 enables us to dominate inequalities 2.26 and replace them with a lifted set of inequalities as *ADC*;

$$\sum_{j\in J}\sum_{k\in K}\mu_{jkt}\,z_{jkt}=\sum_{i\in I}\lambda_{it}\qquad;\qquad\forall\quad t\in T\tag{2.27}$$

that guarantees in each time period, total amount of utilization of all facilities is exactly equal with aggregation of demands of all customers.

In opposite of *SI*, the only way of having *ADC* is adding them directly to the model as constraints, because, *ADC* are redundant even for LP relaxation of the model. However, adding them to the model enables MIP solvers to generate *Cover cuts* that further strengthen the formulation. (Jena, Cordeau, and Gendron, 2015)

Mathematical Proof of ADC Validity:

$$eq\ 2.13:\ \sum_{i\in I}\lambda_{it}\,x_{ijt}=\sum_{k\in K}\mu_{jkt}\,z_{jkt}\quad\Longrightarrow\quad\sum_{j\in J}\sum_{i\in I}\lambda_{it}\,x_{ijt}=\sum_{j\in J}\sum_{k\in K}\mu_{jkt}\,z_{jkt}$$

$$\Longrightarrow\quad\sum_{i\in I}\lambda_{it}\sum_{j\in J}x_{ijt}=\sum_{j\in J}\sum_{k\in K}\mu_{jkt}\,z_{jkt}\quad\xRightarrow[\sum_{j\in J}x_{ijt}=1]{eq\ 2.15:}\quad\sum_{i\in I}\lambda_{it}=\sum_{j\in J}\sum_{k\in K}\mu_{jkt}\,z_{jkt}$$

$$\xRightarrow[z_{jkt}\leq\sum_{\acute{k}\in K}y_{j\acute{k}kt}]{eq\ 2.10:}\quad\sum_{i\in I}\lambda_{it}\leq\sum_{j\in J}\sum_{\acute{k}\in K}\sum_{k\in K}\mu_{jkt}\,y_{j\acute{k}kt}$$

2.7 Lifting Linearization Inequalities

As is mentioned in section 2.4.2, the formulation has inequality 2.16 for each $j\in J,\ h\in H$ and $t\in T$. As a consequence, it is intended to have a *basic mixed integer cut* for each line of this set of inequalities. In this regard, we have;

$$\sum_{k\in K} z_{jkt} \le \frac{(\sum_{k\in K} w_{jkt}) + (R^h)^2}{(1+R^h)^2} \xrightarrow[w_{j0t}=0]{z_{j0t}=0} \sum_{k\in K\setminus\{0\}} z_{jkt} \le \frac{\sum_{k\in K\setminus\{0\}} w_{jkt}}{(1+R^h)^2} + \frac{(R^h)^2}{(1+R^h)^2}$$

$$\Longrightarrow \sum_{k\in K\setminus\{0\}} z_{jkt} - \sum_{\acute{k}\in K}\sum_{k\in K\setminus\{0\}} y_{j\acute{k}kt} \le \frac{\sum_{k\in K\setminus\{0\}} w_{jkt}}{(1+R^h)^2} + \frac{(R^h)^2}{(1+R^h)^2} - \sum_{\acute{k}\in K}\sum_{k\in K\setminus\{0\}} y_{j\acute{k}kt}$$

$$\Longrightarrow \sum_{k\in K\setminus\{0\}} \left(z_{jkt} - \sum_{\acute{k}\in K} y_{j\acute{k}kt} \right) \le \frac{\sum_{k\in K\setminus\{0\}} w_{jkt}}{(1+R^h)^2} + \frac{(R^h)^2}{(1+R^h)^2} - \sum_{\acute{k}\in K}\sum_{k\in K\setminus\{0\}} y_{j\acute{k}kt}$$

$$\Longrightarrow \sum_{\acute{k}\in K}\sum_{k\in K\setminus\{0\}} y_{j\acute{k}kt} \le \sum_{k\in K\setminus\{0\}} \left(\sum_{\acute{k}\in K} y_{j\acute{k}kt} - z_{jkt} \right) + \frac{\sum_{k\in K\setminus\{0\}} w_{jkt}}{(1+R^h)^2} + \frac{(R^h)^2}{(1+R^h)^2}$$

Having considered 2.10, the term $\sum_{k\in K\setminus\{0\}} \left(\sum_{\acute{k}\in K} y_{j\acute{k}kt} - z_{jkt} \right)$ has a positive value.

$$Lets \begin{cases} \eta = \sum_{\acute{k}\in K}\sum_{k\in K\setminus\{0\}} y_{j\acute{k}kt} \ \in Z^+ \\[2ex] u = \sum_{k\in K\setminus\{0\}} (\sum_{\acute{k}\in K} y_{j\acute{k}kt} - z_{jkt}) + \frac{\sum_{k\in K\setminus\{0\}} w_{jkt}}{(1+R^h)^2} \ \ge 0 \\[2ex] b = \frac{(R^h)^2}{(1+R^h)^2} \ \notin Z \end{cases}$$

$$\Longrightarrow \quad \eta \le u + b \quad \xrightarrow{\textit{corollary of proposition 8.6}\text{, Wolsey (1998)}} \quad \eta \le \lfloor b \rfloor + \frac{u}{1 - b + \lfloor b \rfloor}$$

$$\xrightarrow{\lfloor b \rfloor = 0} \quad \eta \le \frac{u}{1-b} \quad \xrightarrow{0 \le b < 1} \quad (1-b)\,\eta \le u$$

$$\Longrightarrow (1 - \frac{(R^h)^2}{(1+R^h)^2}) \sum_{\acute{k}\in K}\sum_{k\in K\setminus\{0\}} y_{j\acute{k}kt} \le \sum_{k\in K\setminus\{0\}} (\sum_{\acute{k}\in K} y_{j\acute{k}kt} - z_{jkt}) + \frac{\sum_{k\in K\setminus\{0\}} w_{jkt}}{(1+R^h)^2}$$

which is equal with

$$\sum_{k\in K\setminus\{0\}} z_{jkt} \le \frac{\sum_{k\in K\setminus\{0\}} w_{jkt}}{(1+R^h)^2} + \left(\frac{(R^h)^2}{(1+R^h)^2} \right) \sum_{\acute{k}\in K}\sum_{k\in K\setminus\{0\}} y_{j\acute{k}kt} \qquad (2.28)$$

that is valid for each $j \in J,\ h \in H$ and $t \in T$.

As inequalities 2.28 are *mixed integer* cuts generated for the model, they tighten LP relaxation of it. Therefore, there are various approaches to incorporate them such as adding them to the model as constraints or adding them in a Branch-and-Cut manner only when they are violated.

Furthermore, having regarded that $z_{j0t} = w_{j0t} = 0$ and $\sum_{\acute{k}\in K}\sum_{k\in K\setminus\{0\}} y_{j\acute{k}kt} \leq 1$, it could be noticed that constraints 2.16 are redundant and dominated by incorporation of inequalities 2.28 into the model. Hence, it is declared as a solution to replace 2.16 with 2.28 as constraints of the model that makes its LP relaxation tighter. However, it remains as a rational solution to have 2.16 and 2.28 in the model, simultaneously, because, involving the dominated set of constraints enables MIP solvers to generate more number of *Mixed Integer Rounding (MIR)* cuts, rather than it when the model includes only the lifted set of inequalities.

Due to the fact that constraints 2.16 are added to the model as a consequence of linearization, it could be claimed that this lifting is a contribution to the linearization method introduced by Elhedhli, 2005.

As a conclusion of section 2.6, the mathematical model written in the page 27 is promoted to the model illustrated in the next page. The new model is called the tightened model and is considered as the input to the MIP solver.

Tightened Model:

minimize $Z_H\left(x^H, y^H, z^H, w^H\right) =$

$$\sum_{j\in J}\sum_{\acute{k}\in K}\sum_{k\in K}\sum_{t\in T} f_{j\acute{k}kt}\, y_{j\acute{k}kt} + \sum_{j\in J}\sum_{t\in t}\left(\sum_{i\in I} c_{ijt}\,\lambda_{it}\,x_{ijt} + \sum_{k\in K} p_{jkt}\,\mu_{jkt}\,z_{jkt}\right)$$
$$+\sum_{j\in J}\sum_{t\in T}\frac{h_{jt}}{2}\left(\sum_{k\in K}\left((1+C^2_{s_{jt}})\,w_{jkt} + (1-C^2_{s_{jt}})\,z_{jkt}\right)\right)$$

subject to:

$$\sum_{i\in I}\lambda_{it}\,x_{ijt} = \sum_{k\in K}\mu_{jkt}\,z_{jkt} \quad ; \quad \forall \quad j\in J,\ t\in T$$

$$\sum_{k\in K} y_{jk^jk1} = 1 \quad ; \quad \forall \quad j\in J$$

$$\sum_{\acute{k}\in K} y_{j\acute{k}k(t-1)} = \sum_{\acute{k}\in K} y_{jk\acute{k}t} \quad ; \quad \forall \quad j\in J,\ k\in K,\ t\in T\backslash\{1\}$$

$$\sum_{j\in J} x_{ijt} = 1 \quad ; \quad \forall \quad i\in I,\ t\in T$$

$$\sum_{k\in K\backslash\{0\}} z_{jkt} \le \frac{\sum_{k\in K\backslash\{0\}} w_{jkt}}{(1+R^h)^2} + \left(\frac{(R^h)^2}{(1+R^h)^2}\right)\sum_{\acute{k}\in K}\sum_{k\in K\backslash\{0\}} y_{j\acute{k}kt} \quad ; \quad \forall \quad j\in J,\ h\in H,\ t\in T$$

$$z_{jkt} \le \textstyle\sum_{\acute{k}\in K} y_{j\acute{k}kt}, \quad z_{j0t} = 0 \quad ; \quad \forall \quad j\in J,\ k\in K\backslash\{0\},\ t\in T$$

$$w_{jkt} \le M \textstyle\sum_{\acute{k}\in K} y_{j\acute{k}kt}, \quad w_{j0t} = 0 \quad ; \quad \forall \quad j\in J,\ k\in K\backslash\{0\},\ t\in T$$

$$x_{ijt} \le \sum_{\acute{k}\in K}\sum_{k\in K\backslash\{0\}} y_{j\acute{k}kt} \quad ; \quad \forall \quad i\in I,\ j\in J,\ t\in T$$

$$\sum_{j\in J}\sum_{k\in K}\mu_{jkt}\,z_{jkt} = \sum_{i\in I}\lambda_{it} \quad ; \quad \forall \quad t\in T$$

$$y_{j\acute{k}kt} \in \{0,1\} \quad ; \quad \forall \quad j\in J,\ \acute{k}\in K,\ k\in K,\ t\in T$$

$$0 \le x_{ijt} \le 1 \quad ; \quad \forall \quad i\in I,\ j\in J,\ t\in T$$

$$0 \le z_{jkt} \le 1 \quad ; \quad \forall \quad j\in J,\ k\in K,\ t\in T$$

$$0 \le w_{jkt} \quad ; \quad \forall \quad j\in J,\ k\in K,\ t\in T$$

3 Solution Methods

This chapter is divided into two sections. In the first section, two methods are presented to solve the piecewise linearized model with exact solutions. Actually, two families of user-cuts are added to the CPLEX pool of cuts to make the branch-and-cut more efficient. The first presented user-cut is an inequality specific for facility location problems, while the second is generic for MIP model adapted to the formulation. As the first inequality could be generated in non-polynomial number, several heuristics are described to separate them. At the end of the chapter, we propose a decent heuristic in sake of obtaining a close-to-exact solution of the original (MINLP) problem in shorter time, where a parameter is defined to control the interaction between speed and quality of the heuristic solutions. This method starts with a less fine linearization to approximate the optimal solution in the initial step. Then, by reducing the model difficulties, more accurate approximations of the model get solved. It is shown, in several cases, this heuristics results in better solutions than those obtained by the exact methods.

As is mentioned in section 2.4.3, the optimal solution of the linearized model provides an approximation of it of the nonlinear one (the original problem). Therefore, $Z_H\left(x^H, y^H, z^H, w^H\right)$ might have a gap less than or equal to ε with $Z\left(x, y\right)$. Moreover, as is shown in sections 2.6 and 2.7, the LP relaxation of the linearized model is tightened to make MIP solvers more efficient to solve the model. Thus, in this chapter, whenever we say the model, it stands for the tightened linearized model which is illustrated in the page 37.

3.1 ϵ-optimally Solution Methods

Although the fact that the linearized model is an approximation of the original problem, in this section, we solve this MIP problem ϵ-optimally to obtain the best LB of the original problem optimal solution.

3.1.1 Mixed-Dicuts

To tighten to LP relaxation of the model, a set of Mixed-Dicut inequalities are proposed to the formulation. This name comes from a paper by Marchand et al., 2002.

Proposition 3.1: Having $S_1 \subseteq I$ and $S_2 \subseteq J$, inequalities

$$\sum_{i\in S_1}\sum_{j\in S_2}\lambda_{it}\, x_{ijt} + \sum_{j\in J\setminus S_2}\sum_{k\in K} min\{\mu_{jkt}\ ,\ \sum_{i\in S_1}\lambda_{it}\} \sum_{\acute{k}\in K} y_{j\acute{k}kt} \geq \sum_{i\in S_1}\lambda_{it} \quad ; \quad \forall\, t\in T \tag{3.1}$$

are valid for $X = \{x^H,\ y^H,\ z^H,\ w^H\}$ denotes the polyhedral set of the feasible solutions of the model.

Mixed-dicut inequalities (3.1) states that in each time period, if the demands of a set of customers are not satisfied by a selection of facilities, other facilities must be capable enough to satisfy them. The validity proof of these inequalities is shown in the appendix B.

As mixed-dicuts are valid for any $S_1 \subseteq I$ and $S_2 \subseteq J$, their number are not polynomial. Therefore, it would not be reasonable to add them directly to the model as constraints. To involve mixed-dicuts, we add them as cutting planes in the branch-and-bound provided by CPLEX, in addition of the cuts that CPLEX adds by its default.

Separation of Mixed-Dicuts

As can be seen in equation 3.1, by selecting different subsets of I and J, there are $2^{|I|+|J|}$ number of mixed-dicuts for each t in T which is a non-polynomial number that could be extremely massive for large-scale instances. Consequently, in each selected node of the branch-and-bound tree, we use a *Cut Callback* to find the most violated mixed-dicuts for each t, where $\bar{x}_{ijt}$ and $\bar{y}_{j\acute{k}kt}$ are constant values (input) which denote the temporary values of x_{ijt} and $y_{j\acute{k}kt}$ in the LP solution.

Two distinct procedures are presented to find the most violated mixed-dicuts which differ in quality and time of the separation procedure.

Exact Separations as Subproblem

By adding a violated Mixed-Dicut as a cut, the LP relaxation of the model gets tighter. As a result, a higher LB is obtained in each node of the CPLEX branch-and-cut tree that leads to finding the optimal solution faster. The most efficient cuts are those with the most violation before separation.

To have the best selection of $S_1 \subseteq I$ and $S_2 \subseteq J$ that leads to separate the most violated Mixed-Dicut, we can define a MIP model and solve it iteratively in each node of the branch-and-cut tree as a subproblem. Defining a_i and b_j as binary decision variables which are equal to 1 if and only if $i \in S_1$ and $j \in S_2$, respectively, the model is as follows;

minimize f^t(a, b)=

$$\sum_{i\in I}\sum_{j\in J} a_i\, b_j\, \lambda_{it}\, \bar{x}_{ijt} + \sum_{j\in J}(1-b_j)\sum_{k\in K}\Psi_k \sum_{\acute{k}\in K}\bar{y}_{j\acute{k}kt} - \sum_{i\in I} a_i\, \lambda_{it} \tag{3.2}$$

subject to:

$$a_i \in \{0,1\} \quad ; \quad \forall \quad i \in I \tag{3.3}$$

$$b_j \in \{0,1\} \quad ; \quad \forall \quad j \in J \tag{3.4}$$

while $\Psi_k = min\{\mu_{jkt}\ ,\ \sum_{i\in S_1}\lambda_{it}\}$. As expression 3.2 is not linear, it cannot be solved by CPLEX. In this end, auxiliary variables ab_{ij} are used to be replaced with expressions $a_i * b_j$, so that

$$a_i + b_j - 1 \le ab_{ij} \le (a_i + b_j)/2$$

In addition, Ψ_k is rewritten as below;

$$\Psi_k = min\{\mu_{jkt}\ ,\ \sum_{i\in S_1} \lambda_{it}\} = \mu_{jkt}(1-\delta_k) + (\sum_{i\in I} a_i\ \lambda_{it})\delta_k = \mu_{jkt} - \delta_k\ \mu_{jkt} + \sum_{i\in I} a\delta_{ik}\ \lambda_{it}$$

when $a_i + \delta_k - 1 \leq a\delta_{ik} \leq (a_i + \delta_k)/2$. As the result, the model could be rewritten as a linear MIP as follows;

minimize f_l^t(a, b)=

$$\sum_{i\in I}\sum_{j\in J} ab_{ij}\ \lambda_{it}\ \bar{x}_{ijt}$$
$$+\sum_{j\in J}\sum_{k\in K}\left(\mu_{jkt} - \delta_k\ \mu_{jkt} + \sum_{i\in I} a\delta_{ik}\ \lambda_{it} - b_j\ \mu_{jkt} + b\delta_{jk}\ \mu_{jkt} - \sum_{i\in I} ba\delta_{jik}\ \lambda_{it}\right)\sum_{\acute{k}\in K}\bar{y}_{j\acute{k}kt}$$
$$-\sum_{i\in I} a_i\ \lambda_{it} \tag{3.5}$$

subject to:

$$\begin{aligned}
&ab_{ij} \geq a_i + b_j - 1 \\ &ab_{ij} \leq (a_i + b_j)/2 && ; && \forall\ \ i\in I,\ j \in J && (3.6)\\
&a\delta_{ik} \geq a_i + \delta_k - 1 \\ &a\delta_{ik} \leq (a_i + \delta_k)/2 && ; && \forall\ \ i\in I,\ k \in K && (3.7)\\
&b\delta_{jk} \geq b_j + \delta_k - 1 \\ &b\delta_{jk} \leq (b_j + \delta_k)/2 && ; && \forall\ \ j\in J,\ k \in K && (3.8)\\
&ba\delta_{jik} \geq b_j + a\delta_{ik} - 1 \\ &ba\delta_{jik} \leq (b_j + a\delta_{ik})/2 && ; && \forall\ \ i\in I,\ j\in J,\ k \in K && (3.9)\\
&a_i \in \{0,1\} && ; && \forall\ \ i\in I && (3.3)\\
&b_j \in \{0,1\} && ; && \forall\ \ j\in J && (3.4)\\
&\delta_k \in \{0,1\} && ; && \forall\ \ k\in K && (3.10)\\
&ab_{ij} \in \{0,1\} && ; && \forall\ \ i\in I,\ j \in J && (3.11)\\
&a\delta_{ik} \in \{0,1\} && ; && \forall\ \ i\in I,\ k \in K && (3.12)\\
&b\delta_{jk} \in \{0,1\} && ; && \forall\ \ j\in J,\ k \in K && (3.13)\\
&ba\delta_{jik} \in \{0,1\} && ; && \forall\ \ i\in I,\ j\in J,\ k \in K && (3.14)
\end{aligned}$$

where 3.6 guarantee that $ab_{ij} = 1$ if and only if $a_i = b_j = 1$. Constraints 3.6-3.9 have similar interpretations. As is stated by 3.3, 3.4 and 3.10-3.14, there are seven sets of binary decision variables in this model. However, there are only a_i and b_j which are meaningful and independent. Values of a_i and b_j determine the sets $S_1 \subseteq I$ and $S_2 \subseteq J$. Other binary variables are auxiliary which are decided by CPLEX. As it is a minimization problem, decision variables δ_k would be set in such a way that set $\Psi_k = min\{\mu_{jkt}\ ,\ \sum_{i \in S_1} \lambda_{it}\}$ correctly.

Solving this model as a subproblem in some selected nodes of the branch-and-cut tree, we separate the most violated Mixed-Dicut for each $t \in T$ in the temporary solution of the LP model. If $f_l^{t*}(a,\ b) \geq 0$, there is no violation of 3.1 in the LP solution.

We also can have the *First Improvement* strategy for separation, which means adding the mixed-dicut once it is observed. To do so, we must change the CPLEX configuration for the subproblem as setting *cutoff* parameter to $-\xi$ (very small negative number) and *absolute gap* to ∞ (infinity). However, we prefer to keep the CPLEX configuration as default for the subproblem to have the *Best Improvement* strategy, which means separating the most violated mixed-dicut.

Efficient Separation Heuristics

As the reason of adding Mixed-Dicuts is acceleration in solving the problem, a rapid separation procedure is desired, while, finding the sets $S_1 \subseteq I$ and $S_2 \subseteq J$ as they correspond the most violation of 3.1 is a NP-hard problem. In this regard, we propose two various heuristics to separate Mixed-Dicuts. Although, the heuristics do not guarantee to obtain the most violated Mixed-Dicut, they find a violated Mixed-Dicut faster than the exact procedure. Thus, in a given time, more number of Mixed-Dicuts could be added to the model.

The separation heuristic which is outlined in algorithm 1 in the page 44 is applied for each $t \in T$. To initialize the algorithm, we put all $i \in I$ in S_1 and all $j \in J$ out of s_2. Then, we search for the largest open facility which is not included in S_2. Having found such a facility, it gets included in S_2 and all the customers whose demands (or a proportion of demands) are allocated to this facility get excluded from S_1. We repeat this operations to observe all the $j \in J$ out of S_2.

In contrast, in algorithm 2 illustrated in the page 2, we put all $i \in I$ out of S_1 and all $j \in J$ in s_2 to initialize the algorithm. Then, we search for the smallest open facility which is included in S_2. Having found such a facility, it gets excluded form S_2 and all the customers whose demands (or a proportion of demands) are not allocated to this facility get included in S_1. We repeat this operations to observe all the $j \in J$ in S_2.

We take advantage of S_{sep} or $\acute{S}_{sep}$ as user-defined parameters to give more variety to the heuristics by playing with their values.

If the algorithm 1 or 2 gets terminated by operating line 24 or $\hat{f}^t \geq 0$, there is no violation found of expression 3.1 for the corresponding $t \in T$. Otherwise, $-\hat{f}^t$ would be the value of the observed violation.

We also can have the *First Improvement* strategy for separation, which means adding the Mixed-Dicut once it is observed. To do so, we must remove lines 8 and 19 in the algorithm 1 or 2. However, we prefer to operate lines 9-27 in a loop which observe all $j \in J$ to have the *Best Improvement* strategy, which means separating the most violated Mixed-Dicut.

In this paper, we set $S_{sep} = \acute{S}_{sep} = 0.5$.

Adding Mixed-Dicuts

Having performed the separation, if no violation found, the algorithm closes the Cut Callback and lets CPLEX decide to pass the node or add its heuristics before that to obtain a new UB. Otherwise, when a violation of 3.1 is found at least for one t in T, the corresponding Mixed-Dicuts is added as purge-able user-cuts to the LP relaxation of the model. In this case, CPLEX could remove user-cuts after processing several nodes. Once user-cuts of each t are added, the node is re-optimized to obtain a new LB and the algorithm searches for the most violated Mixed-Dicuts, again. Thus, in each iteration, at most $\mid T \mid$ user-cuts could be added. These iterations happen in some selected nodes till there is no more violation of 3.1 or the total number of added Mixed-Dicuts are equal to a specific number ($\theta^{9-\sqrt{depth\ of\ the\ node}}$). This is a stopping criteria that is explained in the following.

