A MEMETIC ALGORITHM FOR LOGISTICS NETWORK DESIGN PROBLEMS

by

Cristiano Baptista Faria Maciel

Submitted to the Department of Mathematics and Statistics
for the degree of

Master of Economic and Corporate Decision-Making

at the

LISBON SCHOOL OF ECONOMICS & MANAGEMENT

October 2014

Certified by . . . .
Prof. Dr. Teresa Melo
Thesis Supervisor

A MEMETIC ALGORITHM FOR LOGISTICS NETWORK DESIGN PROBLEMS

by

Cristiano Baptista Faria Maciel

Submitted to the Department of Mathematics and Statistics
on October 15, 2014, in partial fulfillment of the
requirements for the degree of
Master of Economic and Corporate Decision-Making

Abstract

This thesis describes a memetic algorithm applied to the design of a three-echelon logistics network over multiple periods with transportation mode selection and outsourcing. The memetic algorithm can be applied to an existing supply chain in order to obtain an optimized configuration or, if required, it can be used to define a new logistics network. In addition, production can be outsourced and direct shipments of products to customer zones are possible. In this problem, the capacity of an existing or new facility can be expanded over the time horizon. In this case, the facility cannot be closed. Existing facilities, once closed, cannot be reopened. New facilities cannot be closed, once opened. The heuristic is able to determine the number and locations of facilities (i.e. plants and warehouses), capacity levels as well as the flow of products throughout the supply chain.

KEYWORDS: Memetic Algorithm, Genetic Algorithm, Outsourcing.

Acknowledgments

To Prof. Teresa Melo, that was always available to help and guide me during this process. Without her support, comments and suggestions this work would not be possible to finish.

To the Saarland University of Applied Sciences that provided unique conditions to work and live during this period at Saarbrücken, Germany.

To my family and friends that despite the distance, constantly give me support and motivation to accomplish my purposes.

To everyone, a grateful thanks.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Logistics network design (LND) is a strategic planning process that aims to provide an exceptional performance in order to achieve shareholder goals and clients satisfaction. As result, LND is dedicated to process optimization by taking the following decisions:

- Defining the optimal number and location of facilities (e.g., manufacturing plants and warehouses)
- Allocating capacity and technology requirements to facilities
- Deciding on the flow of products throughout the supply chain

In order to be competitive a company may consider redesigning its supply chain or designing a new chain, depending on the company business requirements and strategy objectives. If a company decides to enter a new market or to expand its product segments, then designing a new logistics network (LND) is the most advantageous decision. In contrast, if there is change in the market and business conditions, in order to deal with increasing costs or new business requirements a company is impelled to adopt a restructuration of its supply chain through logistics network re-design (LRD). Strategic network decisions have a direct influence on the competitiveness of a company and it is crucial for management to have a framework for successful tactical and operational supply chain processes. As highlighted by Ballou [4] and Harrison [14], a network re-design project can result in a 5 to 15 percent reduction of the overall logistics costs, with 10 percent being often achieved. In this thesis, an

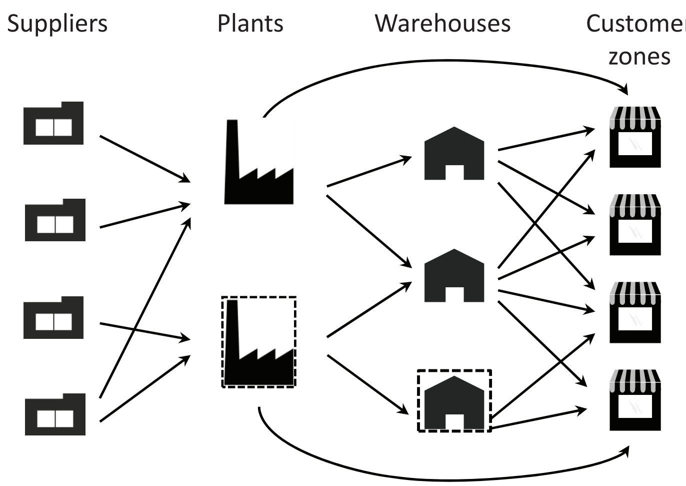

Figure 1-1: Possible configuration of a multi-echelon logistics network.