To add Mixed-Dicuts, we have to stipulate two facts;

- The nodes of the branch-and-cut tree where the Mixed-Dicuts are added.
- The maximum number of added user-cuts in each selected node, as a stopping criteria

Algorithm 1 : Mixed-Dicuts Heuristic Separation 1 (inclusive)

Require: Choose a fractional parameter $0.5 \leq S_{sep} < 1$

initialize $maxViolation = 0$
for all $i \in I$ **do**
    $a_i = 1$
end for
for all $j \in J$ **do**
    $b_j = 0$
end for
while $\sum_{j \in J} b_j < \mid J \mid$ **do**
    $largestOpenfacility = 0$
    for all $j \in J$ **do**
        if $b_j = 0$ and $largestOpenfacility < \sum_{\acute{k} \in K} \sum_{k \in K} \mu_{jkt} \, \bar{y}_{j\acute{k}kt}$ **then**
            $largestOpenfacility \longleftarrow \sum_{\acute{k} \in K} \sum_{k \in K} \mu_{jkt} \, \bar{y}_{j\acute{k}kt}$
            $\hat{j} \longleftarrow j$
        end if
    end for
    if any $\hat{j}$ is found **then**
        $b_{\hat{j}} \longleftarrow 1$
        for all $i \in I$ **do**
            if $\bar{x}_{i\hat{j}t} > S_{sep}$ **then**
                $a_i \longleftarrow 0$
            end if
        end for
    else
        terminate
    end if
    calculate $\Psi_k = min\{\mu_{jkt} \, , \, \sum_{i \in I} a_i \, \lambda_{it}\}$ for each $k \in K$
    calculate $\hat{f}^t = \sum_{i \in I} \sum_{j \in J} a_i \, b_j \, \lambda_{it} \, \bar{x}_{ijt} + \sum_{j \in J} (1 - b_j) \sum_{k \in K} \Psi_k \sum_{\acute{k} \in K} \bar{y}_{j\acute{k}kt} - \sum_{i \in I} a_i \, \lambda_{it}$
    if $\hat{f}^t < maxViolation$ **then**
        $maxViolation \longleftarrow \hat{f}^t$
        save the corresponding S_1 and S_2 as the best selection
    end if
end while

Algorithm 2 : Mixed-Dicuts Heuristic Separation 2 (exclusive)

Require: Choose a fractional parameter $0 < \acute{S}_{sep} \leq 0.5$

initialize $maxViolation = 0$
for all $i \in I$ **do**
    $a_i = 0$
end for
for all $j \in J$ **do**
    $b_j = 1$
end for
while $\sum_{j \in J} b_j > 0$ **do**
    $smallestOpenfacility = \infty$
    for all $j \in J$ **do**
        if $b_j = 1$ and $smallestOpenfacility > \sum_{\acute{k} \in K} \sum_{k \in K \setminus \{0\}} \mu_{jkt}\, \bar{y}_{j\acute{k}kt}$ **then**
            $smallestOpenfacility \longleftarrow \sum_{\acute{k} \in K} \sum_{k \in K \setminus \{0\}} \mu_{jkt}\, \bar{y}_{j\acute{k}kt}$
            $\hat{j} \longleftarrow j$
        end if
    end for
    if any $\hat{j}$ is found **then**
        $b_{\hat{j}} \longleftarrow 0$
        for all $i \in I$ **do**
            if $\bar{x}_{i\hat{j}t} < S_{sep}$ **then**
                $a_i \longleftarrow 1$
            end if
        end for
    else
        terminate
    end if
    calculate $\Psi_k = min\{\mu_{jkt}\ ,\ \sum_{i \in I} a_i\, \lambda_{it}\}$ for each $k \in K$
    calculate $\hat{f}^t = \sum\limits_{i \in I} \sum\limits_{j \in J} a_i\, b_j\, \lambda_{it}\, \bar{x}_{ijt} + \sum\limits_{j \in J} (1 - b_j) \sum\limits_{k \in K} \Psi_k \sum\limits_{\acute{k} \in K} \bar{y}_{j\acute{k}kt} - \sum\limits_{i \in I} a_i\, \lambda_{it}$
    if $\hat{f}^t < maxViolation$ **then**
        $maxViolation \longleftarrow \hat{f}^t$
        save the corresponding S_1 and S_2 as the best selection
    end if
end while

To specify these facts, we use a dynamic parameter θ while Mixed-Dicuts separations happen in the nodes with depth 0(root), 1, θ, θ^2, θ^3, ... , θ^8 and the maximum numbers of added user-cuts in these nodes are θ^9, θ^8, θ^7, ... , θ, 1.
During tuning nodes of the CPLEX branch-and-cut tree, the algorithm 3 outlined in the page 46 is performed inside a Cut Callback once CPLEX takes a new node to tune, where *nDepth* and *fSeparation* denote the depth of the node and the objective function value of the Separation, respectively. The reason of having such an algorithm is that adding Mixed-Dicuts has the most efficiency when they are added in the initial nodes of the branch-and-cut tree, especially in the root.

Algorithm 3 : Adding Mixed-Dicuts during probing nodes

```
Require: Choose an integer parameter θ > 1
initialize setParam = 0
while setParam < 9 do
  if nDepth=0 then
    nGenerated=0 ; voilationFound=1
    while nGenerated<θ^9  and  violationFound=1 do
      do Mixed-Dicut Separation
      if fSeparation<0 then
        add the violated Mixed-Dicuts
        nGenerated=nGenerated+number of added cuts
      else
        voilationFound=0
      end if
    end while
  else if nDepth=θ^setParam then
    nGenerated=0 ; voilationFound=1
    while nGenerated<θ^(8−setParam)  and  violationFound=1 do
      do Mixed-Dicut Separation
      if fSeparation<0 then
        add the violated Mixed-Dicuts
        nGenerated=nGenerated+number of added cuts
      else
        voilationFound=0
      end if
    end while
  else if nDepth>θ^setParam then
    setParam = setParam + 1
  end if
end while
```

In this paper, we set $\theta = 2$, thus, we investigate for the most violated Mixed-Dicuts in the nodes of the branch-and-cut tree with the following depths;

$$\begin{cases} \text{depth} = 0 \text{ (root)} & ; \quad \text{maximum number of user-cuts in each node is} \quad 512 \\ \text{depth} = 1 & ; \quad \text{maximum number of user-cuts in each node is} \quad 256 \\ \text{depth} = 2 & ; \quad \text{maximum number of user-cuts in each node is} \quad 128 \\ \text{depth} = 4 & ; \quad \text{maximum number of user-cuts in each node is} \quad 64 \\ \text{depth} = 8 & ; \quad \text{maximum number of user-cuts in each node is} \quad 32 \\ \dots & \\ \text{depth} = 256 & ; \quad \text{maximum number of user-cuts in each node is} \quad 1 \end{cases}$$

3.1.2 Enhanced-MIR Cuts

In this section, more valid inequalities are proposed to enrich the CPLEX branch-and-cut. These valid inequalities are called *enhanced-MIR* in the literature (Bodur and Luedtke, 2016).

Proposition 3.2: Inequalities

$$\sum_{\acute{k}\in K} y_{j\acute{k}kt} - z_{jkt} + f(\alpha-\beta)\lceil\alpha\rceil \sum_{\hat{k}\in K} y_{\hat{j}k\hat{j}\hat{k}1} \geq \alpha + f(\beta-\alpha)\left(\lceil\beta-\alpha\rceil - \lceil\beta\rceil \sum_{\check{k}\in K} y_{\check{j}k\check{j}\check{k}1}\right) \tag{3.15}$$

$$\sum_{\acute{k}\in K} y_{j\acute{k}kt} - z_{jkt} + \alpha + f(\alpha-\beta)\lfloor\beta-\alpha\rfloor \geq f(\beta-\alpha)\lfloor\alpha\rfloor \sum_{\hat{k}\in K} y_{\hat{j}k\hat{j}\hat{k}1} + f(\alpha-\beta)\lfloor\beta\rfloor \sum_{\check{k}\in K} y_{\check{j}k\check{j}\check{k}1} \tag{3.16}$$

$$M\sum_{\acute{k}\in K} y_{j\acute{k}kt} - w_{jkt} + \alpha + f(\alpha-\beta)\lfloor\beta-\alpha\rfloor \geq f(\beta-\alpha)\lfloor\alpha\rfloor \sum_{\hat{k}\in K} y_{\hat{j}\hat{j}\hat{k}1} + f(\alpha-\beta)\lfloor\beta\rfloor \sum_{\check{k}\in K} y_{\check{j}k\check{j}\check{k}1} \tag{3.17}$$

$$M\sum_{\acute{k}\in K} y_{j\acute{k}kt} - w_{jkt} + f(\alpha-\beta)\lceil\alpha\rceil \sum_{\hat{k}\in K} y_{\hat{j}k\hat{j}\hat{k}1} \geq \alpha + f(\beta-\alpha)\left(\lceil\beta-\alpha\rceil - \lceil\beta\rceil \sum_{\check{k}\in K} y_{\check{j}k\check{j}\check{k}1}\right) \tag{3.18}$$

are valid for the piecewise linearized formulation for all $j, \hat{j}, \check{j} \in J$, $k \in K$ and $t \in T$ where $f(\alpha-\beta) = \alpha - \beta - \lfloor\alpha-\beta\rfloor > 0$, $f(\beta-\alpha) = \beta - \alpha - \lfloor\beta-\alpha\rfloor > 0$ and α and β are positive continuous values.

Separation of Enhanced-MIR Cuts

Validity of inequalities 3.15-3.17 is proved in appendix B. They could be added to the CPLEX branch-and-cut to tighten the LP relaxation of the formulation by cutting of integrally relaxed polyhedron. To generate enhanced-MIR cuts, values of α and β must be determined. Numerical values of Location and Capacity Selection variables and Utilization variables are given in each node as temporary values of the LP solution. In sake of the most violated enhanced-MIR cut in each node, several subproblems are solved to decide about the values of α and β . For example, to separate cuts 3.15, we solve the following model;

minimize $f_{MIR_1}(\alpha, \beta) =$

$$\sum_{\acute{k} \in K} \bar{y}_{j\acute{k}kt} - \bar{z}_{jkt} + f(\alpha - \beta)\lceil \alpha \rceil \sum_{\hat{k} \in K} \bar{y}_{\hat{j}k\hat{j}\hat{k}1} - \alpha - f(\beta - \alpha)\left(\lceil \beta - \alpha \rceil - \lceil \beta \rceil \sum_{\check{k} \in K} \bar{y}_{\check{j}k\check{j}\check{k}1} \right)$$

subject to:

$\alpha > 0$

$\beta > 0$

Solving it for each $j, \hat{j}, \check{j}$ in J, k in K and t in T, we find the most violation of inequality 3.15. In other words, this minimization problem is solved $\mid J \mid^3 * \mid K \mid * \mid T \mid$ times as a subproblem by CPLEX to separate inequalities 3.15. As it seems time-demanding, we separate these cuts only in the root (node 0) of the branch-and-cut tree to keep tuning fast after branching. Due to nonlinearity, it is not possible to solve this subproblem exactly by CPLEX. Therefore, the best values of α and β are estimated by some heuristics. However, no violation of inequalities 3.15-3.17 is found by the heuristics.

Adding Enhanced-MIR Cuts

After solving the LP relaxation in the root, if the lowest record of f_{MIR_1} be negative, which interprets the violation, we add the corresponding inequality as a purge-able cut to the root node and re-optimize the LP relaxation. Algorithm 4 illustrated in the page 49 demonstrates our method to separate enhanced-MIR cuts.

Algorithm 4 : Separating Enhanced-MIR Cuts in the Root

for all inequalities 3.15-3.17 **do**
  $ViolationFound \longleftarrow 1$
  while $ViolationFound = 1$ **do**
    $MostViolation \longleftarrow 0$
    for all $\hat{j} \in J$ **do**
      for all $\breve{j} \in J$ **do**
        for all $j \in J$ **do**
          for all $k \in K$ **do**
            for all $t \in T$ **do**
              solve the corresponding subproblem
              if $f_{MIR} < MostViolation$ **then**
                $MostViolation \longleftarrow f_{MIR}$
                save the cut with the obtained α^* and β^* and the corresponding $j, \hat{j}, \breve{j}, k$ and t
              end if
            end for
          end for
        end for
      end for
    end for
    if $MostViolation < 0$ **then**
      add the saved cut
    else
      $ViolationFound \longleftarrow 0$
    end if
  end while
end for

3.2 Heuristic Methods to Solve the MIP

As is shown in the table 4.9, large scaled test problems are not necessarily solvable in the time limitation. Thus, we have to capture the incumbent UB as the best known solution (if it exists). In this regard, several heuristics are presented in this paper to obtain a close-to-optimal or at least a feasible solution, when it is not possible by the exact methods.

Due to the objective of these heuristics that is a fast approximation, the linearization gap (ε) is set to 10^{-1} in advance of execution. To implement the heuristics, we let CPLEX to tune the root node of the branch-and-cut tree. By the way, most of the CPLEX cuts are added and its heuristics are applied at least for several time. Subsequently, an acceptable LB is earned. Hence, we execute our heuristics right before tuning node 1 and then, CPLEX pursues its regular tuning as is specified in the configuration. The tuning is accelerated after the root node, because, the heuristics fix several variables or tighten them adding some global cuts explained in the following;

- **Fixing close-to-one binary variables to 1:**

Integrity of binary variables is relaxed in the LP relaxation. However, if the value of a binary variable in the LP relaxation solution is close to 1, its integer value is likely to be 1 in the MIP solution. We fix such a binary variable in advance of being decided by CPLEX by adding a global cut to the branch-and-cut.

$$\text{if} \quad \bar{y}_{j\acute{k}kt} > 1 - \omega_1 \qquad \text{then add} \quad y_{j\acute{k}kt} = 1 \quad ; \quad \forall\, j \in J,\ \acute{k} \in K,\ k \in K,\ t \in T \tag{3.19}$$

where $\bar{y}_{j\acute{k}kt}$ is the value of $y_{j\acute{k}kt}$ in the LP relaxation solution after tuning the root node and $0 < \omega_1 \leq 0.5$ is a constant value.

- **Fixing close-to-zero binary variables to 0:**

Similarly, we add;

$$\text{if} \quad \bar{y}_{j\acute{k}kt} < \omega_1 \qquad \text{then add} \quad y_{j\acute{k}kt} = 0 \quad ; \quad \forall\, j \in J,\ \acute{k} \in K,\ k \in K,\ t \in T \tag{3.20}$$

- **Fixing several Allocation variables:**

As the same as many benchmarks for a Facility Location problem, in our test problems, each candidate facility location is a demand point of a customer as well. In other words, the location of facilities are the location of the first $\mid J \mid$ number of $i \in I$ (more details about test problems is provided in section 4.1). Consequently, it is observed experimentally that in the optimal solution, whole demand of customer j is allocated to facility j if and only if facility j is open. Regarding the fact, such global cuts are proposed;

$$\text{if} \quad \bar{y}_{j\acute{k}kt} > 1 - \omega_1 \qquad \text{then add} \quad x_{jjt} = 1 \quad ; \quad \forall\, j \in J,\ \acute{k} \in K,\ k \in K \backslash \{0\},\ t \in T \tag{3.21}$$

- **Tightening Utilization variables:**

As is shown in section 4.6, utilization figures have a relatively uniform trend in open facilities in the optimal solutions. Having such an inspection, we estimate an UB for z_{jkt} after finding their values in the LP relaxation solution and picking the maximum of them. The following global cuts are added after tuning the root node.

$$z_{jkt} \leq ((1 + \omega_2) * max\{\bar{z}_{jkt}\}) * \sum_{\acute{k} \in K} y_{j\acute{k}kt} \quad ; \quad \forall\, j \in J,\ k \in K \backslash \{0\},\ t \in T \tag{3.22}$$

where $max\{\bar{z}_{jkt}\}$ is the largest z_{jkt} in all $j \in J,\ k \in K,\ t \in T$ and $0 \leq \omega_2 \leq 0.5$ is a constant value.

- **Tightening auxiliary variables:**

Due to the fact that the value of w_{jkt} is correlated with it of z_{jkt}, we have a set of similar cuts as is demonstrated below;

$$w_{jkt} \leq ((1 + \omega_2) * max\{\bar{w}_{jkt}\}) * \sum_{\acute{k} \in K} y_{j\acute{k}kt} \quad ; \quad \forall\, j \in J,\ k \in K \backslash \{0\},\ t \in T \tag{3.23}$$

where $max\{\bar{w}_{jkt}\}$ is the largest w_{jkt} in all $j \in J,\ k \in K$ and $t \in T$.

As mentioned before, these heuristics are applied after that the root node is tuned by CPLEX. Nonetheless, in very large scaled test problems, CPLEX may be unable to tune the root node within the time limitation. Consequently, no heuristic is applied. Thus, to activate heuristics, the algorithm considers a time limitation specifically for the root node which is 75% of the time limitation given to CPLEX to solve the problem.

In this paper, time limitation to solve the linearized model in 2 hours. So, if any of the heuristics 3.19-3.23 are applied, tuning the root node also has a time limitation which is 90 minutes. It means that if the root node is not tuned within 90 minutes by CPLEX, the algorithm forces CPLEX to pass the root with any LB/UB taken by the time. In the following, heuristics cuts are added as global cuts and then, tuning nodes is resumed for up to 30 minutes.

3.2.1 MIP Heuristics Aggressiveness

To be in control of the aggressiveness of heuristics 3.19-3.23, we simply define a parameter $\Omega = 100\,\omega_1 + 10\,(0.5 - \omega_2)$. As is clear Ω has positive and negative correlation with ω_1 and ω_2, respectively, while ω_1 and ω_2 are particularized independently to apply the heuristics. Obviously, large values of Ω lead to quickness in finding the solution. However, they increase the probability of missing the optimal solution. $\Omega = 0$ means no usage of heuristics 3.19-3.23 and lack of time limitation to tune the root node.

In this thesis, to solve the approximation (the linearized model), we set $\omega_1 = 0.03$ and $\omega_2 = 0.2$. Therefore, the following global cuts are added in the node 1 of the branch-and-cut tree.

$$\begin{cases} \text{if} \quad \bar{y}_{j\acute{k}kt} > 0.97 \quad \text{then add} \quad y_{j\acute{k}kt} = 1 & ; \quad \forall\, j \in J,\ \acute{k} \in K,\ k \in K,\ t \in T \\ \text{if} \quad \bar{y}_{j\acute{k}kt} < 0.03 \quad \text{then add} \quad y_{j\acute{k}kt} = 0 & ; \quad \forall\, j \in J,\ \acute{k} \in K,\ k \in K,\ t \in T \\ \text{if} \quad \bar{y}_{j\acute{k}kt} > 0.97 \quad \text{then add} \quad x_{ijt} = 1 & ; \quad \forall\, j \in J,\ \acute{k} \in K,\ k \in K\backslash\{0\},\ t \in T \\ z_{jkt} \leq (1.2 * max\{\bar{z}_{jkt}\}) * \sum_{\acute{k} \in K} y_{j\acute{k}kt} & ; \quad \forall\, j \in J,\ k \in K\backslash\{0\},\ t \in T \\ w_{jkt} \leq (1.2 * max\{\bar{w}_{jkt}\}) * \sum_{\acute{k} \in K} y_{j\acute{k}kt} & ; \quad \forall\, j \in J,\ k \in K\backslash\{0\},\ t \in T \end{cases}$$

where $\Omega = 6$.

3.3 Solution Methods for the Original Problem

In conclusion of chapter 2 and section 3.1, we found the best known setting to solve the linearized model. Regarding the *proposition 2.2* mentioned in the page 29, we know that solving the MIP, provides a lower bound (LB) and an upper bound (UB) for the mixed-integer nonlinear program (MINLP) model.

In other words, in section 3.1 of this paper, the aim is to find the MINLP LB as quick as possible, while, in this section 3.3, the focus is on obtaining an appropriate UB for the MINLP illustrated in the page **??**. In this end, an iterative solution method is presented to solve the original problem. In each iteration, the linearization is done as is explained in the section 2.4. In this method, the algorithm commences with a less accurate linearization, then, it increases the accuracy (decreases the linearization gap, ε) gradually within consecutive iterations. Ideally, once the linearization gap reaches to zero (or an acceptable small value), the optimal solution of the original problem is acquired. The interesting point is that by running the first iteration, the location decision (or capacity selection decision either) is fixed. Exploitation of this point, in addition of some other considerable advantages of starting with an amenable linearization leads to an efficient solution method which is capable of finding exact or approximate solutions of the problem rapidly.

3.3.1 Iterative Solution Method

In this method, firstly, the algorithm does linearization as explained in the section 2.4.2 with $\varepsilon = 10^{-1}$, regardless of the acceptable gap value. The resulting $\rho_{jt} - R_{jt}$ graphs are illustrated in figure 3.1. Clearly, solving the MIP with this linearization would be easier for the solver, because, as is explained in section 2.4.3, in this case, $\mid H \mid = 3$ where $H = \{0\ ,\ 1.14\ ,\ 6.41\}$. As a consequence, the number of constraints 2.16 is reduced to $3 * \mid J \mid * \mid T \mid$. The purpose is to pass the first iteration quickly enough. Then, the linearization gap is decreased to $\varepsilon = 10^{-2}$ and the algorithm starts the second iteration. In this step, the linearized MIP model is solved with the new linearization. Nevertheless, this time, we have critical information about the optimal solution which is acquired from the solution in the previous iteration and is delineated in the upcoming sections. Similarly, in each of the following iterations, the algorithm does linearization with $\varepsilon = 10^{-Iteration\ Number}$. It stops when the linearized model is solved with the acceptable linearization gap. For example, if the acceptable linearization gap is 10^{-3} or 10^{-6}, the algorithm iterations number would be 3 or 6, respectively. Ideally, if the number of iterations goes to infinity, the linearization gap is absolute 0. Thus, the nonlinear model corresponding to the original problem would be solved, exactly.

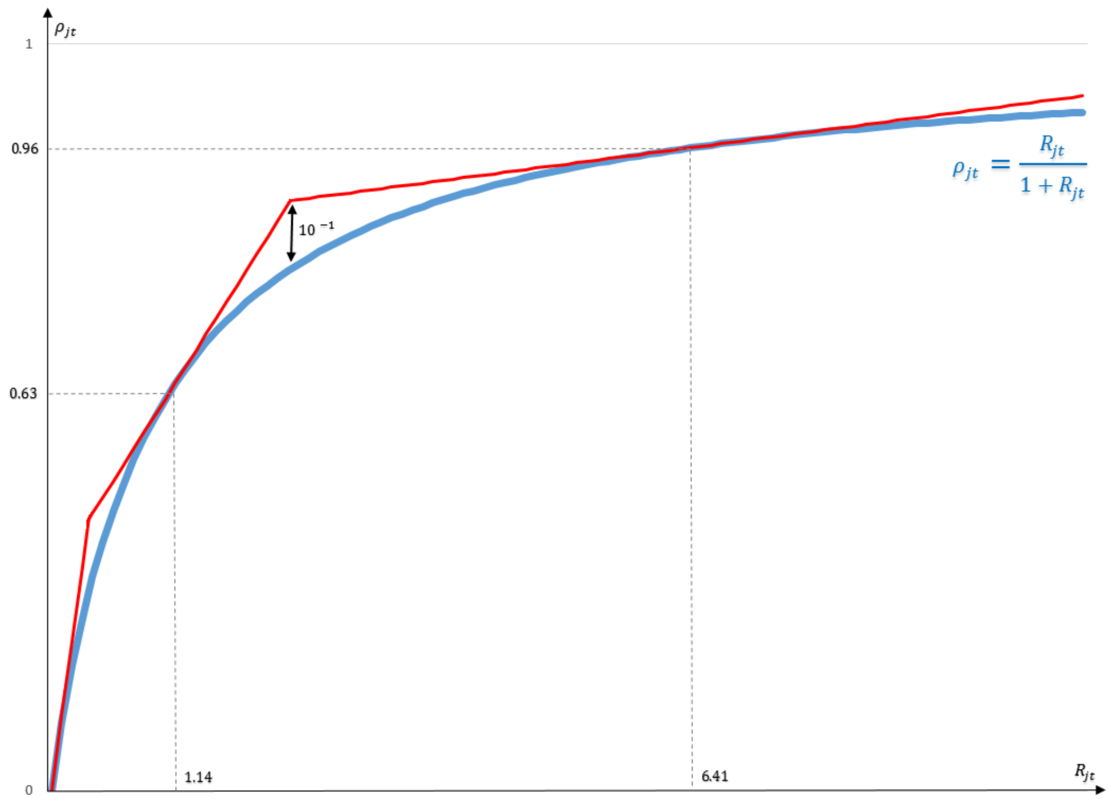


FIGURE 3.1: Linearization of $\rho_{jt} - R_{jt}$ graphs with $\varepsilon = 10^{-1}$

In this paper, we consider linearization gap $\varepsilon = 10^{-6}$ as the exact solution, while the default acceptable absolute MIP gap of CPLEX is 10^{-6} as well.

Tightening Utilization Variables

Having solved the linearized model with any linearization gap, an interval for each R^*_{jt} could be specified which is between the two closest linearization points (R^h). Similarly, an interval for each ρ^*_{jt} could be defined while $\rho(R^h) \leq \rho^*_{jt} < \rho(R^{h+1})$ that actually implies $\frac{R^h}{1+R^h} \leq \rho^*_{jt} < \frac{R^{h+1}}{1+R^{h+1}}$. Thus, as we have $\rho_{jt} = \sum_{\acute{k}\in K} z_{jkt}\, y_{j\acute{k}kt}$ and $R_{jt} = \sum_{\acute{k}\in K} w_{jkt}\, y_{j\acute{k}kt}$, if solving the linearized model in the first iteration results in $R^h \leq \sum_{k\in K} w^*_{jkt} < R^{h+1}$ for any $h \in H$, we set $LB_{R_{jt}} = R^h$ and $UB_{R_{jt}} = R^{h+1}$ as lower bound and upper bound of R_{jt} , respectively. Then, we can have the cutting planes written below in the following iterations for each $j \in J$ and $t \in T$.

$$\sum_{k\in K} w_{jkt} \geq LB_{R_{jt}} \quad ; \quad \forall \quad j \in J,\ t \in T \tag{3.24}$$

$$\sum_{k\in K} z_{jkt} \geq \frac{LB_{R_{jt}}}{1 + LB_{R_{jt}}} \quad ; \quad \forall \quad j \in J,\ t \in T \tag{3.25}$$

$$\sum_{k\in K} w_{jkt} < UB_{R_{jt}} \quad ; \quad \forall \quad j \in J,\ t \in T \tag{3.26}$$

$$\sum_{k\in K} z_{jkt} < \frac{UB_{R_{jt}}}{1 + UB_{R_{jt}}} \quad ; \quad \forall \quad j \in J,\ t \in T \tag{3.27}$$

The algorithm adds these inequalities as constraints to the linearized model.

Therefore, at the end of each iteration, the linearized MIP model formulation gets considerably tighter.

Fixing the Location Decision

As is illustrated in the figure 3.1, linearization with $\varepsilon = 10^{-1}$ partitions possible values of *utilization variable* (z_{jkt}) into three ranges; $[0\ ,\ 0.63)$, $[0.63\ ,\ 0.96)$ and $[0.96\ ,\ 1)$. Therefore, at the end of the first iteration, a closed facility is not differentiated with an open facility with utilization less than $\%63$. To exploit the advantages of such a partitioning as much as possible, we consider an additional tangent point in the $\rho_{jt} - R_{jt}$ graphs. Within the first iteration of the algorithm, we add $R^{add} = \vartheta$ to the set H where ϑ is very small positive value. Clearly,

$\mid H \mid = 4$ when $H = \{0\ ,\ \vartheta\ ,\ 1.14\ ,\ 6.41\}$. As a consequence, two new partitions $[0\ ,\ \frac{\vartheta}{1+\vartheta})$ and $[\frac{\vartheta}{1+\vartheta}\ ,\ 0.63)$ would be replaced with the partition $[0\ ,\ 0.63)$. As $\frac{\vartheta}{1+\vartheta}$ is very small positive number, $LB_{R_{jt}} = 0$ if and only if $\sum_{k \in K} w^*_{jkt} = 0$ which interprets closeness of a facility. In better words, finding $LB_{R_{jt}}$ for each $j \in J$ and $t \in T$ enables us to not only adding inequalities 3.24-3.27 withing the next iteration, but also implicitly specifying openness or closeness of each facility in each time period for all the following algorithm iterations. Thus, having solved the linearized model in the first iteration resulted in $LB_{R_{jt}} = 0$ or $LB_{R_{jt}} > 0$, we can have cutting planes written below within all the following iterations.

$$\begin{cases} \text{if} \quad LB_{R_{jt}} = 0 \quad ; \quad \sum_{\acute{k} \in K} \sum_{k \in K \setminus \{0\}} y_{j\acute{k}kt} = 0 \\ \text{if} \quad LB_{R_{jt}} > 0 \quad ; \quad \sum_{\acute{k} \in K} \sum_{k \in K \setminus \{0\}} y_{j\acute{k}kt} = 1 \end{cases} \quad ; \qquad \forall \quad j \in J,\ t \in T \tag{3.28}$$

To add these cuts more efficiently, we write them as follows;

$$\text{if} \quad LB_{R_{jt}} = 0 \quad ; \quad y_{j\acute{k}kt} = 0 \quad ; \quad \forall \quad j \in J,\ \acute{k} \in K,\ k \in K \setminus \{0\},\ t \in T \tag{3.29}$$

$$\text{if} \quad LB_{R_{jt}} > 0 \quad ; \quad y_{j\acute{k}0t} = 0 \quad ; \quad \forall \quad j \in J,\ \acute{k} \in K,\ t \in T \tag{3.30}$$

The algorithm adds these equalities as constraints to the linearized model that lead to fixing a significant number of binary variables.

In this thesis, $\vartheta = 0.07$ if the number of customers is less than 500 ($\mid I \mid < 500$) or the number of facilities is less than 100 ($\mid J \mid < 100$). Otherwise, it is set $\vartheta = 0.002$. Actually, $\frac{\vartheta}{1+\vartheta}$ assures the minimum allowed percent of utilization of open facilities.

As a result, by solving the linearized model with $\varepsilon = 10^{-1}$ in the first iteration, an approximation of the original problem optimal solution is found in the location. However, the capacity level selection and the allocation decisions, which are strongly correlated with the congestion cost, remain to be determined in the following iterations with more accuracy.

Reduction in Linearization Constraints Number

Since the beginning of the second iteration, the function $\rho_{jt}(R_{jt}) = \frac{R_{jt}}{1+R_{jt}}$ is not required to be linearized with $R_{jt} \in [0, \infty]$ if and only if $min \left\{ LB_{R_{jt}} > 0\ ;\ j \in J\ ,\ t \in T \right\} >$

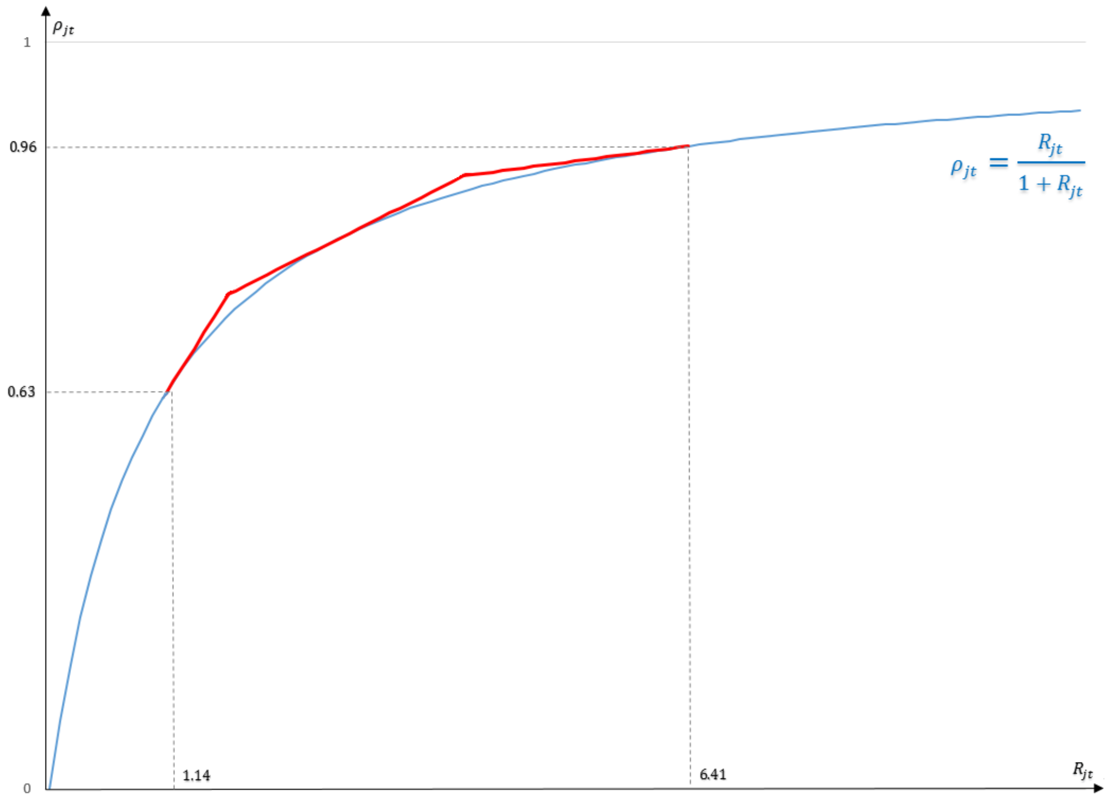


FIGURE 3.2: Shrinkage in the piecewise linearized function as a result of bounding ρ_{jt} and R_{jt}

ϑ or $max \left\{ UB_{\rho_{jt}} \;\; ; \;\; j \in J \; , \; t \in T \right\} < 1$. Having set LB_R and UB_R the lowest and largest positive values of $LB_{R_{jt}}$ and $UB_{R_{jt}}$ among all the $j \in J$ and $t \in T$, respectively, at the beginning of each iteration, we linearize function $\rho_{jt}(R_{jt}) = \frac{R_{jt}}{1+R_{jt}}$ only in $R_{jt} \in [LB_R, UB_R]$. Such a shrinkage in the piecewise linearization function leads to a significant reduction in number of lazy constraints 2.16 within resolving the linearized MIP model, because, it clearly reduces $\mid H \mid$ where the number of these constraints is $\mid J \mid * \mid H \mid * \mid T \mid$ in each iteration.

For example, if $LB_R = 1.14$ and $UB_R = 6.41$ at the end of the first iteration, the $\rho_{jt} - R_{jt}$ graphs would be as illustrated in the figure 3.2 before resolving the model in the second iteration.

Starting Solution

In CPLEX point of view, a new MIP model gets solved in each iteration, without any memory of the solved models in the previous iterations, while, the obtained solution in each iteration is an approximation of the exact solution in the following iteration. Hence, to accelerate the CPLEX performance since the second iteration, the algorithm uses the earned optimal values of x_{ijt}, $y_{j\acute{k}kt}$ and w_{jkt} variables in each iteration as the starting solution within the next iteration. Due to having constraints that certain $\sum_{i\in I} \lambda_{it}\ x_{ijt} = \sum_{k\in K} \mu_{jkt}\ z_{jkt}$, we deliberately skip specifying values of z_{jkt} variables simultaneously with it of x_{ijt} variables, because, it might leads to an infeasible starting solution which is not our preference.

Having defined the starting solution for CPLEX, its *effort level* and *AdvInd (Advanced Start Switch)* is set to 3 and 2, respectively. In such a case, CPLEX solves a sub-MIP, retains the current incumbent, re-applies presolve, and starts a new search from a new root. It results in smaller Branch-and-Cut trees in iterations after the the first one.

Algorithm 5 : Iterative Solution Method Algorithm

```
Require: Choose an acceptable linearization tolerance ε
initialize Iteration Nubmber = 1
while 10^(−Iteration Number) ≥ ε do
    if Iteratoin Number > 1 then
        Shrinkage the piecewise linearized functions
        Starting Solution ⟵ Solution obtained from the previous iteration
    end if
    Do linearization with gap 10^(−Iteration Number)
    if Iteration Number = 1 then
        Solve the linearized MIP model
    else
        Solve the sub-MIP model
    end if
    for all j ∈ J and t ∈ T do
        bound ρ_jt and R_jt
        if Iteration Number = 1 then
            Fix the Openness/Closeness of facility j in time period t
        end if
    end for
end while
```

A summary of the iterative solution method algorithm is illustrated in algorithm 5. Considering that this method is a heuristic solution method, it is aimed to define a controlling parameter to manage interaction between quality and taken time of solutions. In this method, it is the linearization gap in the first iteration that controls aggressiveness of the heuristic. Enlargement in the initial value of ε leads to easier problems in the first iteration (because of less number of linearization constraints). Consequently, the first iteration, which takes most of the algorithm CPU time, would be processed faster, but with less accurate approximation, when all of the following iterations are based on this approximation. As a result, it may affect the quality of the final solution.

Due to the focus of this paper which is on a strategical problem, the main concern is the quality of the heuristic. Thus, in this paper, the initial value of ε is 10^{-1} that is small enough to have an appropriate approximation. Table 4.10 approves this claim.

3.3.2 Aggressively Fixing Iterative Solution Method

In this section, we present a more aggressive version of the heuristic method explained in section 3.3.1. The main difference between these two version is at the beginning of the second iteration, when the algorithm fixes openness or closeness of all facilities, according the approximate solution obtained in the first iteration. In the aggressive version, the algorithm fixes all the binary decision variables as they are estimated in the first iteration. In other words, the aggressively fixing algorithm does not only fix openness/closeness of all facilities, implicitly, but also fixes the acquired capacity level in each facility.

Such a slight modification enables the algorithm to fix location and capacity selection decisions in the first iteration. As an interesting consequence, the sub-problem solved in the following iterations is an allocation problem which corresponds a linear continuous problem and determines the allocation and congestion costs, simultaneously.

In mathematical modeling point of view, in the aggressive version of the iterative solution method, cutting planes 3.29 and 3.30 are replaced with constraints written below;

$$y_{j\acute{k}kt} = \ddot{y}_{j\acute{k}kt} \qquad ; \qquad \forall \quad j \in J,\ \acute{k} \in K,\ k \in K,\ t \in T \tag{3.31}$$

where $\ddot{y}_{j\acute{k}kt}$ is a constant value equal with the optimal value of $y_{j\acute{k}kt}$ obtained in the first iteration (${y^*}_{j\acute{k}kt}$).

To have a more efficient execution of the iterations followed by the first, the model is rewritten as is shown below;

minimize $Z_H\left(x^H, z^H, w^H\right) =$

$$
\begin{aligned}
&\textit{resulted location cost in the first iteration}\\
&+\sum_{j\in J}\sum_{t\in t}\left(\sum_{i\in I} c_{ijt}\,\lambda_{it}\,x_{ijt} + \sum_{k\in K} p_{jkt}\,\mu_{jkt}\,z_{jkt}\right)\\
&+\sum_{j\in J}\sum_{t\in T}\frac{h_{jt}}{2}\left(\sum_{k\in K}\left(\left(1+C_{s_{jt}}^2\right) w_{jkt} + \left(1-C_{s_{jt}}^2\right) z_{jkt}\right)\right)
\end{aligned}
$$

subject to:

$$\sum_{i\in I}\lambda_{it}\,x_{ijt} = \sum_{k\in K}\mu_{jkt}\,z_{jkt} \qquad ;\ \forall \quad j\in J,\ t\in T$$

$$\sum_{j\in J} x_{ijt} = 1 \qquad ;\ \forall \quad i\in I,\ t\in T$$

$$\sum_{k\in K\backslash\{0\}} z_{jkt} \le \frac{\sum_{k\in K\backslash\{0\}} w_{jkt}}{(1+R^h)^2} + \left(\frac{(R^h)^2}{(1+R^h)^2}\right)\sum_{\acute{k}\in K}\sum_{k\in K\backslash\{0\}} y_{j\acute{k}kt} \qquad ;\ \forall \quad j\in J,\ h\in H,\ t\in T$$

$$\begin{aligned} &z_{jkt} \le \textstyle\sum_{\acute{k}\in K} \ddot{y}_{j\acute{k}kt}\\ &z_{j0t} = 0\end{aligned} \qquad ;\ \forall \quad j\in J,\ k\in K\backslash\{0\},\ t\in T$$

$$\begin{aligned} &w_{jkt} \le M \textstyle\sum_{\acute{k}\in K} \ddot{y}_{j\acute{k}kt}\\ &w_{j0t} = 0\end{aligned} \qquad ;\ \forall \quad j\in J,\ k\in K\backslash\{0\},\ t\in T$$

$$x_{ijt} \le \sum_{\acute{k}\in K}\sum_{k\in K\backslash\{0\}} \ddot{y}_{j\acute{k}kt} \qquad ;\ \forall \quad i\in I,\ j\in J,\ t\in T$$

$$\sum_{j\in J}\sum_{k\in K}\mu_{jkt}\,z_{jkt} = \sum_{i\in I}\lambda_{it} \qquad ;\ \forall \quad t\in T$$

$$0 \le x_{ijt} \le 1 \qquad ;\ \forall \quad i\in I,\ j\in J,\ t\in T$$

$$0 \le z_{jkt} \le 1 \qquad ;\ \forall \quad j\in J,\ k\in K,\ t\in T$$

$$0 \le w_{jkt} \qquad ;\ \forall \quad j\in J,\ k\in K,\ t\in T$$

As could be seen, there is no binary decision variable in this LP model and constraints 2.4 and 2.5 are eliminated.

Such a fixing accelerates CPLEX performance since the second iteration. A summary of this heuristic solution method for the original problem is illustrated

Algorithm 6 : Fixing Iterative Solution Method Algorithm

Require: Choose an acceptable linearization tolerance ε
initialize $Iteration\ Nubmber = 1$
while $10^{-Iteration\ Number} \geq \varepsilon$ **do**
  if $Iteratoin\ Number > 1$ **then**
    Shrinkage the piecewise linearized functions
    Starting Solution $\longleftarrow$ Solution obtained from the previous iteration
  end if
  Do linearization with gap $10^{-Iteration\ Number}$
  if $Iteration\ Number = 1$ **then**
    Solve the linearized MIP model
  else
    Solve the fixed MIP (LP) model
  end if
  for all $j \in J$ and $t \in T$ **do**
    bound z_{jkt} and w_{jkt}
    if $Iteration\ Number = 1$ **then**
      Fix the binary decision variables
    end if
  end for
end while

in algorithm 6. The implementation of this algorithm has another difference with that is explained in section 3.3.1 which is the starting solution for iterations followed by the first. The *effort level* of the starting solution in this algorithm is 2. It means that CPLEX solves a fixed MIP model (actually a LP model which is not necessarily the relaxation of the corresponding MIP model) since the second iteration.

Another benefit of solving the problem iteratively is that at the end of each iteration, the algorithm can compare the optimal objective function value of the linearized model ($Z_H^*(x^H, y^H, z^H, w^H)$) with the corresponding original objective function value ($Z(x^H, y^H)$), which are LB and UB for the optimal original cost ($Z^*(x, y)$), respectively. If the gap was ignorable, the algorithm could be terminated. Having considered *proposition* 2.2 mentioned in section 2.4.3, if the linearized model is not solvable within the time limitation, it is the obtained LB $Z_H^*(x^H, y^H, z^H, w^H)$ which is valid as a LB for $Z^*(x, y)$.

In this paper, the iterative algorithm is terminated after 8 iterations, when the linearization gap (ε value) in the last iteration is 10^{-8}.

4 Computational Results

This chapter indicates the numerical result of modeling, tightening and solution methods. Furthermore, the most difficult set of constraints of the model is defined as *Lazy Constraints*. It accelerates solving the LP relaxation of the formulation. Moreover, we find a problem-specific CPLEX configuration and replace it with the default CPLEX configuration in sake of the best execution of our model. At the end of the chapter, the instances that the model is tested on them are described, where a sensitivity analysis is also provided.

In this research, all the solution procedures were coded in C++ and the MIP problems were solved using IBM ILOG CPLEX 12.7.1 interfaced in Visual Studio 2015/C++ on a Dell Precision Tower 3620 PC with 4 Cores 3.60 GHz processor and 16 GB of RAM.

CPLEX takes advantage of a branch-and-cut method to solve such a combinatorial optimization problem, while in each node, the LP relaxation of the model is solved to obtain a lower bound (LB) and some heuristics are applied to obtain an upper bound (UB) for the objective function value of the optimal solution.

Furthermore, to have a fare benchmarking, we observe the CPLEX performance with an empty Control Callback, because, using a Control Callback changes some default configurations of CPLEX such as setting *multi-threads* to *single-thread* computing and *dynamic search* in the tree to *traditional branch-and-bound*. Hence, to compare the efficiency of our algorithm with it of CPLEX 12.7.1 default algorithm in theoretical point of view, we consider CPLEX with an empty Callback as the basis, because, its default takes advantages of parallel programming as well.

4.1 Test Problems

In this section, a few of notation relating to either of the formulation or numerical tables are explained. To know more about the randomly generated test problems used in this study, it is referred to e-companion paper of Jena, Cordeau, and Gendron, 2015. The paper explains how the test problems are created for a dynamic facility location problem which determines location, allocation and capacity acquisition decisions in a multi-period horizon. In this thesis, as a contribution to that paper, we also consider congestion. Thus, h_j denoting holding cost in location j is added to the parameters described in that paper.

4.1.1 Fixed Costs Matrix

In addition of other input data of the problem, we define the following fixed costs to characterize possible transitions of capacity levels in facility;

- f^o_{jkt} and f^c_{jkt} are the costs to open and close the facility j with capacity level k in period t, respectively.

- f^e_{jkt} and f^r_{jkt} are the costs to expand and reduce the capacity of open facility j by k levels in period t, respectively.

- f^m_{jkt} is the cost to maintain the facility j in capacity level k in period t.

Then, we set the fixed costs matrix of $f_{j\acute{k}kt}$ as demonstrated below as coefficients of location and capacity selection variables ($y_{j\acute{k}kt}$) used in the formulation.

$$f_{j\acute{k}kt} = \begin{cases} f^{o}_{jkt} + f^{m}_{jkt} & ; \quad \text{if} \quad \acute{k} = 0 \text{ and } k > 0 \\ f^{c}_{jkt} & ; \quad \text{if} \quad \acute{k} > 0 \text{ and } k = 0 \\ f^{e}_{j(k-\acute{k})t} + f^{m}_{jkt} & ; \quad \text{if} \quad \acute{k} > 0 \text{ and } k > \acute{k} \\ f^{e}_{j(\acute{k}-k)t} + f^{m}_{jkt} & ; \quad \text{if} \quad \acute{k} > k \text{ and } k > 0 \\ f^{m}_{jkt} & ; \quad \text{if} \quad \acute{k} = k \end{cases}$$

As is mentioned in section 2.3, $f_{j\acute{k}kt}$ denotes the fixed cost of facility j to hold capacity level k in period t while it has been $\acute{k}$ in the previous time period.