integrated and comprehensive view of the supply chain is taken by considering suppliers, manufacturing facilities, warehouses, transportation channels, and customer zones as shown in Figure 1-1. In an LRD approach, a network is already in place with a number of plants and warehouses being operated at fixed locations (these are the facilities marked by the dashed lines in the figure). A variety of decisions have to be made regarding facility location and logistics functions along the supply chain, such as opening new plants and/or warehouses at potential sites (the facilities without dashed lines in Figure 1-1) and selecting their capacity levels from a set of available discrete sizes. This is motivated by the fact that capacity is often purchased in the form of equipment which is only available at a few discrete sizes. As strategic planning for multiple time periods is considered, capacity can be acquired more than once over the time horizon in the same location. Existing plants and warehouses can be closed throughout the planning horizon. Alternatively, their capacities can be expanded if a company forecasts future expansion into new markets or to meet increasing product demand in current markets. In an LND approach, the scope of the location decisions is limited to deciding on the optimal size, number, and location of new facilities. Logistics decisions, the second group of key business decisions, involve supplier selection in conjunction with procurement as well as production and distribution decisions. Furthermore, a strategic choice between in-house manufacturing, outsourcing or a mixed approach is to be taken. In the network depicted in Figure 1-1, multiple types of products are manufactured at plants by processing sub-assemblies and components, hereafter called raw materials. The latter can be procured from various suppliers taking into account their availability and cost. Finished products can be delivered to warehouses or shipped directly to customer zones. The flow of goods throughout the network and the use of transportation modes are to be determined in each time period. In addition, end products can also be purchased from external sources and consolidated at the warehouses. The objective is to determine the optimal network configuration over a planning horizon so as to minimize the total cost. The remainder of this thesis is organized as follows. In chapter 2, we review the relevant literature dedicated to LND/LRD and describe its relation to our new model. Chapter 3 introduces a mixed-integer linear programming formulation for logistics network design and re-design. In chapter 4, the memetic algorithm is described. In chapter 5, the heuristic performance is analyzed for a set of randomly generated instances. Moreover, possible ways of improving the heuristic performance are presented. In chapter 6, the main conclusions are exposed and recommendations for future research are provided.

Chapter 2

Literature review

Beginning with the work of Geoffrion and Graves [12] on multi-commodity distribution network design in 1974, a large number of optimization-based approaches have been proposed for the design of logistics networks as shown by the recent surveys of Melo et al. [17] and Mula et al. [22]. These works have resulted in significant improvements in the modeling of these problems as well as in algorithmic and computational efficiency. One of the reasons that contributes to such a large number of literature references is the variety of characteristics that can be taken into account in LND problems: type of planning horizon (single or multi-period), facility location and sizing, number of echelons and distribution levels, multi-stage production taking the bill of materials (BOM) into account, and transportation mode selection, among others.

Although the timing of facility locations and expansions over an extended time horizon is of major importance to decision-makers in strategic network design, the majority of the literature addresses single-period problems, e.g., Babazadeh et al. [2], Cordeau et al. [7], Elhedhli and Goffin [10], Eskigun et al. [11], Olivares et al. [23], and Sadjady and Davoudpour [24]. Our research is different in that a multi-period planning horizon is considered. Unlike our work, in some multi-period LND problems facility sizing is static, meaning that facilities cannot have their capacities expanded or contracted over the planning horizon. The model proposed by Gourdin and Klopfenstein [13] falls into this category.

We will focus next on multi-period LND and LRD problems with dynamic facility sizing decisions. In particular, we will discuss the extent to which the features of the model to be detailed in Section 3 differ from those reported so far in the literature.

To re-design a two-layer network, Antunes and Peeters [1] suggest a modeling framework that allows opening new facilities and closing existing locations, as well as reducing and expanding capacity. Budget constraints are taken into account over the time horizon. To find the minimum discounted cost solution, the authors propose a simulated annealing approach.

Melo et al. [16] study the re-design of a multi-echelon network considering facility expansion and contraction. This feature is modeled through relocating capacity from existing facilities to new facilities over the planning horizon. Network re-design decisions (opening, closing, and relocating facilities) are subject to budget constraints in each time period. General purpose optimization software is used to solve small and medium-sized problem instances. Melo et al. [18, 19] also developed heuristic procedures for this special form of network re-design. The numerical experiments indicate that good solutions can be obtained for large-sized instances within acceptable computational time.