Demand Types

In order to test the model and the algorithm on various instances, two different kinds of demand is considered in test problems. In the cases denoted by demand type 1 (regular), trend of each individual customer is regular. It means that within different time periods, demand of each individual customer is valued by a normal distribution with a same mean, while it is the opposite in the cases denoted by demand type 2 (irregular), where the fluctuation of demand of each individual customer could be high, within time periods.

Map Types

For each of the different problem sizes, demand zones have been randomly generated following a continuous uniform distribution, while the first $| J |$ points of $| I |$ demand zones have additionally been defined as candidate facility locations and therefore coincide with the customer demand points. The networks are generated on squares of the three sizes: 300km, 380km and 450km which are denoted by map type 1, 2 and 3, respectively.

4.2 Numerical Result of the Formulation Tightening

4.2.1 Result of adding SI

Tables 4.1 and 4.2 illustrate CPLEX performance with different selections of *SI* inequalities. As could be noticed, including inequalities 2.23 is vital, because, even the smallest and simplest test problem is not solved within the time limitation (2 hours), if the formulation does not include inequalities 2.23. Furthermore, adding all sets of *SI* is not interesting, because, it makes the model too large that results in solver performance deceleration and also solutions quality deterioration. Although the fact that having equalities 2.24 leads to less number of tuned nodes, adding only 2.23 is the best selection to reduce the CPU time and improve the LB in cases that the problem is not solved optimally.

4.2.2 Result of adding ADC

As the number of this set of inequalities is polynomial, they are added directly to the model. Tables 4.3 indicates the contribution of having inequalities 2.27 and 2.26, compared with the best known formulation so far, which is illustrated in table 4.2 as well. As is shown, benefits of incorporating ADC are not consistent. As it increases the average CPU time, adding ADC to the formulation is skipped.

4.2.3 Result of Lifting Linearization Constraints

Table 4.4 demonstrates two different methods of having linearization constraints. As could be noticed, substituting 2.16 with 2.28 tremendously reinforces the solver execution.

4.3 Numerical Result of Solving Methods

4.3.1 Result of Mixed-Dicuts

Having implemented algorithm 3 described in section 3.1.1 with either of separation procedures, no violation of inequalities 3.1 is found. Table 4.5 indicates the lack of effect on CPLEX execution when adding Mixed-Dicuts is involved.

No violation of Mixed-Dicuts is found when Strong Inequalities (SI) 2.23 are added to the model as constraints. If we sacrifice the best known model (which is discussed in section 2.6) and remove inequalities 2.23, violation of Mixed-Dicuts could be seen and cuts 3.1 are involved in the Branch-and-Cut algorithm. However, the drawback of removing inequalities 2.23 is much more significant than the benefit of adding Mixed-Dicuts. Tables 4.6 illustrates that it does not

TABLE 4.1: Impact of adding SI with other forms than 2.23

Test Problems				**no SI is added**					**inequalities 2.24 are added**				
$\mid I \mid , \mid J \mid$	**Holding Cost (h)**	**Demands Type**	**Optimal Solution Value**	**MIP LB value**	**MIP UB value**	**MIP gap**	**CPU time (sec.)**	**Nodes**	**MIP LB value**	**MIP UB value**	**MIP gap**	**CPU time (sec.)**	**Nodes**
50, 50	50000	1	2,838,580	2,730,580	2,838,580	3.80 %	time	50256	2,741,470	2,844,790	3.63 %	time	21609
		2	2,861,670	2,706,700	2,867,980	5.62 %	time	71861	2,742,920	2,861,670	4.15 %	time	36840
	75000	1	3,070,350	2,912,820	3,075,580	5.29 %	time	51400	2,909,790	3,073,940	5.34 %	time	41400
		2	3,111,200	2,936,920	3,111,200	5.60 %	time	38216	2,905,500	3,148,470	7.72 %	time	40200
250, 50	50000	1	7,088,550	6,896,340	7,138,820	3.40 %	time	23500	6,914,300	7,142,620	3.20 %	time	5844
		2	6,715,500	6,587,460	6,720,150	1.97 %	time	21290	6,586,470	6,715,500	1.92 %	time	8017
	75000	1	7,413,160	7,125,620	7,549,930	5.62 %	time	29700	7,152,350	7,533,430	5.06 %	time	14808
		2	7,015,600	6,847,040	7,035,670	2.68 %	time	13300	6,826,280	7,071,450	3.47 %	time	11901
250, 100	50000	1	7,043,580	6,605,480	7,125,110	7.29 %	time	4000	6,730,550	7,227,550	6.88 %	time	284
		2	6,674,100	6,322,380	6,678,420	5.33 %	time	7675	6,396,540	6,711,130	4.69 %	time	159
	75000	1	7,365,700	6,936,030	7,449,650	6.89 %	time	4872	6,990,570	7,764,870	9.97 %	time	98
		2	6,973,230	6,672,810	7,182,290	7.09 %	time	3966	6,658,530	8,224,860	19.04 %	time	127
					average:	**5.05 %**				**average:**	**6.26 %**		

$\mid I \mid , \mid J \mid$	Holding Cost (h)	Demands Type	Optimal Solution Value	**inequalities 2.25 are added** MIP LB value	MIP UB value	MIP gap	CPU time (sec.)	Nodes	**both of 2.24 and 2.25 are added** MIP LB value	MIP UB value	MIP gap	CPU time (sec.)	Nodes
50,50	50000	1	2,838,580	2,699,340	2,838,920	4.92 %	time	64000	2,727,650	2,838,580	3.91 %	time	22900
		2	2,861,670	2,743,870	2,861,670	4.12 %	time	28962	2,706,100	2,867,980	5.64 %	time	46800
	75000	1	3,070,350	2,907,630	3,070,350	5.30 %	time	54324	2,897,090	3,122,140	7.21 %	time	34071
		2	3,111,200	2,893,520	3,111,200	7.00 %	time	58400	2,940,850	3,123,000	5.83 %	time	24650
250,50	50000	1	7,088,550	6,905,240	7,088,550	2.59 %	time	17800	6,918,770	7,144,750	3.16 %	time	6031
		2	6,715,500	6,605,880	6,715,500	1.63 %	time	21437	6,559,860	6,762,980	3.00 %	time	5268
	75000	1	7,413,160	7,164,370	7,495,570	4.42 %	time	13600	7,109,960	7,539,460	5.70 %	time	7245
		2	7,015,600	6,730,200	7,050,270	4.54 %	time	12564	6,717,050	7,077,000	5.09 %	time	11260
250,100	50000	1	7,043,580	6,627,890	7,182,830	7.73 %	time	6600	6,727,020	7,058,730	4.70 %	time	762
		2	6,674,100	6,297,740	6,685,570	5.80 %	time	4800	6,393,390	7,280,580	12.19 %	time	90
	75000	1	7,365,700	6,954,890	7,448,280	6.62 %	time	3800	7,013,480	7,382,770	5.00 %	time	107
		2	6,973,230	6,683,130	6,990,750	4.40 %	time	4000	6,681,220	7,110,490	6.04 %	time	245
					average:	**4.92 %**				**average:**	**5.62 %**		

(For all cases; $\mid K \mid = 5$, $\mid T \mid = 4$, $C_s = 0.5$, $\varepsilon = 10^{-3}$, Map Type=3)

(Time limitation is 7200 seconds.)

TABLE 4.2: Impact of adding SI with the form 2.23

Test Problems				all kinds of SI are added					both of 2.23 and 2.24 are added				
| I | , | J |	Holding Cost (h)	Demands Type	Optimal Solution Value	MIP LB value	MIP UB value	MIP gap	CPU time (sec.)	Nodes	MIP LB value	MIP UB value	MIP gap	CPU time (sec.)	Nodes
50 , 50	50000	1	2,838,580		optimal		292	552		optimal		299	545
		2	2,861,670		optimal		896	3430		optimal		709	2699
	75000	1	3,070,350		optimal		1317	3893		optimal		1155	2056
		2	3,111,200		optimal		4219	9892		optimal		2646	9334
250 , 50	50000	1	7,088,550		optimal		760	49		optimal		891	42
		2	6,715,500		optimal		759	19		optimal		712	45
	75000	1	7,413,160		optimal		1730	875		optimal		1381	730
		2	7,015,600		optimal		1197	118		optimal		1094	454
250 , 100	50000	1	7,043,580		optimal		3793	641		optimal		3346	377
		2	6,674,100		optimal		3800	401		optimal		4373	701
	75000	1	7,365,700	7,307,610	7,365,700	0.79 %	time	1075	7,312,520	7,365,920	0.72 %	time	1420
		2	6,973,230	6,930,090	7,008,950	1.25 %	time	601	6,937,340	6,991,120	0.01 %	time	1000
						average:	**2764**				**average:**	**2584**	

				both of 2.23 and 2.25 are added					inequalities 2.23 are added				
50 , 50	50000	1	2,838,580		optimal		353	1660		optimal		340	1044
		2	2,861,670		optimal		491	2922		optimal		280	1462
	75000	1	3,070,350		optimal		665	2270		optimal		657	2775
		2	3,111,200		optimal		2731	10021		optimal		1612	8937
250 , 50	50000	1	7,088,550		optimal		395	68		optimal		380	100
		2	6,715,500		optimal		305	43		optimal		308	28
	75000	1	7,413,160		optimal		744	1091		optimal		650	503
		2	7,015,600		optimal		588	333		optimal		650	329
250 , 100	50000	1	7,043,580		optimal		1844	683		optimal		1705	418
		2	6,674,100		optimal		1368	358		optimal		1863	303
	75000	1	7,365,700	7,323,310	7,373,490	0.68 %	time	2800	7,338,850	7,365,700	0.36 %	time	2762
		2	6,973,230	6,957,560	6,973,230	0.00 %	time	2883		optimal		5140	1985
						average:	**1990**				**average:**	**1732**	

(For all cases; $\mid K \mid = 5$, $\mid T \mid = 4$, $C_s = 0.5$, $\varepsilon = 10^{-3}$, Map Type=3)

(Time limitation is 7200 seconds.)

TABLE 4.3: Impact of adding ADC

Test Problems				no ADC is added					inequalities 2.26 are added					inequalities 2.27 are added				
$\mid I \mid , \mid J \mid$	Holding Cost (h)	Demands Type	Optimal Solution Value	MIP LB value	MIP UB value	MIP gap	CPU time (sec.)	Flow Cover cuts generated	MIP LB value	MIP UB value	MIP gap	CPU time (sec.)	Flow Cover cuts generated	MIP LB value	MIP UB value	MIP gap	CPU time (sec.)	Flow Cover cuts generated
50 , 50	50000	1	2,838,580		optimal		340	202 (39%)		optimal		197	189 (36%)		optimal		288	175 (35%)
		2	2,861,670		optimal		280	196 (35%)		optimal		423	194 (35%)		optimal		488	198 (36%)
	75000	1	3,070,350		optimal		657	226 (39%)		optimal		391	237 (41%)		optimal		817	234 (37%)
		2	3,111,200		optimal		1612	250 (39%)		optimal		1877	238 (37%)		optimal		1725	247 (38%)
250 , 50	50000	1	7,088,550		optimal		380	135 (41%)		optimal		340	127 (38%)		optimal		235	112 (35%)
		2	6,715,500		optimal		308	125 (42%)		optimal		302	119 (42%)		optimal		269	130 (45%)
	75000	1	7,413,160		optimal		650	142 (31%)		optimal		704	126 (29%)		optimal		645	129 (28%)
		2	7,015,600		optimal		650	157 (35%)		optimal		512	141 (32%)		optimal		611	148 (34%)
250 , 100	50000	1	7,043,580		optimal		1705	204 (38%)		optimal		1867	215 (42%)		optimal		1772	237 (42%)
		2	6,674,100		optimal		1863	187 (41%)		optimal		1715	206 (44%)		optimal		1894	197 (45%)
	75000	1	7,365,700	7,338,850	7,365,700	0.36 %	time	188 (24%)	7,327,350	7,365,700	0.52 %	time	191 (25%)	7,343,960	7,365,700	0.30 %	time	185 (25%)
		2	6,973,230		optimal		5140	222 (33%)	6,952,800	6,973,230	0.29 %	time	233 (34%)	6,959,110	6,973,230	0.20 %	time	237 (34%)
						average:	**1732**				**average:**	**1894**				**average:**	**1929**	

(For all cases; $\mid K \mid= 5$, $\mid T \mid= 4$, $C_s = 0.5$, $\varepsilon = 10^{-3}$, Map Type=3)

(Time limitation is 7200 seconds.)

TABLE 4.4: Impact of lifting linearization constraints

Test Problems					having inequalities 2.16					having inequalities 2.28				
$\lvert I \rvert$, $\lvert J \rvert$	Holding Cost (h)	Demands Type	Map Type	**Optimal Solution Value**	MIP LB value	MIP UB value	MIP gap	CPU time (sec.)	Nodes	MIP LB value	MIP UB value	MIP gap	CPU time (sec.)	Nodes
50 , 50	50000	1	1	2,457,150		optimal		1448	4828		optimal		210	305
			2	2,712,890		optimal		133	354		optimal		34	0
			3	2,838,580		optimal		312	1044		optimal		44	3
		2	1	2,443,180		optimal		938	4137		optimal		147	130
			2	2,721,570		optimal		144	533		optimal		81	75
			3	2,861,670		optimal		292	1462		optimal		82	39
	75000	1	1	2,624,330		optimal		1370	5998		optimal		296	318
			2	2,946,210		optimal		834	2998		optimal		116	55
			3	3,070,350		optimal		693	2775		optimal		112	44
		2	1	2,605,920		optimal		1122	4294		optimal		524	1072
			2	2,965,450		optimal		1205	5137		optimal		444	1240
			3	3,111,200		optimal		1706	8937		optimal		375	816
	100000	1	1	2,750,950		optimal		1546	4858		optimal		280	254
			2	3,101,890		optimal		1200	4388		optimal		291	414
			3	3,283,420		optimal		3715	17952		optimal		692	1349
		2	1	2,688,720		optimal		434	1295		optimal		211	220
			2	3,131,160		optimal		1202	4553		optimal		419	845
			3	3,288,410		optimal		3850	14638		optimal		1017	2588
	150000	1	1	2,906,440		optimal		416	801		optimal		472	563
			2	3,391,380		optimal		3067	11189		optimal		928	2318
			3	3,609,420	3,557,440	3,609,420	1.44 %	time	21403		optimal		2445	5747
		2	1	2,854,320		optimal		283	445		optimal		233	134
			2	3,400,260		optimal		3682	13908		optimal		2887	8248
			3	unknown	3,476,910	3,602,940	3.50 %	time	26003	3,589,760	3,592,060	0.06 %	time	16356
250 , 50	50000	1	1	5,696,480		optimal		1072	960		optimal		507	21
			2	6,380,410		optimal		347	65		optimal		230	7
			3	7,088,550		optimal		382	100		optimal		232	0
		2	1	5,384,890		optimal		552	177		optimal		347	5
			2	6,014,170		optimal		308	22		optimal		221	0
			3	6,715,500		optimal		313	28		optimal		180	0
	75000	1	1	6,037,000	6,003,240	6,045,250	0.69 %	time	5383		optimal		1336	228
			2	6,707,030		optimal		579	294		optimal		361	6
			3	7,413,160		optimal		654	503		optimal		378	9
		2	1	5,718,990		optimal		3735	4694		optimal		1107	265
			2	6,327,310		optimal		528	193		optimal		230	1
			3	7,015,600		optimal		657	329		optimal		252	0
	100000	1	1	6,357,050	6,227,080	6,403,950	2.76 %	time	4885		optimal		4198	1476
			2	7,033,280		optimal		2587	3399		optimal		805	57
			3	7,737,360		optimal		2875	3899		optimal		692	199
		2	1	6,025,820	5,915,440	6,043,450	2.12 %	time	4522		optimal		2999	1056
			2	6,640,160		optimal		2200	2645		optimal		653	73
			3	7,315,410		optimal		1412	1214		optimal		587	23
	150000	1	1	unknown	6,614,540	7,046,340	6.13 %	time	2840	6,718,370	7,037,490	4.53 %	time	1653
			2	7,662,390	7,531,050	7,662,390	1.71 %	time	7831		optimal		3664	1227
			3	8,368,670	8,241,170	8,369,550	1.53 %	time	7362		optimal		4760	2144
		2	1	unknown	6,288,080	6,585,500	4.52 %	time	4208	6,377,230	6,597,430	3.34 %	time	1484
			2	7,234,490	7,110,250	7,234,490	1.72 %	time	4768		optimal		3750	1172
			3	7,914,250	7,796,830	7,914,250	1.48 %	time	7687		optimal		2863	1414
250 , 100	50000	1	1	5,615,110	5,565,510	5,630,600	1.16 %	time	1681		optimal		1944	83
			2	6,272,190		optimal		1129	87		optimal		487	0
			3	7,043,580		optimal		1721	418		optimal		832	14
		2	1	5,319,800	5,285,700	5,319,800	0.64 %	time	2299		optimal		2026	121
			2	5,930,430		optimal		1090	62		optimal		501	0
			3	6,674,100		optimal		1885	303		optimal		717	11
	75000	1	1	unknown	5,848,200	6,009,530	2.68 %	time	927	5,933,150	5,953,440	0.34 %	time	100
			2	6,609,690		optimal		4207	1854		optimal		1182	21
			3	7,365,700	7,338,710	7,365,700	0.37 %	time	2740		optimal		1628	16
		2	1	unknown	5,559,070	5,732,890	3.03 %	time	1204	5,615,100	5,663,050	0.85 %	time	184
			2	6,244,240		optimal		5265	1810		optimal		1103	10
			3	6,973,230		optimal		5180	1985		optimal		965	6
	100000	1	1	unknown	6,129,200	7,154,020	14.33 %	time	138	6,200,870	6,496,190	4.55 %	time	0
			2	6,941,910	6,862,250	6,963,430	1.45 %	time	899		optimal		6086	413
			3	7,687,690	7,604,600	7,695,370	1.18 %	time	2767		optimal		5510	190
		2	1	unknown	5,807,860	6,124,490	5.17 %	time	138	5,872,070	6,113,200	3.94 %	time	8
			2	6,557,620	6,505,940	6,560,230	0.83 %	time	2000		optimal		4523	102
			3	7,271,670	7,218,130	7,271,670	0.74 %	time	2099		optimal		3374	94
	150000	1	1	unknown	6,477,940	9,041,130	28.35 %	time	54	6,526,270	7,314,020	10.77 %	time	0
			2	unknown	7,332,550	7,944,080	7.70 %	time	283	7,431,400	7,634,330	2.66 %	time	4
			3	unknown	8,118,310	8,453,850	3.97 %	time	380	8,201,380	8,344,400	1.71 %	time	9
		2	1	unknown	6,155,720	6,752,430	8.84 %	time	56	6,215,140	6,340,200	1.97 %	time	0
			2	unknown	6,920,350	7,392,770	6.39 %	time	95	7,049,860	7,323,260	3.73 %	time	2
			3	unknown	7,686,490	8,062,140	4.66 %	time	328	7,771,970	7,879,180	1.36 %	time	10

500 , 100	50000	1	1	8,260,370		optimal		4083	49		optimal		2685	0
			2	9,479,780		optimal		3147	2		optimal		1845	0
			3	10,649,100		optimal		2885	0		optimal		1710	0
		2	1	7,787,230		optimal		4583	213		optimal		2563	5
			2	8,986,610		optimal		3415	41		optimal		2513	0
			3	10,053,800		optimal		2694	0		optimal		1879	0
	75000	1	1	8,555,230	8,543,050	8,555,230	0.14 %	time	435		optimal		3467	0
			2	9,762,540		optimal		4207	118		optimal		2095	0
			3	10,932,200		optimal		3529	22		optimal		1772	0
		2	1	8,059,270	8,045,870	8,059,310	0.17 %	time	747		optimal		3281	12
			2	9,251,560		optimal		6236	1381		optimal		2872	16
			3	10,313,900		optimal		3886	36		optimal		1874	0
	100000	1	1	unknown	8,807,130	8,853,940	0.53 %	time	121	8,840,940	8,849,750	0.10 %	time	0
			2	10,045,200		optimal		5831	596		optimal		3513	5
			3	11,215,200		optimal		4902	368		optimal		2527	0
		2	1	unknown	8,281,390	8,335,170	0.65 %	time	41	8,325,170	8,330,960	0.07 %	time	22
			2	9,515,170	9,471,190	9,526,070	0.58 %	time	570		optimal		5469	77
			3	10,573,900		optimal		4416	78		optimal		2223	0
	150000	1	1	unknown		out of memory		time	0	9,263,730	9,572,490	3.23 %	time	0
			2	unknown	10,551,400	10,610,400	0.56 %	time	40	10,595,300	10,610,400	0.14 %	time	0
			3	unknown	11,708,200	11,781,200	0.62 %	time	341	11,773,000	11,781,200	0.07 %	time	5
		2	1	unknown	8,764,040	12,348,800	29.03 %	time	0	8,677,250	8,982,710	3.40 %	time	0
			2	unknown	9,950,200	10,090,500	1.39 %	time	10	10,002,600	10,747,300	6.93 %	time	0
			3	unknown	11,017,300	11,094,000	0.69 %	time	73	11,079,700	11,094,000	0.13 %	time	1
							average:	**4197**				**average:**	**2782**	

(For all cases; $\mid K \mid = 5 \, , \mid T \mid = 4 \, , C_s = 0.5 \, , \varepsilon = 10^{-3}$)

(Time limitation is 7200 seconds.)

worth to disregard the best form of SI. As a result, we keep the best known formulation and skip Mixed-Dicuts.

4.3.2 Result of enhanced-MIR

Although the fact that the number of each of inequalities 3.15-3.17 is polynomial, it is not a rational solution to add them directly to the formulation, because of several factors mentioned in the following. Firstly, there are specific values for the best α and β which lead to the most violation of the enhanced-MIR. Moreover, the numerical values of the variables of the LP solution affect the best α and β. So, it is not the most efficient method to specify α and β values in advance and add enhanced-MIR cuts to formulation. Furthermore, number of these inequalities are large and adding them directly increases the number of rows of the formulation considerably, while CPLEX cuts (mostly MIR cuts in this case) are also supposed to be added for each row. Consequently, it is too expensive and worthless to add inequalities 3.15-3.17 directly to the model. Table 4.7 affirms this claim, where α and β are valued randomly before the solver execution. Moreover, as this table indicates, no violation of enhanced-MIR cuts is found, when we try to separate them inside a CPLEX Cut-callback. In conclusion, as the same as Mixed-Dicuts, involving inequalities 3.15-3.17 has no contribution to the solution method.

TABLE 4.5: Impact of adding Mixed-Dicuts

Test Problems				**without Mixed-Dicuts**				**with Mixed-Dicuts separation**				
$\mid I \mid , \mid J \mid$	**Holding Cost (h)**	**Demands Type**	**Optimal Solution Value**	**MIP LB value**	**MIP UB value**	**MIP gap**	**CPU time (sec.)**	**MIP LB value**	**MIP UB value**	**MIP gap**	**CPU time (sec.)**	**added Mixed-Dicuts**
50 , 50	50000	1	2,838,580		optimal		45		optimal		42	0
		2	2,861,670		optimal		91		optimal		83	0
	75000	1	3,070,350		optimal		121		optimal		119	0
		2	3,111,200		optimal		375		optimal		383	0
250 , 50	50000	1	7,088,550		optimal		230		optimal		233	0
		2	6,715,500		optimal		178		optimal		180	0
	75000	1	7,413,160		optimal		379		optimal		379	0
		2	7,015,600		optimal		253		optimal		253	0
250 , 100	50000	1	7,043,580		optimal		831		optimal		833	0
		2	6,674,100		optimal		717		optimal		720	0
	75000	1	7,365,700		optimal		1618		optimal		1632	0
		2	6,973,230		optimal		965		optimal		970	0
						average:	**484**			**average:**	**486**	

(For all cases; $\mid K \mid = 5$, $\mid T \mid = 4$, $C_s = 0.5$, $\varepsilon = 10^{-3}$, Map Type=3)

(Time limitation is 7200 seconds.)

TABLE 4.6: Impact of adding Mixed-Dicuts when SI is in other forms than 2.23

Test Problems				having Mixed-Dicuts when no SI is added					having Mixed-Dicuts when SI are in form 2.24				
$\mid I \mid , \mid J \mid$	Holding Cost (h)	Demands Type	Optimal Solution Value	MIP LB value	MIP UB value	MIP gap	CPU time (sec.)	added Mixed-Dicuts	MIP LB value	MIP UB value	MIP gap	CPU time (sec.)	added Mixed-Dicuts
50 , 50	50000	1	2,838,580	2,534,460	2,844,790	10.91 %	time	85	2,553,600	2,888,230	11.59 %	time	83
		2	2,861,670	2,617,470	2,898,640	9.70 %	7203	108	2,573,340	2,932,320	12.24 %	time	77
	75000	1	3,070,350	2,688,970	3,122,970	13.90 %	time	90	2,688,540	3,102,870	13.35 %	time	68
		2	3,111,200	2,709,650	3,125,800	13.31 %	time	88	2,702,980	3,140,860	13.94 %	time	90
250 , 50	50000	1	7,088,550	6,330,950	7,251,610	12.70 %	time	77	6,282,830	7,332,920	14.32 %	time	137
		2	6,715,500	6,093,730	6,848,430	11.02 %	time	183	6,028,120	7,165,540	15.87 %	time	125
	75000	1	7,413,160	6,516,740	7,798,940	16.44 %	time	93	6,472,310	7,848,730	17.54 %	time	115
		2	7,015,600	6,203,670	7,087,220	12.47 %	time	278	6,172,370	8,588,730	28.13 %	time	176
250 , 100	50000	1	7,043,580	5,796,730	8,129,270	28.69 %	time	138	5,808,140	8,465,560	31.39 %	time	101
		2	6,674,100	5,722,020	8,601,970	33.48 %	time	212	5,682,970	9,134,580	37.79 %	time	122
	75000	1	7,365,700	5,977,910	9,796,230	38.98 %	time	115	5,923,020	78,765,400	92.48 %	time	87
		2	6,973,230	5,840,880	8,359,100	30.13 %	time	148	5,788,830	79,693,000	92.74 %	time	149
					average:	**19.31 %**				**average:**	**31.78 %**		

Test Problems				having Mixed-Dicuts when SI are in form 2.25					having Mixed-Dicuts when SI are in both forms 2.24 and 2.25				
50 , 50	50000	1	2,838,580	2,543,090	2,847,130	10.68 %	time	83	2,529,870	2,881,980	12.22 %	time	78
		2	2,861,670	2,582,500	3,017,350	14.41 %	time	88	2,602,660	2,884,370	9.77 %	time	95
	75000	1	3,070,350	2,672,870	3,107,550	13.99 %	time	87	2,688,700	3,102,760	13.34 %	time	72
		2	3,111,200	2,673,050	3,125,300	14.47 %	time	85	2,707,430	3,169,130	14.57 %	time	91
250 , 50	50000	1	7,088,550	6,203,610	7,175,250	13.54 %	time	153	6,242,720	7,394,910	15.58 %	time	146
		2	6,715,500	5,917,570	7,845,870	24.58 %	time	217	6,069,300	7,478,350	18.84 %	time	131
	75000	1	7,413,160	6,395,080	7,509,670	14.84 %	time	96	6,398,660	7,954,610	19.56 %	time	94
		2	7,015,600	6,142,550	7,628,900	19.48 %	7203	155	6,211,420	7,288,430	14.78 %	time	209
250 , 100	50000	1	7,043,580	5,794,010	7,803,790	25.75 %	time	126	5,804,980	12,426,300	53.28 %	time	95
		2	6,674,100	5,696,390	7,361,660	22.62 %	time	157	5,633,140	17,598,500	67.99 %	time	145
	75000	1	7,365,700	5,962,940	8,631,740	30.92 %	time	121	5,944,050	12,460,400	52.30 %	time	111
		2	6,973,230	5,894,440	8,457,270	30.30 %	time	133	5,752,140	12,104,600	52.48 %	time	99
					average:	**19.63 %**				**average:**	**28.73 %**		

(For all cases; $\mid K \mid = 5$, $\mid T \mid = 4$, $C_s = 0.5$, $\varepsilon = 10^{-3}$, Map Type=3)

(Time limitation is 7200 seconds.)

However, there is competitive configuration when constraints 2.16 are not lifted (not substituted with 2.28) and inequalities 3.15-3.17 are added directly to the model as constraints. Table 4.8 shows the impact of having enhanced-MIR cuts, in general. However, in this formulation, it is extravagant to lift linearization constraints and involve enhanced-MIR inequalities at a same time. As a result, due to complexity of generating enhanced-MIR cuts rather than replacing constraints 2.16 by 2.28, we keep the tightest formulation as the input model for the solver and skip enhanced-MIR cuts. In other words, it is concluded that tightening formulation makes the algorithm needless to incorporate sophisticated user-cuts.

4.3.3 Result of Heuristics for the Linearized Model

Table 4.9 demonstrates the impact of applying heuristics on the quality and time of the solutions. It could be noticed that the heuristic solution method is able to find a same solution as the optimal one in 75% of the instances, while it takes less solving time compared with exact methods in 65% of the instances. Although the fact, even the LP relaxation of extra-large instances is not solved, where *memory* is reported as the solution value.

Another point which could be seen is that in several instances, the obtained solution value by heuristics is higher than it by the exact method. Nevertheless, it is not necessarily a disadvantage for heuristics. As mention in *proposition 2.2* , the optimal solution value of the linearized model (MIP) is a lower bound for the original cost (MINLP optimal solution value). However, as these solutions are obtained heuristically and naturally do not match the optimal solutions, they are not valid lower bound. Thus, the improvement of the original cost is still under question. This consideration is discussed in chapter 3.3 with more details (see table 4.11).

4.3.4 Result of Heuristics for the Original Model

As is demonstrated in table 4.10, the fixing iterative solution method finds the optimal or close-to-optimal solutions, quicker than the exact method. The interesting point is that this method is promisingly capable of solving the original problem with more quality (less value of the original objective function) in considerably less time, compared with solution methods which set the desired

TABLE 4.7: Impact of adding enhanced-MIR inequalities

Test Problems				without enhanced-MIR inequalities					inequalities 3.15-3.17 are added directly to the model					inequalities 3.15-3.17 are added by separation				
$\mid I \mid , \mid J \mid$	Holding Cost (h)	Demands Type	Optimal Solution Value	MIP LB value	MIP UB value	MIP gap	CPU time (sec.)	generated MIR cuts by CPLEX	MIP LB value	MIP UB value	MIP gap	CPU time (sec.)	generated MIR cuts by CPLEX	MIP LB value	MIP UB value	MIP gap	CPU time (sec.)	added enhanced-MIR cuts
50 , 50	50000	1	2,838,580		optimal		45	353 (93%)		optimal		48	353 (93%)		optimal		53	0
		2	2,861,670		optimal		91	316 (91%)		optimal		169	335 (92%)		optimal		96	0
	75000	1	3,070,350		optimal		121	509 (93%)		optimal		230	464 (94%)		optimal		132	0
		2	3,111,200		optimal		375	384 (90%)		optimal		426	414 (88%)		optimal		386	0
250 , 50	50000	1	7,088,550		optimal		230	189 (72%)		optimal		210	188 (71%)		optimal		241	0
		2	6,715,500		optimal		178	136 (63%)		optimal		178	115 (68%)		optimal		192	0
	75000	1	7,413,160		optimal		379	370 (95%)		optimal		397	345 (96%)		optimal		394	0
		2	7,015,600		optimal		253	363 (85%)		optimal		216	371 (84%)		optimal		264	0
250 , 100	50000	1	7,043,580		optimal		831	292 (82%)		optimal		863	296 (82%)		optimal		845	0
		2	6,674,100		optimal		717	135 (94%)		optimal		709	129 (93%)		optimal		731	0
	75000	1	7,365,700		optimal		1618	518 (93%)		optimal		1911	534 (95%)		optimal		1667	0
		2	6,973,230		optimal		965	573 (94%)		optimal		1233	547 (94%)		optimal		991	0
						average:	**484**				**average:**	**549**				**average:**	**499**	

(For all cases; $\mid K \mid = 5$, $\mid T \mid = 4$, $C_s = 0.5$, $\varepsilon = 10^{-3}$, Map Type=3)

(Time limitation is 7200 seconds.)

TABLE 4.8: Impact of adding enhanced-MIR inequalities directly to the model when linearization constraints are not lifted

Test Problems				**no enhanced-MIR when linearization constraints are 2.16**					**inequalities 3.15-3.17 are added directly to the model when linearization constraints are 2.16**				
$\mid I \mid , \mid J \mid$	**Holding Cost (h)**	**Demands Type**	**Optimal Solution Value**	**MIP LB value**	**MIP UB value**	**MIP gap**	**CPU time (sec.)**	**Nodes**	**MIP LB value**	**MIP UB value**	**MIP gap**	**CPU time (sec.)**	**Nodes**
50 , 50	50000	1	2,838,580		optimal		312	1044		optimal		60	13
		2	2,861,670		optimal		292	1462		optimal		126	119
	75000	1	3,070,350		optimal		693	2775		optimal		90	45
		2	3,111,200		optimal		1706	8937		optimal		425	739
	100000	1	3,283,420		optimal		3715	17952		optimal		462	827
		2	3,288,410		optimal		3850	14638		optimal		476	755
	150000	1	3,609,420	3,557,440	3,609,420	1.44 %	time	21403		optimal		1438	3792
		2	3,592,060	3,476,910	3,602,940	3.50 %	time	26003		optimal		2056	4525
250 , 50	50000	1	7,088,550		optimal		382	100		optimal		240	0
		2	6,715,500		optimal		313	28		optimal		171	0
	75000	1	7,413,160		optimal		654	503		optimal		374	27
		2	7,015,600		optimal		657	329		optimal		234	0
	100000	1	7,737,360		optimal		2875	3899		optimal		742	146
		2	7,315,410		optimal		1412	1214		optimal		492	17
	150000	1	8,368,670	8,241,170	8,369,550	1.53 %	time	7362	8,306,380	8,368,670	0.74 %	time	3242
		2	7,914,250	7,796,830	7,914,250	1.48 %	time	7687		optimal		4677	2732
250 , 100	50000	1	7,043,580		optimal		1721	418		optimal		760	8
		2	6,674,100		optimal		1885	303		optimal		655	13
	75000	1	7,365,700	7,338,710	7,365,700	0.37 %	time	2740		optimal		1686	34
		2	6,973,230		optimal		5180	1985		optimal		1177	9
	100000	1	7,687,690	7,604,600	7,695,370	1.18 %	time	2767		optimal		5508	325
		2	7,271,670	7,218,130	7,271,670	0.74 %	time	2099		optimal		3479	130
	150000	1	unknown	8,118,310	8,453,850	3.97 %	time	380	8,198,450	8,378,030	2.14 %	time	13
		2	unknown	7,686,490	8,062,140	4.66 %	time	328	7,745,790	7,996,410	3.13 %	time	101
500 , 100	50000	1	10,649,100		optimal		2885	0		optimal		1699	0
		2	10,053,800		optimal		2694	0		optimal		1861	0
	75000	1	10,932,200		optimal		3529	22		optimal		1660	0
		2	10,313,900		optimal		3886	36		optimal		1829	0
	100000	1	11,215,200		optimal		4902	368		optimal		2010	0
		2	10,573,900		optimal		4416	78		optimal		2300	0
	150000	1	unknown	11,708,200	11,781,200	0.62 %	time	341	11,771,500	11,781,200	0.08 %	time	20
		2	unknown	11,017,300	11,094,000	0.69 %	time	73	11,073,600	11,094,000	0.18 %	time	7
						average:	**3974**				**average:**	**2272**	

(For all cases; $\mid K \mid = 5$, $\mid T \mid = 4$, $C_s = 0.5$, $\varepsilon = 10^{-3}$, Map Type=3)

(Time limitation is 7200 seconds.)

TABLE 4.9: Estimating the optimal objective function value of the linearized model (MIP) by heuristics

Test Problems					**MIP solved exactly with $\varepsilon = 10^{-1}$**		**MIP solved heuristically with $\Omega = 6$**	
$\mid T \mid$	$\mid I \mid , \mid J \mid$	**Holding Cost (h)**	**Demands Type**	**Optimal Solution Value**	**MIP UB value**	**CPU time (sec.)**	**obtained MIP objective function value**	**CPU time (sec.)**
4	250 , 100	100000	1	6,916,200	optimal	1606	6,917,180	1144
			2	6,684,980	optimal	5763	6,914,060	2177
		150000	1	7,105,020	optimal	2326	7,105,020	1379
			2	unknown	6,911,760	time	7,140,490	1223
	500 , 100	100000	1	10,470,000	optimal	1321	10,470,000	1319
			2	9,904,730	optimal	1560	9,904,730	1552
		150000	1	10,654,300	optimal	1950	10,654,300	1700
			2	10,079,900	optimal	1968	10,079,900	1956
	1000 , 100	100000	1	16,488,800	optimal	3892	16,488,800	3733
			2	15,528,700	optimal	5138	15,528,700	4984
		150000	1	16,668,500	optimal	5000	16,668,500	4878
			2	15,692,900	optimal	5791	memory	
					average:	**3626**	**average:**	**2770**
8	250 , 100	100000	1	unknown	13,077,500	time	13,053,000	6946
			2	unknown	12,874,400	time	12,874,400	4920
		150000	1	unknown	14,348,300	time	13,418,700	5598
			2	unknown	13,632,800	time	13,632,800	time
	500 , 100	100000	1	19,800,600	optimal	4912	19,800,600	4972
			2	18,663,000	optimal	6875	memory	
		150000	1	20,162,200	optimal	4478	20,162,200	4651
			2	18,993,800	optimal	6042	18,993,800	5371
	1000 , 100	100000	1	unknown	memory		memory	
			2	unknown	memory		memory	
		150000	1	unknown	memory		memory	
			2	unknown	memory		memory	
					average:	**6659**	**average:**	**6305**

(For all cases; $\mid K \mid = 5$, $C_s = 0.5$, Map Type=3)

(Time limitation is 7200 seconds.)

linearization gap at the beginning and solve the model in one iteration. As is indicated, in 58% of relatively difficult test problems ($|T| = 4$), the best known solution is obtained by the heuristic, while this figure is 35% in difficult test problems ($|T| = 8$).

4.4 Problem-specific Solver Setting

4.4.1 Lazy Constraints

One of the methods that leads to acceleration in solving a linearized model is relaxing a bunch of difficult constraints and add them as *lazy constraints*. In this case, lazy constraints are ignored in building the LP relaxation of the model and the LB is obtained regardless of them in each node. In the following, to obtain the UB, the solutions which violate the lazy constraints are rejected. As a subsequent, the UB remains valid for the optimal solution value. This method is implemented within a *LazyConstraintsCallback* in CPLEX.

In this paper, two different selections of lazy constraints are presented. In the first one, only constraints 2.28 are considered as lazy constraints, because, most of the difficulty of the model comes from these constraints. In other words, solving the LP relaxation of the model is easier in case of disregarding linearization constraints. Table 4.12 shows that converting constraints 2.28 to lazy constraints is a beneficial solution.

In the second type of selection, constrains 2.10 and 2.11 are added as Lazy Constraints, as well as 2.28. In this case, the LP relaxation of the model in exactly the same as it for a traditional dynamic facility location, regardless of the congestion facts. The polyhedron of the LP relaxation would be identical with it of the Generalized Modular Formulation (GMF), which is addressed by Jena, Cordeau, and Gendron, 2015 and the tightness of the polyhedron is demonstrated by them. Table 4.12 illustrates the comparison of these two types of selection of lazy constraints and their impact on the CPLEX execution. Clearly, it is only the first selection that leads to an advantageous solution. As a result, constraints 2.28 are considered as lazy constraints in the algorithm.

TABLE 4.10: Original costs obtained by the iterative method

Test Problems					MIP solved exactly with $\varepsilon = 10^{-1}$			MIP solved exactly with $\varepsilon = 10^{-3}$			MINLP solved iteratively	
$\mid T \mid$	$\mid I \mid , \mid J \mid$	Holding Cost (h)	Demands Type	Map Type	MINLP LB value	MINLP UB value	CPU time (sec.)	MINLP LB value	MINLP UB value	CPU time (sec.)	obtained cost	CPU time (sec.)
4	250 , 50	100000	1	1	5,659,630	6,520,120	3248	6,357,050	**6,359,350**	3087	6,376,540	1377
				2	6,325,920	7,163,870	557	7,033,280	**7,034,980**	705	7,065,400	250
				3	6,965,830	7,803,120	376	7,737,360	7,739,130	780	**7,738,930**	283
			2	1	5,518,580	6,151,360	2562	6,025,820	26,465,200	2761	**6,027,520**	2176
				2	6,133,130	6,769,450	1078	6,640,160	6,642,110	559	**6,641,880**	404
				3	6,720,880	7,366,880	500	7,315,410	7,316,910	607	**7,316,830**	346
		150000	1	1	5,846,200	7,136,940	3711	6,809,920	54,700,800	time	**6,992,620**	time
				2	6,514,930	7,772,600	845	7,662,390	**7,664,750**	3123	7,738,810	744
				3	7,156,350	8,447,510	454	8,368,670	**8,371,040**	3221	8,387,820	711
			2	1	5,611,180	6,682,510	time	6,425,340	**6,533,840**	time	6,620,070	time
				2	6,308,730	15,026,000	time	7,234,490	**7,236,890**	2425	7,253,050	2102
				3	6,944,800	8,005,300	3837	7,914,250	7,916,920	1790	**7,916,650**	1140
	250 , 100	100000	1	1	5,566,930	6,448,380	time	6,226,770	6,498,130	time	**6,306,190**	time
				2	6,247,090	7,101,460	1901	6,941,910	**6,943,620**	3449	6,948,600	1450
				3	6,916,200	**7,753,300**	1606	7,687,690	68,159,800	5198	18,613,100	1841
			2	1	5,396,080	6,113,060	time	5,886,620	**6,035,500**	time	6,079,200	time
				2	6,064,600	6,690,280	4899	6,557,620	**6,559,270**	3920	6,561,810	3305
				3	6,684,980	7,334,350	5763	7,271,670	7,273,490	3409	**7,273,370**	2016
		150000	1	1	5,713,730	15,104,400	time	6,597,810	7,318,990	time	**7,070,720**	time
				2	6,435,370	7,725,600	3008	7,471,870	**7,696,440**	time	12,671,900	4835
				3	7,105,020	8,396,490	2326	8,226,520	8,333,920	time	**8,333,660**	6993
			2	1	5,556,610	**6,609,560**	time	6,233,850	9,141,880	time	6,793,560	time
				2	6,177,540	7,288,670	time	7,087,600	7,371,630	time	**7,195,080**	time
				3	6,790,140	7,983,690	time	7,784,650	7,937,270	time	**7,870,050**	6711
	500 , 100	100000	1	1	8,079,790	8,890,330	2174	8,844,800	8,851,490	time	**8,851,450**	3611
				2	9,277,790	10,050,800	1495	10,045,200	10,047,100	3112	**10,047,000**	1849
				3	10,470,000	11,228,000	1321	11,215,200	11,217,200	2422	**11,217,100**	1757
			2	1	7,715,570	8,352,780	4397	8,326,400	8,332,290	time	**8,332,240**	4470
				2	8,844,960	9,527,400	2819	9,515,170	**9,516,590**	5915	9,544,110	2837
				3	9,904,730	10,589,700	1560	10,573,900	10,575,500	2565	**10,575,500**	1834
		150000	1	1	8,260,450	9,484,840	2236	9,161,890	9,682,360	time	**9,443,300**	time
				2	9,457,090	10,616,600	1500	10,593,700	10,613,200	time	**10,613,100**	3368
				3	10,654,300	11,799,300	1950	11,781,200	11,784,100	6574	**11,784,100**	2894
			2	1	7,877,660	**8,974,220**	time	8,686,580	8,981,570	time	96,955,300	time
				2	9,023,890	**10,050,300**	3724	10,008,100	10,398,500	time	10,089,700	6286
				3	10,079,900	11,113,300	1968	11,080,700	11,096,400	time	**11,096,400**	3080
						average:	**3517**		**average:**	**4945**	**average:**	**3708**
8	250 , 50	100000	1	1	10,501,400	**12,224,100**	5444	11,970,400	14,464,300	time	14,913,400	5117
				2	11,783,600	13,486,400	1159	13,326,300	**13,329,500**	4465	13,356,100	1645
				3	13,113,300	14,796,500	1194	14,726,300	**14,729,700**	3978	14,730,600	990
			2	1	10,159,600	13,762,400	time	11,239,900	18,228,900	time	**11,322,100**	time
				2	11,418,600	12,691,400	2793	12,521,800	27,844,800	6212	**12,679,300**	5758
				3	12,606,400	13,901,600	3281	13,790,700	**13,794,400**	3499	13,798,500	3969
		150000	1	1	10,804,000	13,586,500	time	12,846,000	13,627,000	time	**13,335,500**	time
				2	12,161,100	14,715,500	1809	14,363,300	**14,668,100**	time	14,673,500	5713
				3	13,495,400	16,042,800	1663	15,809,000	16,003,800	time	**16,003,400**	4489
			2	1	10,469,100	12,887,800	time	11,947,900	13,020,800	time	**12,695,900**	time
				2	11,703,400	13,937,500	time	13,483,200	18,680,200	time	**13,800,500**	time
				3	12,895,300	15,105,800	time	14,804,100	**14,969,000**	time	15,003,500	time
	250 , 100	100000	1	1	10,301,000	**12,201,300**	time		memory			memory
				2	11,620,900	13,398,900	time	12,751,600	**13,176,700**	time	20,649,600	time
				3	12,987,400	**14,746,600**	time	14,239,500	16,027,000	time		memory
			2	1	9,962,890	**12,081,900**	time		memory			memory
				2	11,244,100	**12,534,100**	time		memory		13,030,300	time
				3	12,444,300	13,801,000	time	13,437,900	**13,764,200**	time	13,787,300	time
		150000	1	1		memory			memory			memory
				2	11,988,800	14,719,500	time		memory		**14,537,400**	time
				3	13,347,600	16,242,900	time		memory		**16,097,000**	time
			2	1		memory			memory			memory
				2		memory			memory		**15,457,000**	time
				3	12,778,400	15,022,800	time	13,687,600	**14,975,500**	time		memory
	500 , 100	100000	1	1	15,353,400	16,982,400	6052	16,220,400	**16,914,300**	time	16,915,200	time
				2	17,596,700	19,159,200	4074		memory		**19,149,800**	time
				3	19,800,600	**21,305,700**	4912	20,775,600	21,306,800	time	21,306,800	time
			2	1	14,573,200	15,870,800	time		memory		**15,845,700**	time
				2	16,650,500	17,971,800	time	17,419,800	**17,946,000**	time	17,951,600	time
				3	18,663,000	20,021,600	6875	19,610,000	**19,986,400**	time	19,991,000	time
		150000	1	1	15,696,600	18,194,900	time		memory		**18,107,500**	time
				2	17,958,300	20,302,100	3884	18,887,100	**20,292,600**	time	20,294,600	time
				3	20,162,200	**22,419,900**	4478	21,123,800	22,429,900	time	22,430,200	time
			2	1	14,875,600	**16,979,200**	time		memory			memory
				2	16,980,000	19,086,000	time	17,576,100	**18,998,700**	time	19,004,800	time
				3	18,993,800	**21,026,100**	6042	19,788,100	21,029,500	time	21,036,000	time
						average:	**5891**		**average:**	**6905**	**average:**	**6569**

(For all cases; $\mid K \mid = 5$, $C_s = 0.5$)

(Time limitation is 7200 seconds.)