Hinojosa et al. [15] address a multi-echelon, multi-commodity network re-design problem with inventory strategic decisions at warehouses and outsourcing of demand. Commodities flow from manufacturing plants to customers via warehouses. Outsourced commodities are delivered directly to customers. An initial network configuration is considered that gradually changes over a multi-period horizon through opening new facilities and closing existing facilities. A lower bound is imposed on the number of plants and warehouses operating in the first and last time periods. The problem is solved using a Lagrangean-based procedure through decomposition into simpler subproblems. A heuristic method is then applied to obtain a feasible solution.

Thanh et al. [25] consider a multi-period, multi-product logistics network comprising suppliers, plants, warehouses (public and private), and customers. Strategic decisions include facility location and capacity acquisition as well as supplier selection, production, distribution, and inventory planning. In particular, plants and warehouses can have their capacities expanded (but not contracted) over the time horizon. Production decisions take into account the BOM and intermediate components can be sub-contracted to an external plant. Furthermore, products flow downstream not only between adjacent supply chain layers but also directly from plants to customers. In addition, plants may exchange components. To identify the least-cost network configuration, Thanh et al. [27] propose an LP-rounding heuristic. This method is later improved by combining it with DC programming (cf. Thanh et al. [26]).

Bashiri and Badri [5] address the problem of designing a new supply chain network with a similar topology to that considered in [25]. The objective is to find the network configuration that maximizes the total net profit subject to a given budget in each time period. In this work, demand requirements may not be completely satisfied. Strategic decisions include opening and expanding new plants and private (company-owned) warehouses. In addition, public warehouses can be hired for a pre-specified number of time periods and variable costs are charged for their operation. The proposed model is extended in Bashiri et al. [6] through introducing different time scales for strategic and tactical decisions. In particular, the latter are made in each time period, whereas network design decisions are only made over a subset of periods of the planning horizon. The models presented in [5, 6] are solved with a general-purpose optimization solver. Later, Badri et al. [3] develop a Lagrangean-based approach. Feasible solutions are obtained by dualising the budget constraints for opening new facilities or expanding the capacities of existing facilities and adding some constraints to the subproblems to guarantee feasibility.

Dias et al. [9] proposed a memetic algorithm for an LND problem involving opening, closing, and re-opening facilities over a planning horizon with multiple periods. Various mechanisms were employed to search the solution space. In addition,the decision taken is able to set the search area and establish limits to the objective function.

In order to relate the existing literature to the LND and LRD problems that are studied in this paper, a classification of the aforementioned works is given in Table 2.1. This table is not intended to provide an exhaustive list of the features described in this section but rather to illustrate the extent to which our research generalizes the existing literature.

The characteristics of the surveyed LND and LRD problems are classified according to five categories. The category "Network" comprises the number of layers in the supply chain, including the customer level (column 2) and the number of layers involving location decisions (column 3). Furthermore, it is specified whether products can be distributed directly from higher level facilities to customer locations (column 4). The second category (column 5) refers to the type of planning horizon. The category "Facility sizing" summarizes the type of capacity decisions that can be made. To this end, column 6 indicates if the size of each facility is limited (G: global capacities) and if capacity levels are selected from a set of available discrete sizes (M: modular capacities). In column 7, the type of capacity planning is given: expansion (E) and/or downsizing (D) over the time horizon. Moreover, to distinguish network design from network re-design problems, the latter are highlighted with the letter C, meaning that existing facilities can be closed. The category in column 8 refers to the selection of transportation modes. Finally, the category "Products" outlines characteristics related to the number of products (column 9), multi-stage production taking into account the BOM (column 10), satisfaction of demand requirements (column 11), and the possibility of outsourcing components or end products as an alternative to in-house manufacturing (column 11).

The last row of Table 2.1 highlights the main features of the model to be detailed in Section 3. It can be seen that in our model various features are considered simultaneously in a multi-period planning context. To the best of our knowledge, this integration of practically relevant features into a single model has not been addressed in the related literature so far.