TABLE 4.11: Original costs of the most difficult test problems obtained by various solution method

Test Problems				MIP solved exactly with $\varepsilon = 10^{-1}$			MIP solved exactly with $\varepsilon = 10^{-3}$			MINLP solved iteratively		MIP solved heuristically with $\Omega = 6$	
$\mid T \mid$	$\mid I \mid , \mid J \mid$	Holding Cost (h)	Demands Type	MINLP LB value	MINLP UB value	CPU time (sec.)	MINLP LB value	MINLP UB value	CPU time (sec.)	obtained cost	CPU time (sec.)	obtained MINLP value	CPU time (sec.)
4	250 , 100	100000	1	6,916,200	7,753,300	1606	7,687,690	68,159,800	5198	69,749,300	1841	7,751,940	1144
			2	6,684,980	7,334,350	5763	7,271,670	7,273,490	3409	7,273,380	2016	7,653,970	2177
		150000	1	7,105,020	8,396,490	2326	8,226,520	8,333,920	time	8,333,660	6993	8,396,490	1379
			2	6,790,140	7,983,690	time	7,784,650	7,937,270	time	7,870,050	6711	7,914,620	1223
	500 , 100	100000	1	10,470,000	11,228,000	1321	11,215,200	11,217,200	2422	11,217,100	1757	11,228,000	1319
			2	9,904,730	10,589,700	1560	10,573,900	10,575,500	2565	10,575,500	1834	10,589,700	1552
		150000	1	10,654,300	11,799,300	1950	11,781,200	11,784,100	6574	11,784,100	2894	11,799,300	1700
			2	10,079,900	11,113,300	1968	11,080,700	11,096,400	time	11,096,400	3080	11,113,300	1956
	1000 , 100	100000	1	16,488,800	17,189,400	3892		memory		17,189,400	5004	17,189,400	3733
			2	15,528,700	16,174,700	5138		memory		21,562,300	4461	16,174,700	4984
		150000	1	16,668,500	17,719,400	5000		memory		40,792,600	time	17,719,400	4878
			2	15,692,900	16,662,000	5791		memory		17,020,600	time	memory	
					average:	**3626**		**average:**	**5881**	**average:**	**4249**	**average:**	**2770**
8	250 , 100	100000	1	12,987,400	14,746,600	time	14,239,500	16,027,000	time	memory		98,771,500	6946
			2	12,444,300	13,801,000	time	13,437,900	13,764,200	time	13,787,300	time	13,801,000	4920
		150000	1	13,347,600	16,242,900	time		memory		16,097,000	time	52,230,100	5598
			2	12,778,400	15,022,800	time	13,687,600	14,975,500	time	memory		15,022,800	time
	500 , 100	100000	1	19,800,600	21,305,700	4912	20,775,600	21,306,800	time	21,306,800	time	21,305,700	4972
			2	18,663,000	20,021,600	6875	19,610,000	19,986,400	time	19,991,000	time	memory	
		150000	1	20,162,200	22,419,900	4478	21,123,800	22,429,900	time	22,430,200	time	22,419,900	4651
			2	18,993,800	21,026,100	6042	19,788,100	21,029,500	time	21,036,000	time	21,034,700	5371
	1000 , 100	100000	1		memory			memory		memory		memory	
			2		memory			memory		memory		memory	
		150000	1		memory			memory		memory		memory	
			2		memory			memory		memory		memory	
					average:	**6659**		**average:**	**7200**	**average:**	**7200**	**average:**	**6305**

(For all cases; $\mid K \mid = 5$, $C_s = 0.5$, Map Type=3)

(Time limitation is 7200 seconds.)

4.4.2 CPLEX Configurations

Having observed the CPLEX performance in solving a wide range of test problems, we are able to tune some parameter of CPLEX MIP solver for a problem-specific configuration.

4.4.2.1 CPLEX cuts

By the default, CPLEX automatically chooses the number of each type of cuts added to the branch-and-cut. As is noted, no other types of cuts from *MIR*, *Gomory Fractional*, *Flow Cover* and *Implied Bounding* is applicable for our formulation. Consequently, we disable generation the non-applicable cuts to save the corresponding separation times in the problem-specific configurations. Furthermore, it is observed that the numbers of generated *Gomory Fractional* and *MIR* cuts are many more than them for *Flow Cover* and *Implied Bound* cuts. Therefore, we set *Gomory Fractional* and *MIR* cuts generation to aggressive mode and *Implied Bound* cuts to moderate. Generation of *Flow Cover* cuts remains automatic by CPLEX. This problem-specific configuration, illustrated below, is the best found cut configuration among many customized configurations and is gained try and false on a bunch of test problems, while, in CPLEX default, all of these parameter are equal to 0 (automatic).

$$\text{CPLEX cuts parameters} \begin{cases} Gomory\ Fractional = 2 & \text{(aggressive)} \\ MIR = 2 & \text{(aggressive)} \\ Implied\ Bound = 1 & \text{(moderate)} \\ Flow\ Cover = 0 & \text{(automatic)} \\ others = -1 & \text{(disabled)} \end{cases}$$

Table 4.13 demonstrates the efficiency of the problem-specific configuration. As a result, the best know configuration is considered as the setting of the algorithm.

4.4.2.2 CPLEX heuristics

Having inspected the CPLEX performance, tuning the root node is so difficult for the large-sized test problems. Moreover, most of this time is spent to implement CPLEX heuristics in search of an UB. As a consequence, we disable CPLEX heuristics in the root and enable them again in the node 1, while the frequency

TABLE 4.12: Impact of having lazy constraints

Test Problems			Optimal Solution Value	all constraints are normal					inequalities 2.28 are lazy constraints					inequalities 2.28, 2.10 and 2.11 are lazy constraints				
$\mid I \mid , \mid J \mid$	Holding Cost (h)	Demands Type		MIP LB value	MIP UB value	MIP gap	CPU time (sec.)	Nodes	MIP LB value	MIP UB value	MIP gap	CPU time (sec.)	Nodes	MIP LB value	MIP UB value	MIP gap	CPU time (sec.)	Nodes
50 , 50	50000	1	2,838,580		optimal		45	3		optimal		32	48		optimal		97	49
		2	2,861,670		optimal		91	39		optimal		91	150		optimal		135	174
	75000	1	3,070,350		optimal		121	44		optimal		98	137		optimal		166	150
		2	3,111,200		optimal		375	816		optimal		345	2317		optimal		705	2851
250 , 50	50000	1	7,088,550		optimal		230	0		optimal		188	3		optimal		115	4
		2	6,715,500		optimal		178	0		optimal		130	8		optimal		110	11
	75000	1	7,413,160		optimal		379	9		optimal		273	13		optimal		322	16
		2	7,015,600		optimal		253	0		optimal		220	1		optimal		186	2
250 , 100	50000	1	7,043,580		optimal		831	14		optimal		864	8		optimal		917	10
		2	6,674,100		optimal		717	11		optimal		509	57		optimal		985	58
	75000	1	7,365,700		optimal		1618	16		optimal		1473	53		optimal		2403	56
		2	6,973,230		optimal		965	6		optimal		1013	5		optimal		1327	7
						average:	**484**				**average:**	**436**				**average:**	**622**	

(For all cases; $\mid K \mid = 5$, $\mid T \mid = 4$, $C_s = 0.5$, $\varepsilon = 10^{-3}$, Map Type=3)

(Time limitation is 7200 seconds.)

TABLE 4.13: Impact of customizing CPLEX configuration

Test Problems				**CPLEX default configuration**						**CPLEX heuristics are disables in root node**					**customized problem-specific cut configuration**					
$\mid I \mid , \mid J \mid$	**Holding Cost (h)**	**Demands Type**	**Optimal Solution Value**	**MIP LB value**	**MIP UB value**	**MIP gap**	**CPU time (sec.)**	**Nodes**	**number of generated cuts by CPLEX**	**MIP LB value**	**MIP UB value**	**MIP gap**	**CPU time (sec.)**	**Nodes**	**MIP LB value**	**MIP UB value**	**MIP gap**	**CPU time (sec.)**	**Nodes**	**number of generated cuts by CPLEX**
50 , 50	50000	1	2,838,580		optimal		32	48	379		optimal		50	15		optimal		29	12	432
		2	2,861,670		optimal		91	150	349		optimal		155	227		optimal		84	118	368
	75000	1	3,070,350		optimal		98	137	545		optimal		111	39		optimal		95	12	675
		2	3,111,200		optimal		345	2317	428		optimal		707	1489		optimal		307	977	488
250 , 50	50000	1	7,088,550		optimal		188	3	263		optimal		238	5		optimal		168	0	228
		2	6,715,500		optimal		130	8	217		optimal		186	8		optimal		113	0	219
	75000	1	7,413,160		optimal		273	13	391		optimal		374	37		optimal		245	10	350
		2	7,015,600		optimal		220	1	426		optimal		314	26		optimal		218	0	365
250 , 100	50000	1	7,043,580		optimal		864	8	354		optimal		834	25		optimal		829	10	351
		2	6,674,100		optimal		509	57	143		optimal		859	39		optimal		484	11	143
	75000	1	7,365,700		optimal		1473	53	555		optimal		1995	135		optimal		1458	74	619
		2	6,973,230		optimal		1013	5	608		optimal		1522	41		optimal		962	10	587
						average:	**436**					**average:**	**612**				**average:**	**416**		

(For all cases; $\mid K \mid = 5$, $\mid T \mid = 4$, $C_s = 0.5$, $\varepsilon = 10^{-3}$, Map Type=3)

(Time limitation is 7200 seconds.)

to apply the heuristics is chosen by CPLEX (as its default). However, this idea leads to failure, because, in default setting, CPLEX re-optimizes the problem after finding an UB, repeats this procedure for several times and then, passes the root (or any other node). Subsequently, the LB is improved in these repetitions. Nonetheless, when the heuristics are disabled, no UB is found. Thus, the LP relaxations is solved only once and the LB is not improved repetitively. In other words, disabling the CPLEX heuristics does not only skip the UB, but also, deteriorates the LB obtained at the node. Table 4.13 shows that disabling CPLEX heuristics in the root is a failed idea.

4.5 Interpretation of *Big-M*

As is mentioned in section 2.3.1, in the nonlinear MIP model, capacity constraints (2.3) are never biding, because of 2.1 involved in the congestion cost expression of the objective function of the nonlinear model (**??**). It states that no facility utilizes all of its capacity during time periods.

In the linearized model, in spite of nonnegativity of decision variables z_{jkt} and w_{jkt} which are involved in the congestion cost expression of the objective function of the linearized model (2.12), the fact remains guaranteed by constraints 2.19.

In the ideal case of approximation, $w_{jkt} = \frac{z_{jkt}}{1-z_{jkt}}$. Thus, having considered constraints 2.19, we have

$$w_{jkt} \leq M \quad \xRightarrow{\text{ignoring approximation gap}} \quad \frac{z_{jkt}}{1 - z_{jkt}} \leq M \;\Rightarrow\; z_{jkt} \leq \frac{M}{M+1} < 1$$

Hence, proposing *Big-M* as upper bound for w_jkt implicitly ensures the steady state conditions of congested queues.

In this paper, $M = 100$ that states facilities are allowed to utilize at most 99% of their capacities.

4.6 Sensitivity Analysis

By lifting the set of constraints 2.16 and defining them as *lazy constraints*, beside of adding valid inequalities 2.23 and customizing CPLEX configurations as specified in section 4.4.2.1, we solve the model illustrated in the page 37 as quick as possible.

To analyze the sensitivity of the MIP (linearized) model on different input parameters, a huge bunch of test problems are solved by the exact method presented in section 3.1. As table 4.14 demonstrates, difficulty of the MIP model has a strong correlation with the number of time periods ($\mid T \mid$). The second prior factor that burdens the model execution is the size of the test problem. The execution is more time-taking for instances with high number of customers ($\mid I \mid$) and candidate locations ($\mid J \mid$). It is also noticed that variation in number of potential capacity levels ($\mid K \mid$) does change the corresponding optimal solution value. However, increasing $\mid K \mid$ impedes the execution slightly.

Table 4.15 is a selection of several rows of table 4.14 with more information of the optimal solution.

TABLE 4.14: Sensitivity analysis of the MIP model

Test Problems				**Solved exactly with $\varepsilon = 10^{-3}$**			
$\mid I \mid , \mid J \mid$	**Capacity Levels Number ($\mid K \mid$)**	**Time Periods Number ($\mid T \mid$)**	**Holding Cost (h)**	**MIP LB value**	**MIP UB value**	**MINLP UB**	**CPU time (sec.)**
50 , 10	3	4	25000	optimal	2,664,510	2,664,970	2
			50000	optimal	2,947,400	2,948,350	1
			75000	optimal	3,177,900	3,178,860	3
			100000	optimal	3,389,170	3,390,050	7
			150000	optimal	3,697,140	3,698,230	6
		8	25000	optimal	4,994,280	4,995,450	2
			50000	optimal	5,596,150	5,597,680	6
			75000	optimal	6,037,230	6,038,670	12
			100000	optimal	6,378,520	6,380,690	4
			150000	optimal	7,041,190	7,042,590	49
	10	4	25000	optimal	2,664,510	2,664,970	3
			50000	optimal	2,947,400	2,948,350	6
			75000	optimal	3,177,900	3,178,860	5
			100000	optimal	3,389,170	3,390,050	14
			150000	optimal	3,638,730	3,639,310	21
		8	25000	optimal	4,994,280	4,995,450	8
			50000	optimal	5,596,150	5,597,680	30
			75000	optimal	6,037,230	6,038,670	54
			100000	optimal	6,378,520	6,380,690	29
			150000	optimal	7,041,190	7,042,590	898
50 , 50	3	4	25000	optimal	2,522,470	2,523,360	20
			50000	optimal	2,838,580	2,839,310	43
			75000	optimal	3,070,350	3,071,380	69
			100000	optimal	3,283,420	3,284,470	173
			150000	optimal	3,609,420	3,611,450	586
		8	25000	optimal	4,759,210	4,760,120	39
			50000	optimal	5,404,570	5,406,400	835
			75000	optimal	5,860,210	5,861,480	1056
			100000	optimal	6,206,000	6,208,190	1632
			150000	optimal	6,770,310	6,772,470	2543
	10	4	25000	optimal	2,522,470	2,523,360	39
			50000	optimal	2,838,580	2,839,310	256
			75000	optimal	3,070,350	3,071,380	134
			100000	optimal	3,283,420	3,284,470	428
			150000	optimal	3,609,420	3,611,450	5084
		8	25000	optimal	4,759,210	4,760,120	762
			50000	optimal	5,404,570	5,406,400	1071
			75000	optimal	5,860,210	5,861,480	5565
			100000	6,122,550	6,206,000	6,208,190	time
			150000	6,558,850	6,788,040	361,212,000	time
250 , 50	3	4	25000	optimal	6,763,240	6,763,520	123
			50000	optimal	7,088,550	7,089,300	205
			75000	optimal	7,413,160	7,414,410	215
			100000	optimal	7,737,360	7,739,130	492
			150000	optimal	8,368,670	8,371,040	1361
		8	25000	optimal	12,817,900	12,818,500	355
			50000	optimal	13,455,100	13,456,500	418
			75000	optimal	14,091,600	14,093,900	982
			100000	optimal	14,726,300	60,242,400	1906
			150000	15,856,800	15,960,300	15,966,000	time
	10	4	25000	optimal	6,763,240	6,763,520	241
			50000	optimal	7,088,550	7,089,300	383
			75000	optimal	7,413,160	7,414,410	718
			100000	optimal	7,737,360	7,739,130	1129
			150000	8,269,770	8,368,670	80,300,100	time
		8	25000	optimal	12,817,900	12,818,500	927
			50000	optimal	13,455,100	13,456,500	1740
			75000	optimal	14,091,600	14,093,900	3221
			100000	14,665,200	14,727,500	14,731,100	time
			150000	15,525,900	15,998,200	16,003,800	time

250 , 100	3	4	25000	optimal	6,717,940	6,718,290	488
			50000	optimal	7,043,580	7,044,460	700
			75000	optimal	7,365,700	7,366,870	1161
			100000	optimal	7,687,690	26,949,500	2771
			150000	8,241,830	8,331,470	8,333,920	time
		8	25000	optimal	12,741,200	12,742,000	1506
			50000	optimal	13,380,300	13,382,300	3535
			75000	13,989,500	14,011,700	14,014,800	time
			100000	14,467,600	16,022,100	16,027,000	time
			150000	14,695,100	18,610,000	18,617,000	time
	10	4	25000	optimal	6,717,940	6,718,290	990
			50000	optimal	7,043,580	7,044,460	1533
			75000	optimal	7,365,700	7,366,870	2851
			100000	7,669,000	7,687,690	7,689,290	time
			150000	8,163,820	8,415,560	8,418,630	time
		8	25000	optimal	12,741,200	12,742,000	3007
			50000	optimal	13,380,300	185,563,000	6510
			75000	13,725,700	15,277,000	15,279,900	time
			100000	14,025,400	16,022,100	16,027,000	time
			150000	14,300,500	17,507,200	17,514,100	time
500 , 100	3	4	25000	optimal	10,366,000	10,366,500	1213
			50000	optimal	10,649,100	10,650,000	1474
			75000	optimal	10,932,200	10,933,700	1591
			100000	optimal	11,215,200	11,217,200	1708
			150000	optimal	11,781,200	11,784,100	3382
		8	25000	optimal	19,620,800	19,621,600	4674
			50000	optimal	20,181,500	20,183,200	5472
			75000	20,729,200	20,742,200	20,744,900	time
			100000	21,110,700	21,303,200	21,306,800	time
			150000	21,341,200	22,424,500	22,429,900	time
	10	4	25000	optimal	10,366,000	10,366,500	2283
			50000	optimal	10,649,100	10,650,000	2425
			75000	optimal	10,932,200	10,933,700	2520
			100000	optimal	11,215,200	11,217,200	3202
			150000	11,745,800	11,781,200	11,784,100	time
		8	25000	optimal	19,620,800	19,621,600	7126
			50000		memory		
			75000		memory		
			100000		memory		
			150000		memory		
1000 , 100	3	4	25000	optimal	16,393,900	16,394,500	4329
			50000	optimal	16,658,000	16,659,300	4858
			75000	optimal	16,922,100	16,924,000	6623
			100000	17,182,800	17,188,700	17,191,200	time
			150000		memory		
		8	25000		memory		
			50000		memory		
			75000		memory		
			100000		memory		
			150000		memory		
	10	4	25000		memory		
			50000		memory		
			75000		memory		
			100000		memory		
			150000		memory		
		8	25000		memory		
			50000		memory		
			75000		memory		
			100000		memory		
			150000		memory		

(For all cases; $C_s = 0.5$, Demand Type=1 , Map Type=3)

(Time limitation is 7200 seconds.)

TABLE 4.15: Impact of the congestion multiplier on the MIP model optimal solution

Test Problems		**Solved exactly with $\varepsilon = 10^{-3}$**							
$\mid I \mid , \mid J \mid$	**Holding Cost (h)**	**average utilization of facilities**	**Location cost**	**Location cost / Total cost**	**Transportation cost**	**Transportation cost / Total cost**	**Linearized congestion cost**	**Linearized congestion cost / Total cost**	**Total cost (MIP UB)**
50 , 50	25000	0.65	912,000	36 %	1,173,910	47 %	436,567	17 %	2,522,470
	50000	0.49	1,175,400	41 %	1,186,690	42 %	476,488	17 %	2,838,580
	75000	0.48	1,175,400	38 %	1,209,730	39 %	685,211	22 %	3,070,350
	100000	0.39	1,438,800	44 %	1,184,630	36 %	659,988	20 %	3,283,420
	150000	0.39	1,438,800	40 %	1,200,370	33 %	970,244	27 %	3,609,420
250 , 50	25000	0.31	2,432,000	36 %	4,005,430	59 %	325,816	5 %	6,763,240
	50000	0.31	2,432,000	34 %	4,007,090	57 %	649,452	9 %	7,088,550
	75000	0.31	2,432,000	33 %	4,008,230	54 %	972,930	13 %	7,413,160
	100000	0.31	2,432,000	31 %	4,008,790	52 %	1,296,580	17 %	7,737,360
	150000	0.28	2,736,000	33 %	3,763,920	45 %	1,868,750	22 %	8,368,670
250 , 100	25000	0.36	2,128,000	32 %	4,248,860	63 %	341,080	5 %	6,717,940
	50000	0.31	2,432,000	35 %	3,966,590	56 %	644,997	9 %	7,043,580
	75000	0.31	2,432,000	33 %	3,967,520	54 %	966,182	13 %	7,365,700
	100000	0.12	2,432,000	32 %	3,968,030	52 %	1,287,660	17 %	7,687,690
	150000	0.31	2,432,000	29 %	3,968,810	48 %	1,930,660	23 %	8,331,470
500 , 100	25000	0.22	3,344,000	32 %	6,738,930	65 %	283,079	3 %	10,366,000
	50000	0.22	3,344,000	31 %	6,738,930	63 %	566,157	5 %	10,649,100
	75000	0.22	3,344,000	31 %	6,739,050	62 %	849,113	8 %	10,932,200
	100000	0.22	3,344,000	30 %	6,739,050	60 %	1,132,150	10 %	11,215,200
	150000	0.22	3,344,000	28 %	6,739,460	57 %	1,697,790	14 %	11,781,200

(For all cases; $\mid K \mid = 3$, $\mid T \mid = 4$, $C_s = 0.5$, Demand Type=1 , Map Type=3)

(Time limitation is 7200 seconds.)

List of Figures

List of Tables

List of Abbreviations

ADC	**A**ggregated **D**emands **C**onstraint
BTO	**B**uild **T**o **O**rder
DC	**D**istribution **C**entre
DND	**D**istribution **N**etwork **D**esign
FLP	**F**acility **L**ocation **P**roblem
GMC	**G**eneralized **M**odular **C**apacities
MIR	**M**ixed **I**nteger **R**ounding
MINLP	**M**ixed **I**nteger **N**on **L**inear **P**rogramming
MIP	**M**ixed **I**nteger **P**rogramming
LB	**L**ower **B**ound
LP	**L**inear **P**rogramming
LR	**L**agrangean **R**elaxation
SCM	**S**upply **C**hain **M**anagement
SI	**S**trong **I**nequalities
UB	**U**pper **B**ound
WIP	**W**ork **I**n **P**rocess

A : Linearization Accuracy

A priori set of points $\{R^h\}_{h\in H}$ could be generated for the function $\rho_{jt}(R_{jt})$ in such a way that the piecewise linear approximation $\widehat{\rho_{jt}}(R_{jt})$ satisfies

$$0 \leq \widehat{\rho_{jt}}(R_{jt}) - \rho_{jt}(R_{jt}) \leq \varepsilon$$

for all $R_{jt} \geq 0$ and $\varepsilon > 0$ (Elhedhli, 2005) that is called *outer linearization*.

Suppose that the piecewise linear approximation $\widehat{\rho_{jt}}(R_{jt})$ has breakdowns at $R^0 = 0,\ R^1,\ R^2,\ ...,\ R^n_{(\varepsilon)}$ and is tangent to $\rho_{jt}(R_{jt})$ at points $r^1,\ r^2,\ r^3,\ ...,\ r^n_{(\varepsilon)}$ where $R_{l-1} \leq r_l \leq R_l$. The values of R_l and r_l are recursively determined using the fact that $\widehat{\rho_{jt}}(R_{jt})$ is linear between $R_{l-1} \leq r_l \leq R_l$ with a slope of $\acute{\rho}_{jt}(r_l)$. Therefore, given $\widehat{\rho_{jt}}(R_{l-1})$ and R_{l-1} and using the fact that $\rho_{jt}(r_l) = r_l/(1+r_l)^2$, we use the line equation

$$\widehat{\rho_{jt}}(R_{l-1}) = \frac{r_l}{1+r_l} + \frac{1}{(1+r_l)^2}(R_{l-1} - r_l)$$

to find r_l. Then using r_l and $\widehat{\rho_{jt}}(R_l) = \rho_{jt}(R_l) + \varepsilon = R_l/(1+R_l) + \varepsilon$, we again use the line equation

$$\widehat{\rho_{jt}}(R_l) = \frac{r_l}{1+r_l} + \frac{1}{(1+r_l)^2}(R_l - r_l)$$

to find R_l. Finding r_l values to solving for the positive root of

$$[\widehat{\rho_{jt}}(R_{l-1}) - 1]r_l^2 + [2\,\widehat{\rho_{jt}}(R_{l-1})]r_l + \widehat{\rho_{jt}}(R_{l-1}) - R_{l-1} = 0$$

leading to

$$r_l = \frac{-\widehat{\rho_{jt}}(R_{l-1}) + \sqrt{\widehat{\rho_{jt}}(R_{l-1}) - R_{l-1}(1 - \widehat{\rho_{jt}}(R_{l-1}))}}{1 - \widehat{\rho_{jt}}(R_{l-1})}$$

Finding R_l values to solving for the positive root of

$$R_l^2 - [2\, r_l + \varepsilon\, (1 + r_l)^2] R_l + [r_l^2 - \varepsilon\, (1 + r_l)^2] = 0$$

leading to

$$R_l = [r_l + \frac{\varepsilon}{2}(1 + r_l)^2] + \sqrt{[r_l + \frac{\varepsilon}{2}(1 + r_l)^2]^2 - [r_l^2 - \varepsilon\, (1 + r_l)^2]}$$

If $\rho_{jt}(R_l) + \varepsilon \geq 1$, then $l = n(\varepsilon)$. In that case, a slightly better point R_l is chosen by setting $\widehat{\rho_{jt}}(R_l)$ to 1,

$$\widehat{\rho_{jt}}(R_l) = \frac{r_l}{1 + r_l} + \frac{1}{(1 + r_l)^2}(R_l - r_l) = 1$$

and

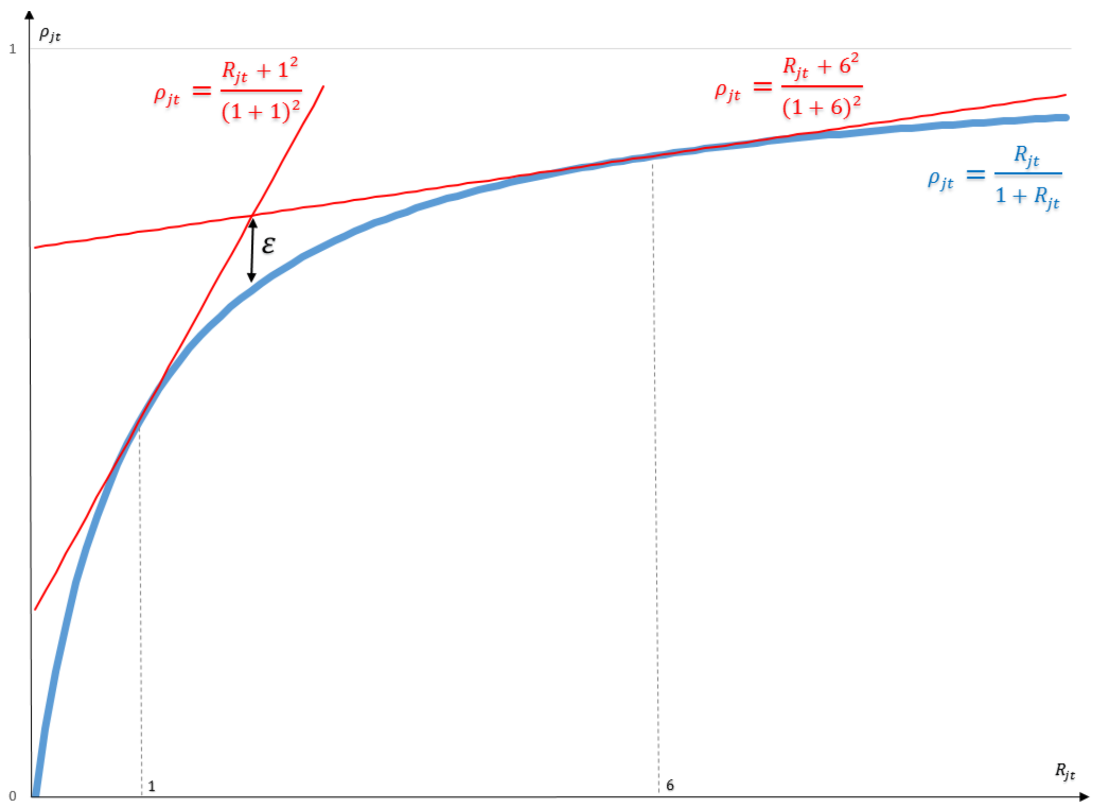
$$R_l = 1 + 2\, r_l = R_{n(\varepsilon)}$$

Finally, for $R \geq R_{n(\varepsilon)}$, $\widehat{\rho_{jt}}(R) = 1$.

In summary, we set ε as an input to the model. Then indices $\{R^h\}_{h \in H}$ are calculated as shown above to build constraints 2.16. Moreover, regarding $\mid H \mid = n(\varepsilon)$, the number of constraints 2.16 could be obtained by $\mid J \mid * \mid H \mid * \mid T \mid$. Similarly, it is possible to propose an approximation which satisfies

$$0 \leq \rho_{jt}(R_{jt}) - \widehat{\rho_{jt}}(R_{jt}) \leq \varepsilon$$

which is called *inner linearization*. However, it this paper, an outer linearization is executed as explained above.



ρ_{jt}
1
$\rho_{jt} = \frac{R_{jt} + 1^2}{(1+1)^2}$
$\rho_{jt} = \frac{R_{jt} + 6^2}{(1+6)^2}$
$\rho_{jt} = \frac{R_{jt}}{1+R_{jt}}$
ε
0
1
6
R_{jt}

B : Mathematical Proof of User-cuts

B.1 Validity Proof of Mixed-Dicuts:

Proof of inequalities 3.1:

Having S_1 and S_2 as subsets of I and J, respectively, for each t in T, we can have;

$$\sum_{i\in S_1}\lambda_{it}=\sum_{i\in S_1}\lambda_{it}\ \xrightarrow[\sum_{j\in J}x_{ijt}=1]{eq\ 2.15:}\ \sum_{i\in S_1}\sum_{j\in J}\lambda_{it}\,x_{ijt}=\sum_{i\in S_1}\lambda_{it}$$
$$\Longrightarrow\ \sum_{i\in S_1}\sum_{j\in S_2}\lambda_{it}\,x_{ijt}+\sum_{i\in S_1}\sum_{j\in J\setminus S_2}\lambda_{it}\,x_{ijt}=\sum_{i\in S_1}\lambda_{it} \tag{B.1}$$
$$\xrightarrow[\sum_{i\in I}\lambda_{it}\,x_{ijt}=\sum_{k\in K}\mu_{jkt}\,z_{jkt}]{eq\ 3.3.1:}\ \sum_{i\in S_1}\sum_{j\in S_2}\lambda_{it}\,x_{ijt}+\sum_{j\in J\setminus S_2}\sum_{k\in K}\mu_{jkt}\,z_{jkt}=\sum_{i\in S_1}\lambda_{it}$$
$$\xrightarrow[z_{jkt}\le\sum_{\acute{k}\in K}y_{j\acute{k}kt}]{eq\ 2.10:}\ \sum_{i\in S_1}\sum_{j\in S_2}\lambda_{it}\,x_{ijt}+\sum_{j\in J\setminus S_2}\sum_{k\in K}\mu_{jkt}\sum_{\acute{k}\in K}y_{j\acute{k}kt}\ge\sum_{i\in S_1}\lambda_{it} \tag{B.2}$$

In addition, by having equation B.1, it could be proven;

$$eq\ B.1:\ \sum_{i\in S_1}\sum_{j\in S_2}\lambda_{it}\,x_{ijt}+\sum_{i\in S_1}\sum_{j\in J\setminus S_2}\lambda_{it}\,x_{ijt}=\sum_{i\in S_1}\lambda_{it}$$
$$\xrightarrow[x_{ijt}\le\sum_{\acute{k}\in K}\sum_{k\in K\setminus\{0\}}y_{j\acute{k}kt}]{eq\ 2.23:}\ \sum_{i\in S_1}\sum_{j\in S_2}\lambda_{it}\,x_{ijt}+\sum_{i\in S_1}\sum_{j\in J\setminus S_2}\lambda_{it}\sum_{\acute{k}\in K}\sum_{k\in K\setminus\{0\}}y_{j\acute{k}kt}\ge\sum_{i\in S_1}\lambda_{it}$$
$$\xrightarrow{\sum_{\acute{k}\in K}y_{j\acute{k}0t}\ge 0}\ \sum_{i\in S_1}\sum_{j\in S_2}\lambda_{it}\,x_{ijt}+\sum_{j\in J\setminus S_2}\sum_{k\in K}\sum_{i\in S_1}\lambda_{it}\sum_{\acute{k}\in K}y_{j\acute{k}kt}\ge\sum_{i\in S_1}\lambda_{it} \tag{B.3}$$

Having equations B.2 and B.3 results in inequalities 3.1.

B.2 Validity Proof of Enhanced-MIR Cuts:

Firstly, lets prove the validity of these inequalities which are called *enhanced-MIR* in the literature (Bodur and Luedtke, 2016).

Having $u+\eta_1 \geq b_1$ and $u+\eta_2 \geq b_2$ while $u \geq 0$, $\eta_1 \in Z$, $\eta_2 \in Z$ and $(b_2-b_1) \notin Z$, the inequality

$$u \geq f(b_2-b_1)\left(\lceil b_2-b_1 \rceil-(\eta_2-\eta_1)\right)+b_1-\eta_1 \tag{B.4}$$

is valid for the integral set where $f(b_2-b_1)=b_2-b_1-\lfloor b_2-b_1 \rfloor>0$ (Bodur and Luedtke, 2016).

Proof:
According to the *Proposition 8.6* of Wolsey, 1998, inequality $x \geq f(b)(\lceil b \rceil - y)$ is valid for any set $X^{\geq}=\{(x,y) \in R_{+}^{1} \times Z^{1} : x+y \geq b\}$ where $f(b)=b-\lceil b \rceil>0$.
Now, let $\varphi=u+\eta_1-b_1$, as $u+\eta_1 \geq b_1$, so $\varphi \geq 0$. In addition, $\eta_1 \in Z$ and $\eta_2 \in Z$, so $\eta_2-\eta_1 \in Z$. Thus, if $(b_2-b_1) \notin Z$, we can write

$$\varphi = u+\eta_1-b_1 \xrightarrow{u+\eta_2 \geq b_2} \varphi \geq b_2-\eta_2+\eta_1-b_1=(b_2-b_1)-(\eta_2-\eta_1)$$

$$\Longrightarrow \quad \varphi+(\eta_2-\eta_1) \geq (b_2-b_1) \quad when \begin{cases} \varphi \geq 0 \\ (\eta_2-\eta_1) \in Z \\ (b_2-b_1) \notin Z \end{cases}$$

$$\xrightarrow{\text{Proposition 8.6 of Wolsey, 1998}} \quad \varphi \geq f(b_2-b_1)\left(\lceil b_2-b_1 \rceil-(\eta_2-\eta_1)\right)$$

$$\xrightarrow{\varphi=u+\eta_1-b_1} \quad u \geq f(b_2-b_1)\left(\lceil b_2-b_1 \rceil-(\eta_2-\eta_1)\right)+b_1-\eta_1$$

Also, having $u+b_1 \geq \eta_1$ and $u+b_2 \geq \eta_2$ while $u \geq 0$, $\eta_1 \in Z$, $\eta_2 \in Z$ and $(b_2-b_1) \notin Z$, the inequality

$$u \geq (1-f(b_2-b_1))\left((\eta_2-\eta_1)-\lfloor b_2-b_1 \rfloor\right)+\eta_1-b_1 \tag{B.5}$$

is valid for the integral set where $f(b_2-b_1)=b_2-b_1-\lfloor b_2-b_1 \rfloor>0$.

Proof:
According to the *Corollary of Proposition 8.6* of Wolsey, 1998, inequality $x \geq (1-f(b))\,(y-\lfloor b \rfloor)$ is valid for any set $X^{\leq}=\{(x,y) \in R_{+}^{1} \times Z^{1} : x+b \geq y\}$ where $f(b)=b-\lceil b \rceil>0$.
Now, let $\varphi=u+b_1-\eta_1$, as $u+b_1 \geq \eta_1$, so $\varphi \geq 0$. In addition, $\eta_1 \in Z$ and $\eta_2 \in Z$,

so $\eta_2 - \eta_1 \in Z$. Thus, if $(b_2 - b_1) \notin Z$, we can write

$$\varphi = u + b_1 - \eta_1 \xrightarrow{u+b_2 \geq \eta_2} \varphi \geq \eta_2 - b_2 + b_1 - \eta_1 = (\eta_2 - \eta_1) - (b_2 - b_1)$$

$$\Longrightarrow \quad \varphi + (b_2 - b_1) \geq (\eta_2 - \eta_1) \quad when \begin{cases} \varphi \geq 0 \\ (\eta_2 - \eta_1) \in Z \\ (b_2 - b_1) \notin Z \end{cases}$$

$$\xrightarrow[\text{Proposition 8.6 of Wolsey, 1998}]{\text{Corollary of}} \quad \varphi \geq (1 - f(b_2 - b_1))\left((\eta_2 - \eta_1) - \lfloor b_2 - b_1 \rfloor\right)$$

$$\xrightarrow{\varphi = u + b_1 - \eta_1} \quad u \geq (1 - f(b_2 - b_1))\left((\eta_2 - \eta_1) - \lfloor b_2 - b_1 \rfloor\right) + \eta_1 - b_1$$

Having proved the validity of enhanced-MIR cuts in generic form, we create this kind of inequalities for our formulation;

Proof of inequalities 3.15:

Having constraints 2.4 ($\sum_{k \in K} y_{jkjk1} = 1$) and continuous values $\alpha, \beta > 0$, we can have expressions $\alpha \sum_{\hat{k} \in K} y_{\hat{j}k\hat{j}\hat{k}1} = \alpha$ and $\beta \sum_{\breve{k} \in K} y_{\breve{j}k\breve{j}\breve{k}1} = \beta$ for each $\hat{j}, \breve{j} \in J$. Subtracting these expression from the constraints 2.10 ($z_{jkt} \leq \sum_{\acute{k} \in K} y_{j\acute{k}kt}$), we have;

$$\begin{cases} z_{jkt} - \alpha \sum_{\hat{k} \in K} y_{\hat{j}k\hat{j}\hat{k}1} \leq \sum_{\acute{k} \in K} y_{j\acute{k}kt} - \alpha & ; \quad \forall\, j, \hat{j} \in J,\ k \in K, t \in T \\ z_{jkt} - \beta \sum_{\breve{k} \in K} y_{\breve{j}k\breve{j}\breve{k}1} \leq \sum_{\acute{k} \in K} y_{j\acute{k}kt} - \beta & ; \quad \forall\, j, \breve{j} \in J,\ k \in K, t \in T \end{cases}$$

$$z_{jkt} - \alpha \sum_{\hat{k} \in K} y_{\hat{j}k\hat{j}\hat{k}1} \leq \sum_{\acute{k} \in K} y_{j\acute{k}kt} - \alpha \quad \Longrightarrow \quad z_{jkt} - \lceil \alpha \rceil \sum_{\hat{k} \in K} y_{\hat{j}k\hat{j}\hat{k}1} \leq \sum_{\acute{k} \in K} y_{j\acute{k}kt} - \alpha$$

$$\Longrightarrow \quad \sum_{\acute{k} \in K} y_{j\acute{k}kt} - z_{jkt} + \lceil \alpha \rceil \sum_{\hat{k} \in K} y_{\hat{j}k\hat{j}\hat{k}1} \geq \alpha$$

Similarly, we have $\sum_{\acute{k}\in K} y_{j\acute{k}kt} - z_{jkt} + \lceil\beta\rceil \sum_{\check{k}\in K} y_{\check{j}k\check{j}\check{k}1} \geq \beta$ while $\sum_{\acute{k}\in K} y_{j\acute{k}kt} - z_{jkt} \geq 0$ because of constraints 2.10 ($z_{jkt} \leq \sum_{\acute{k}\in K} y_{j\acute{k}kt}$). Therefore, if $\alpha - \beta \notin Z$;

$$having \begin{cases} u = \sum_{\acute{k}\in K} y_{j\acute{k}kt} - z_{jkt} \geq 0 \\ \eta_1 = \lceil\alpha\rceil \sum_{\hat{k}\in K} y_{\hat{j}k\hat{j}\hat{k}1} \\ \eta_2 = \lceil\beta\rceil \sum_{\check{k}\in K} y_{\check{j}k\check{j}\check{k}1} \\ b_1 = \alpha \neq \beta = b_2 \end{cases} \implies \begin{cases} u + \eta_1 \geq b_1 \\ u + \eta_2 \geq b_2 \end{cases}$$

$$\overset{B.4}{\implies} \quad u \geq f(b_2 - b_1)\left(\lceil b_2 - b_1\rceil - (\eta_2 - \eta_1)\right) + b_1 - \eta_1 \quad \implies$$

$$\sum_{\acute{k}\in K} y_{j\acute{k}kt} - z_{jkt} \geq$$

$$f(\beta-\alpha)\left(\lceil\beta-\alpha\rceil - \lceil\beta\rceil \sum_{\check{k}\in K} y_{\check{j}k\check{j}\check{k}1} + \lceil\alpha\rceil \sum_{\hat{k}\in K} y_{\hat{j}k\hat{j}\hat{k}1}\right) + \alpha - \lceil\alpha\rceil \sum_{\hat{k}\in K} y_{\hat{j}k\hat{j}\hat{k}1}$$

$$\implies$$

$$\sum_{\acute{k}\in K} y_{j\acute{k}kt} - z_{jkt} + \lceil\alpha\rceil \sum_{\hat{k}\in K} y_{\hat{j}k\hat{j}\hat{k}1} \geq$$

$$\alpha + f(\beta-\alpha)\left(\lceil\beta-\alpha\rceil - \lceil\beta\rceil \sum_{\check{k}\in K} y_{\check{j}k\check{j}\check{k}1} + \lceil\alpha\rceil \sum_{\hat{k}\in K} y_{\hat{j}k\hat{j}\hat{k}1}\right)$$

$$\xrightarrow{1-f(\beta-\alpha)=f(\alpha-\beta)}$$

$$\sum_{\acute{k}\in K} y_{j\acute{k}kt} - z_{jkt} + f(\alpha-\beta)\lceil\alpha\rceil \sum_{\hat{k}\in K} y_{\hat{j}k\hat{j}\hat{k}1} \geq \alpha + f(\beta-\alpha)\left(\lceil\beta-\alpha\rceil - \lceil\beta\rceil \sum_{\check{k}\in K} y_{\check{j}k\check{j}\check{k}1}\right)$$

By replacing constraints 2.11 ($w_{jkt} \leq M\sum_{\acute{k}\in K} y_{j\acute{k}kt}$) by 2.10 ($z_{jkt} \leq \sum_{\acute{k}\in K} y_{j\acute{k}kt}$), validity of inequalities 3.18 could also be mathematically proved in a similar way.

Proof of inequalities 3.16:

Having constraints 2.4 ($\sum_{k\in K} y_{jkjk1} = 1$) and continuous values $\alpha, \beta > 0$, we can have expressions $\alpha \sum_{\hat{k}\in K} y_{\hat{j}k\hat{j}\hat{k}1} = \alpha$ and $\beta \sum_{\breve{k}\in K} y_{\breve{j}k\breve{j}\breve{k}1} = \beta$ for each $\hat{j}, \breve{j} \in J$. Aggregating these expression with the constraints 2.10 ($z_{jkt} \leq \sum_{\acute{k}\in K} y_{j\acute{k}kt}$), we have;

$$\begin{cases} z_{jkt} + \alpha \sum_{\hat{k}\in K} y_{\hat{j}k\hat{j}\hat{k}1} \leq \sum_{\acute{k}\in K} y_{j\acute{k}kt} + \alpha & ; \quad \forall\, j, \hat{j} \in J,\ k \in K, t \in T \\ z_{jkt} + \beta \sum_{\breve{k}\in K} y_{\breve{j}k\breve{j}\breve{k}1} \leq \sum_{\acute{k}\in K} y_{j\acute{k}kt} + \beta & ; \quad \forall\, j, \breve{j} \in J,\ k \in K, t \in T \end{cases}$$

$$\begin{aligned} z_{jkt} + \alpha \sum_{\hat{k}\in K} y_{\hat{j}k\hat{j}\hat{k}1} \leq \sum_{\acute{k}\in K} y_{j\acute{k}kt} + \alpha \quad &\Longrightarrow \quad z_{jkt} + \lfloor\alpha\rfloor \sum_{\hat{k}\in K} y_{\hat{j}k\hat{j}\hat{k}1} \leq \sum_{\acute{k}\in K} y_{j\acute{k}kt} + \alpha \\ &\Longrightarrow \quad \sum_{\acute{k}\in K} y_{j\acute{k}kt} - z_{jkt} + \alpha \geq \lfloor\alpha\rfloor \sum_{\hat{k}\in K} y_{\hat{j}k\hat{j}\hat{k}1} \end{aligned}$$

Similarly, we have $\sum_{\acute{k}\in K} y_{j\acute{k}kt} - z_{jkt} + \beta \geq \lfloor\beta\rfloor \sum_{\breve{k}\in K} y_{\breve{j}k\breve{j}\breve{k}1}$ while $\sum_{\acute{k}\in K} y_{j\acute{k}kt} - z_{jkt} \geq 0$ because of constraints 2.10 ($z_{jkt} \leq \sum_{\acute{k}\in K} y_{j\acute{k}kt}$). Therefore, if $\alpha - \beta \notin Z$;

$$having \begin{cases} u = \sum_{\acute{k}\in K} y_{j\acute{k}kt} - z_{jkt} \geq 0 \\ \eta_1 = \lfloor\alpha\rfloor \sum_{\hat{k}\in K} y_{\hat{j}k\hat{j}\hat{k}1} \\ \eta_2 = \lfloor\beta\rfloor \sum_{\breve{k}\in K} y_{\breve{j}k\breve{j}\breve{k}1} \\ b_1 = \alpha \neq \beta = b_2 \end{cases} \Longrightarrow \begin{cases} u + b_1 \geq \eta_1 \\ u + b_2 \geq \eta_2 \end{cases}$$

$$\overset{B.5}{\Longrightarrow} \quad u \geq (1 - f(b_2 - b_1))\left((\eta_2 - \eta_1) - \lfloor b_2 - b_1 \rfloor\right) + \eta_1 - b_1 \quad \Longrightarrow$$

$$\sum_{\acute{k}\in K} y_{j\acute{k}kt} - z_{jkt} \geq$$

$$(1 - f(\beta - \alpha)) \left(\lfloor\beta\rfloor \sum_{\breve{k}\in K} y_{\breve{j}k\breve{j}\breve{k}1} - \lfloor\alpha\rfloor \sum_{\hat{k}\in K} y_{\hat{j}k\hat{j}\hat{k}1} - \lfloor\beta - \alpha\rfloor \right) + \lfloor\alpha\rfloor \sum_{\hat{k}\in K} y_{\hat{j}k\hat{j}\hat{k}1} - \alpha$$

$$\xrightarrow{1-f(\beta-\alpha)=f(\alpha-\beta)}$$

$$\sum_{\acute{k}\in K} y_{j\acute{k}kt} - z_{jkt} \geq$$

$$f(\alpha - \beta) \left(\lfloor\beta\rfloor \sum_{\breve{k}\in K} y_{\breve{j}k\breve{j}\breve{k}1} - \lfloor\alpha\rfloor \sum_{\hat{k}\in K} y_{\hat{j}k\hat{j}\hat{k}1} - \lfloor\beta - \alpha\rfloor \right) + \lfloor\alpha\rfloor \sum_{\hat{k}\in K} y_{\hat{j}k\hat{j}\hat{k}1} - \alpha$$

$$\xrightarrow{1-f(\beta-\alpha)=f(\alpha-\beta)}$$

$$\sum_{\acute{k}\in K} y_{j\acute{k}kt} - z_{jkt} + \alpha + f(\alpha - \beta)\lfloor\beta - \alpha\rfloor \geq f(\beta - \alpha)\lfloor\alpha\rfloor \sum_{\hat{k}\in K} y_{\hat{j}k\hat{j}\hat{k}1} + f(\alpha - \beta)\lfloor\beta\rfloor \sum_{\breve{k}\in K} y_{\breve{j}k\breve{j}\breve{k}1}$$

By replacing constraints 2.11 ($w_{jkt} \leq M \sum_{\acute{k} \in K} y_{j\acute{k}kt}$) by 2.10 ($z_{jkt} \leq \sum_{\acute{k} \in K} y_{j\acute{k}kt}$), validity of inequalities 3.17 could also be mathematically proved in a similar way.

Bibliography

Aboolian, Robert, Oded Berman, and Zvi Drezner (2008). "Location and allocation of service units on a congested network". In: *IIE Transactions* 40.4, pp. 422–433.

Aboolian, Robert, Oded Berman, and Dmitry Krass (2007). "Competitive facility location model with concave demand". In: *European Journal of Operational Research* 181.2, pp. 598–619.

Aboolian, Robert, Yi Sun, and Gary J Koehler (2009). "A location–allocation problem for a web services provider in a competitive market". In: *European Journal of Operational Research* 194.1, pp. 64–77.

Abouee-Mehrizi, Hossein et al. (2011). "Optimizing capacity, pricing and location decisions on a congested network with balking". In: *Mathematical Methods of Operations Research* 74.2, pp. 233–255.

Ahmadi-Javid, Amir and Pooya Hoseinpour (2017). "Convexification of Queueing Formulas by Mixed-Integer Second-Order Cone Programming: An Application to a Discrete Location Problem with Congestion". In: *arXiv preprint arXiv:1710.05794*.

Alijani, Reza et al. (2017). "Two-sided Facility Location". In: *arXiv:1711.11392*.

Amiri, Ali (2006). "Designing a distribution network in a supply chain system: Formulation and efficient solution procedure". In: *European Journal of Operational Research* 171.2, pp. 567–576.

Arabani, Alireza Boloori and Reza Zanjirani Farahani (2012). "Facility location dynamics: An overview of classifications and applications". In: *Computers & Industrial Engineering* 62.1, pp. 408–420.

Arkat, J and R Jafari (2016). "Network location problem with stochastic and uniformly distributed demands". In: *International Journal of Engineering-Transactions B: Applications* 29.5, p. 654.

Bai, Yun et al. (2011). "Biofuel refinery location and supply chain planning under traffic congestion". In: *Transportation Research Part B: Methodological* 45.1, pp. 162–175.

Ball, Michael O and Feng L Lin (1993). "A reliability model applied to emergency service vehicle location". In: *Operations research* 41.1, pp. 18–36.

Ballou, Ronald H (1968). "Dynamic warehouse location analysis". In: *Journal of Marketing Research*, pp. 271–276.

Baron, Opher, Oded Berman, and Dmitry Krass (2008). "Facility location with stochastic demand and constraints on waiting time". In: *Manufacturing & Service Operations Management* 10.3, pp. 484–505.

Baron, Opher et al. (2007). "The equitable location problem on the plane". In: *European Journal of Operational Research* 183.2, pp. 578–590.

Beamon, Benita M (1998). "Supply chain design and analysis:: Models and methods". In: *International journal of production economics* 55.3, pp. 281–294.

Berman, Oded and Zvi Drezner (2006). "Location of congested capacitated facilities with distance-sensitive demand". In: *IIE Transactions* 38.3, pp. 213–221.

Berman, Oded and Dmitry Krass (2001). "11 Facility Location Problems with Stochastic Demands and Congestion". In: *Facility location: applications and theory*, p. 329.

Berman, Oded, Dmitry Krass, and Mozart BC Menezes (2007). "Reliability issues, strategic co-location and centralization in m-median problems". In: *Operations Research* 55.2, pp. 332–350.

Berman, Oded, Dmitry Krass, and Jiamin Wang (2006). "Locating service facilities to reduce lost demand". In: *IIE Transactions* 38.11, pp. 933–946.

Berman, Oded, Richard C Larson, and Samuel S Chiu (1985). "Optimal server location on a network operating as an M/G/1 queue". In: *Operations research* 33.4, pp. 746–771.

Berman, Oded et al. (2009). "Optimal location with equitable loads". In: *Annals of Operations Research* 167.1, pp. 307–325.

Bodur, Merve and James R. Luedtke (2016). "Mixed-Integer Rounding Enhanced Benders Decomposition for Multiclass Service-System Staffing and Scheduling with Arrival Rate Uncertainty". In: *Management Science*.

Boffey, Brian, Roberto D Galvão, and V Marianov (2010). "Location of single-server immobile facilities subject to a loss constraint". In: *Journal of the Operational Research Society* 61.6, pp. 987–999.

Bundschuh, Markus, Diego Klabjan, and Deborah L Thurston (2003). "Modeling robust and reliable supply chains". In: *Optimization Online e-print*.

Campbell, James F (1990). "Locating transportation terminals to serve an expanding demand". In: *Transportation Research Part B: Methodological* 24.3, pp. 173–192.

Castillo, Ignacio, Armann Ingolfsson, and Thaddeus Sim (2009). "Social optimal location of facilities with fixed servers, stochastic demand, and congestion". In: *Production and Operations Management* 18.6, pp. 721–736.

Chopra, S and P Meindl (2010). *Supply Chain Management.–NY*.

Contreras, Ivan, Moayad Tanash, and Navneet Vidyarthi (2016). "Exact and heuristic approaches for the cycle hub location problem". In: *Annals of Operations Research*, pp. 1–23.

Cordeau, Jean-François, Federico Pasin, and Marius M Solomon (2006). "An integrated model for logistics network design". In: *Annals of operations research* 144.1, pp. 59–82.

Crainic, Teodor Gabriel and Gilbert Laporte (1997). "Planning models for freight transportation". In: *European journal of operational research* 97.3, pp. 409–438.

Current, John, Samuel Ratick, and Charles ReVelle (1998). "Dynamic facility location when the total number of facilities is uncertain: A decision analysis approach". In: *European Journal of Operational Research* 110.3, pp. 597–609.

Daskin, M (1995). *Network and discrete location: models, algorithms and applications, New York, USA: John Wiley& Sons*.

Daskin, Mark S (1982). "Application of an expected covering model to emergency medical service system design". In: *Decision Sciences* 13.3, pp. 416–439.

— (2011). *Network and discrete location: models, algorithms, and applications*. John Wiley & Sons.

Daskin, Mark S, Collette R Coullard, and Zuo-Jun Max Shen (2002). "An inventory-location model: Formulation, solution algorithm and computational results". In: *Annals of operations research* 110.1-4, pp. 83–106.

Dias, Joana, M Eugénia Captivo, and João Clímaco (2006). "Capacitated dynamic location problems with opening, closure and reopening of facilities". In: *IMA Journal of Management Mathematics* 17.4, pp. 317–348.

Dogan, Koray and Marc Goetschalckx (1999). "A primal decomposition method for the integrated design of multi-period production-distribution systems". In: *Iie Transactions* 31.11, pp. 1027–1036.

Drezner, E (1996). "Facility location: A survey of applications and methods". In: *Journal of the Operational Research Society* 47.11, pp. 1421–1421.

Drezner, Tammy and Zvi Drezner (2011). "The gravity multiple server location problem". In: *Computers & Operations Research* 38.3, pp. 694–701.

Drezner, Zvi (1995). "Dynamic facility location: The progressive p-median problem". In: *Location Science* 3.1, pp. 1–7.

Drezner, Zvi and Horst W Hamacher (2001). *Facility location: applications and theory*. Springer Science & Business Media.

Drezner, Zvi and GO Wesolowsky (1991). "Facility location when demand is time dependent". In: *Naval Research Logistics (NRL)* 38.5, pp. 763–777.

Eiselt, Horst A, Michel Gendreau, and Gilbert Laporte (1992). "Location of facilities on a network subject to a single-edge failure". In: *Networks* 22.3, pp. 231–246.

Elhedhli, Samir (2005). "Exact solution of a class of nonlinear knapsack problems". In: *Operations Research Letters* 33.6, pp. 615–624.

— (2006). "Service system design with immobile servers, stochastic demand, and congestion". In: *Manufacturing & Service Operations Management* 8.1, pp. 92–97.

Elhedhli, Samir and Fatma Gzara (2008). "Integrated design of supply chain networks with three echelons, multiple commodities and technology selection". In: *IIE Transactions* 40.1, pp. 31–44.

Elhedhli, Samir and Frank Xiaolong Hu (2005). "Hub-and-spoke network design with congestion". In: *Computers & Operations Research* 32.6, pp. 1615–1632.

Erlebacher, Steven J and Russell D Meller (2000). "The interaction of location and inventory in designing distribution systems". In: *Iie Transactions* 32.2, pp. 155–166.

Eskigun, Erdem (2002). "Outbound supply chain network design for a large-scale automotive company". In: *Purdue University, Theses/Dissertations*.

Eskigun, Erdem et al. (2005). "Outbound supply chain network design with mode selection, lead times and capacitated vehicle distribution centers". In: *European Journal of Operational Research* 165.1, pp. 182–206.

Farahani, Reza Zanjirani, Maryam Abedian, and Sara Sharahi (2009). "Dynamic facility location problem". In: *Facility Location*. Springer, pp. 347–372.

Farahani, Reza Zanjirani and Masoud Hekmatfar (2009). *Facility location: concepts, models, algorithms and case studies*. Springer.

Fischetti, Matteo, Ivana Ljubić, and Markus Sinnl (2016). "Benders decomposition without separability: a computational study for capacitated facility location problems". In: *European Journal of Operational Research* 253.3, pp. 557–569.

Francis, Richard L, Leon Franklin McGinnis, and John A White (1992). *Facility layout and location: an analytical approach*. Pearson College Division.

Gendron, Bernard and Teodor Gabriel Crainic (1994). *Relaxations for multicommodity capacitated network design problems*. Tech. rep. Univ., CRT.

Gross, Donald and Carl M Harris (1998). *Fundamentals of Queuing Theory, 3rd Rev*. John Wiley & Sons.

Hajipour, Vahid, Reza Zanjirani Farahani, and Parviz Fattahi (2016). "Bi-objective vibration damping optimization for congested location–pricing problem". In: *Computers & Operations Research* 70, pp. 87–100.

Hajipour, Vahid, Vahid Khodakarami, and Madjid Tavana (2014). "The redundancy queuing-location-allocation problem: A novel approach". In: *IEEE Transactions on Engineering Management* 61.3, pp. 534–544.

Hajipour, Vahid and Seyed Hamid Reza Pasandideh (2011). "A new multi-objective facility location model within batch arrival queuing framework". In: *World Academy of Science, Engineering and Technology* 78, pp. 1665–1673.

— (2012). "Proposing an adaptive particle swarm optimization for a novel bi-objective queuing facility location model". In: *Economic Computation and Economic Cybernetics Studies and Research* 46.3, pp. 223–240.

Hajipour, Vahid et al. (2016). "Multi-objective multi-layer congested facility location-allocation problem optimization with Pareto-based meta-heuristics". In: *Applied Mathematical Modelling* 40.7, pp. 4948–4969.

Hakimi, S Louis (1964). "Optimum locations of switching centers and the absolute centers and medians of a graph". In: *Operations research* 12.3, pp. 450–459.

Hamacher, Horst W and Zvi Drezner (2002). "Facility location: applications and theory". In: *Science & Business Media: Springer*.

Hamaguchi, Tomoki and Koichi Nakade (2010). "Optimal location of facilities on a network in which each facility is operating as an M/G/1 queue". In: *Journal of Service Science and Management* 3.03, p. 287.

Hijazi, Hassan, Pierre Bonami, and Adam Ouorou (2013). "An outer-inner approximation for separable mixed-integer nonlinear programs". In: *INFORMS Journal on Computing* 26.1, pp. 31–44.

Huang, Simin, Rajan Batta, and Rakesh Nagi (2005). "Distribution network design: Selection and sizing of congested connections". In: *Naval Research Logistics (NRL)* 52.8, pp. 701–712.

Jamil, Mamnoon, Alok Baveja, and Rajan Batta (1999). "The stochastic queue center problem". In: *Computers & Operations Research* 26.14, pp. 1423–1436.

Jayaswal, Sachin and Navneet Vidyarthi (2017). "Facility location under service level constraints for heterogeneous customers". In: *Annals of Operations Research* 253.1, pp. 275–305.

Jena, Sanjay D., Jean-François Cordeau, and Bernard Gendron (2015). "Dynamic facility location with generalized modular capacities". In: *Transportation Science* 49.3, pp. 484–499.

Jouzdani, Javid, Seyed Jafar Sadjadi, and Mohammad Fathian (2013). "Dynamic dairy facility location and supply chain planning under traffic congestion and demand uncertainty: A case study of Tehran". In: *Applied Mathematical Modelling* 37.18, pp. 8467–8483.

Jula, Payman and Robert C Leachman (2011). "Long-and short-run supply-chain optimization models for the allocation and congestion management of containerized imports from Asia to the United States". In: *Transportation Research Part E: Logistics and Transportation Review* 47.5, pp. 593–608.

Kim, S (2013). "A column generation heuristic for congested facility location problem with clearing functions". In: *Journal of the Operational Research Society* 64.12, pp. 1780–1789.

Klose, Andreas and Andreas Drexl (2005). "Facility location models for distribution system design". In: *European journal of operational research* 162.1, pp. 4–29.

Laporte, Gilbert, Stefan Nickel, and Francisco S. da Gama (2015). *Location science*. Springer.

Larson, Richard C (1974). "A hypercube queuing model for facility location and redistricting in urban emergency services". In: *Computers & Operations Research* 1.1, pp. 67–95.

Leachman, Robert C and Payman Jula (2011). "Congestion analysis of waterborne, containerized imports from Asia to the United States". In: *Transportation Research Part E: Logistics and Transportation Review* 47.6, pp. 992–1004.

Manne, Alan S (1961). "Capacity expansion and probabilistic growth". In: *Econometrica: Journal of the Econometric Society*, pp. 632–649.

Marchand, Hugues et al. (2002). "Cutting planes in integer and mixed integer programming". In: *Discrete Applied Mathematics* 123.1, pp. 397–446.

Marianov, V, T Brian Boffey, and Roberto D Galvão (2009). "Optimal location of multi-server congestible facilities operating as M/Er/m/N queues". In: *Journal of the Operational Research Society* 60.5, pp. 674–684.

Marianov, Vladimir and Charles ReVelle (1996). "The queueing maximal availability location problem: a model for the siting of emergency vehicles". In: *European Journal of Operational Research* 93.1, pp. 110–120.

Marianov, Vladimir and Miguel Ríos (2000). "A probabilistic quality of service constraint for a location model of switches in ATM communications networks". In: *Annals of Operations Research* 96.1, pp. 237–243.

Marianov, Vladimir and Daniel Serra (1998). "Probabilistic, maximal covering location—allocation models forcongested systems". In: *Journal of Regional Science* 38.3, pp. 401–424.

— (2011). "Location of multiple-server common service centers or facilities, for minimizing general congestion and travel cost functions". In: *International Regional Science Review* 34.3, pp. 323–338.

Melo, M Teresa, Stefan Nickel, and Francisco Saldanha-Da-Gama (2009). "Facility location and supply chain management–A review". In: *European journal of operational research* 196.2, pp. 401–412.

Miller, Tan C et al. (2007). "Reaction function based dynamic location modeling in Stackelberg–Nash–Cournot competition". In: *Networks and Spatial Economics* 7.1, pp. 77–97.

Min, Hokey and Gengui Zhou (2002). "Supply chain modeling: past, present and future". In: *Computers & industrial engineering* 43.1, pp. 231–249.

Miranda, Pablo A and Rodrigo A Garrido (2006). "A simultaneous inventory control and facility location model with stochastic capacity constraints". In: *Networks and Spatial Economics* 6.1, pp. 39–53.

Mirchandani, Pitu B and Richard L Francis (1990). *Discrete location theory*.

Mirchandani, Pitu B and Amedeo R Odoni (1979). "Locations of medians on stochastic networks". In: *Transportation Science* 13.2, pp. 85–97.

Moghadas, Moeen and Taghizadeh Kakhki (2011). "Maximal covering location-allocation problem with M/M/k queuing system and side constraints". In: *Iranian Journal of Operations Research* 2.2, pp. 1–16.

Nel, Louis D and Charles J Colbourn (1990). "Locating a broadcast facility in an unreliable network". In: *INFOR: Information Systems and Operational Research* 28.4, pp. 363–379.

Nozick, Linda K and Mark A Turnquist (2001). "Inventory, transportation, service quality and the location of distribution centers". In: *European Journal of Operational Research* 129.2, pp. 362–371.

Ortega, Francisco and Laurence A Wolsey (2003). "A branch-and-cut algorithm for the single-commodity, uncapacitated, fixed-charge network flow problem". In: *Networks* 41.3, pp. 143–158.

Owen, Susan Hesse and Mark S Daskin (1998). "Strategic facility location: A review". In: *European journal of operational research* 111.3, pp. 423–447.

Pasandideh, Seyed Hamid Reza and Amirhossain Chambari (2010). "A new model for location-allocation problem within queuing framework". In: *Journal of Optimization in Industrial Engineering*, pp. 53–61.

Rabieyan, Reza and Mehdi Seifbarghy (2010). "Maximal Benefit location problem for a congested system". In: *Journal of Industrial Engineering* 5, pp. 73–83.

ReVelle, Charles and Kathleen Hogan (1989). "The maximum availability location problem". In: *Transportation Science* 23.3, pp. 192–200.

Roodman, Gary M and Leroy B Schwarz (1975). "Optimal and heuristic facility phase-out strategies". In: *AIIE transactions* 7.2, pp. 177–184.

Schilling, David A (1980). "Dynamic location modeling for public-sector facilities: a multi-criteria approach". In: *Decision Sciences* 11.4, pp. 714–724.

Seifbarghy, Mehdi and Aida Mansouri (2016). "Modelling and solving a congested facility location problem considering systems' and customers' objectives". In: *International Journal of Industrial and Systems Engineering* 22.3, pp. 281–304.

Shen, Zuo-Jun Max (2005). "A multi-commodity supply chain design problem". In: *Iie Transactions* 37.8, pp. 753–762.

Shen, Zuo-Jun Max and Lian Qi (2007). "Incorporating inventory and routing costs in strategic location models". In: *European journal of operational research* 179.2, pp. 372–389.

Silva, Francisco and Daniel Serra (2016). "Locating Emergency Services with Different Priorities: The Priority Queuing Covering Location Problem". In: *Operational Research for Emergency Planning in Healthcare: Volume 1*. Springer, pp. 15–35.

Snyder, Lawrence V (2006). "Facility location under uncertainty: a review". In: *IIE Transactions* 38.7, pp. 547–564.

Snyder, Lawrence V and Mark S Daskin (2005). "Reliability models for facility location: the expected failure cost case". In: *Transportation Science* 39.3, pp. 400–416.

Sourirajan, Karthik, Leyla Ozsen, and Reha Uzsoy (2007). "A single-product network design model with lead time and safety stock considerations". In: *IIE Transactions* 39.5, pp. 411–424.

Suzuki, Atsuo and Zvi Drezner (2009). "The minimum equitable radius location problem with continuous demand". In: *European Journal of Operational Research* 195.1, pp. 17–30.

Tavakkoli-Moghaddam, Reza et al. (2017). "Pricing and location decisions in multi-objective facility location problem with M/M/m/k queuing systems". In: *Engineering Optimization* 49.1, pp. 136–160.

Teo, Chung-Piaw and Jia Shu (2004). "Warehouse-retailer network design problem". In: *Operations Research* 52.3, pp. 396–408.

Vidal, Carlos J and Marc Goetschalckx (2000). "Modeling the effect of uncertainties on global logistics systems". In: *Journal of Business Logistics* 21.1, p. 95.

Vidyarthi, N. and S. Jayaswal (2014). "Efficient solution of a class of location-allocation problems with stochastic demand and congestion". In: *Computers and Operations Research* 48. Cited By :4, pp. 20–30.

Vidyarthi, Navneet, Samir Elhedhli, and Elizabeth Jewkes (2009). "Response time reduction in make-to-order and assemble-to-order supply chain design". In: *IIE transactions* 41.5, pp. 448–466.

Vidyarthi, Navneet and Onur Kuzgunkaya (2015). "The impact of directed choice on the design of preventive healthcare facility network under congestion". In: *Health care management science* 18.4, pp. 459–474.

Wang, Qian, Rajan Batta, and Christopher M Rump (2002). "Algorithms for a facility location problem with stochastic customer demand and immobile servers". In: *Annals of operations Research* 111.1, pp. 17–34.

— (2004). "Facility location models for immobile servers with stochastic demand". In: *Naval Research Logistics (NRL)* 51.1, pp. 137–152.

Weaver, Jerry R and Richard L Church (1983). "Computational procedures for location problems on stochastic networks". In: *Transportation Science* 17.2, pp. 168–180.

Weber, Alfred (1909). *Ueber den standort der industrien*. Vol. 2.

Wesolowsky, GO and GW Truscott (1976). "Dynamic location of multiple facilities for a class of continuous problems". In: *AIIE Trans* 8.1, pp. 76–83.

Wolsey, Laurence A. (1998). *Integer programming*. Wiley.

Zaferanieh, Mehdi and Jafar Fathali (2017). "The stochastic queue core problem on a tree". In: *arXiv preprint arXiv:1701.01812*.

Zarrinpoor, Naeme, Mohammad Saber Fallahnezhad, and Mir Saman Pishvaee (2018). "The design of a reliable and robust hierarchical health service network using an accelerated Benders decomposition algorithm". In: *European Journal of Operational Research* 265.3, pp. 1013–1032.

Zhang, Yue, Oded Berman, and Vedat Verter (2009). "Incorporating congestion in preventive healthcare facility network design". In: *European Journal of Operational Research* 198.3, pp. 922–935.

— (2012). "The impact of client choice on preventive healthcare facility network design". In: *OR spectrum* 34.2, pp. 349–370.

Zhang, Yue et al. (2010). "A bilevel model for preventive healthcare facility network design with congestion". In: *IIE Transactions* 42.12, pp. 865–880.