References	Network			Multiple time periods	Facility sizing		Transportation	Products			
	Layers	Location levels	Direct shipments		Capacity type	Capacity changes	Mode selection	Multiple	BOM	Demand satisfaction	Outsourcing
Antunes and Peeters [1]	2	1		✓	M,G	E,D,C				✓	
Babazadeh et al. [2]	3	2	✓		M		✓				✓
Badri et al. [3]	4	2		✓	M,G	E		✓	✓		
Bashiri and Badri [5]	4	2		✓	M,G	E		✓	✓		
Bashiri et al. [6]	4	2		✓	M,G	E		✓	✓		
Cordeau et al. [7]	4	2			G		✓	✓	✓	✓	
Elhedhli and Goffin [10]	3	1			G			✓		✓	
Eskigun et al. [11]	3	1	✓		G		✓			✓	
Gourdin and Klopfenstein [13]	2	1		✓	M,G	E				✓	
Hinojosa et al. [15]	3	2		✓	G	C		✓		✓	✓
Melo et al. [16, 18, 19]	N	N	✓	✓	G	E,C		✓	✓	✓	
Olivares et al. [23]	3	1			G		✓			✓	
Sadjady and Davoudpour [24]	3	2			M		✓	✓		✓	
Thanh et al. [25, 27, 26]	4	2		✓	M,G	E,C		✓	✓	✓	✓
New model (cf. Section 3)	**4**	**2**	✓	✓	**M,G**	**E,C**	✓	✓	✓	✓	✓

C: Closing existing facilities
D: Downsizing (capacity contraction)
E: Capacity expansion
G: Global capacities
M: Modular capacities
N: No limit on the number of network layers and location levels

Table 2.1: Classification of the existing literature

Chapter 3

Mathematical formulation

In this section, we introduce a mathematical model for a comprehensive LRD problem. The formulation integrates location and capacity choices for plants and warehouses with supplier and transportation mode selection as well as outsourcing options over multiple time periods. As will be seen, the model also can be used for designing a new logistics network.

3.1 Input data and decision variables

Table 3.1 describes the index sets that are used, while Table 3.2 summarizes all costs. For notational convenience, we introduce the set OD with all origin-destination pairs in the logistics network (recall Figure 1-1):

$$\begin{aligned} OD \;=\; & \{(s,\,\ell) : s \in S,\, \ell \in L\} \cup \{(\ell,\,w) : \ell \in L,\, w \in W\} \cup \\ & \{(\ell,\,c) : \ell \in L,\, c \in C\} \cup \{(w,\,c) : w \in W,\, c \in C\} \end{aligned}$$

The links $(s,\ \ell)$ are used to transport raw materials $r \in R$, while all other origin-destination pairs concern links to move end products, $p \in P$.

Set	Description
T	Time periods
S	Suppliers
L^e	Existing plants at the beginning of the time horizon
L^n	Potential sites for locating new plants
$L = L^e \cup L^n$	All plant locations
K_L	Discrete capacity sizes that can be installed in plant locations
W^e	Existing warehouses at the beginning of the time horizon
W^n	Potential sites for locating new warehouses
$W = W^e \cup W^n$	All warehouse locations
K_W	Discrete capacity sizes that can be installed in warehouse locations
C	Customer zones
R	Raw materials
P	End products
M	Transportation modes
OD	All origin-destination pairs in the logistics network

Table 3.1: Index sets.

The parameters $FC_{t,j}$ and $SC_{t,j}$ reflect fixed costs associated with location decisions. The first term comprises the fixed costs of setting up an infrastructure for a new plant or warehouse (e.g., property acquisition) in time period t. Facility closing costs (the second term) are incurred when an existing plant or warehouse is closed in period t. Other facility costs concern the acquisition of capacity in both new and existing locations and the operation of that capacity. To this end, $IC_{t,j,k} = I_{t,j,k} + \sum_{\tau=t}^{|T|} OC^n_{\tau,j,k}$, where $I_{t,j,k}$ denotes the fixed installation cost of capacity level $k \in K_L \cup K_W$ in location $j \in L \cup W$ in period $t \in T$, and $OC^n_{\tau,j,k}$ is the fixed operating cost in period $\tau \geq t$. The fixed costs $IC_{t,j,k}$ reflect economies of scale. Expansion plans may result in extending the capacity of existing facilities and/or establishing new facilities with given capacity.

Symbol	Description
$FC_{t,j}$	Fixed cost of establishing a new facility in location $j \in L^n \cup W^n$ in period $t \in T$
$SC_{t,j}$	Fixed cost of closing an existing facility in location $j \in L^e \cup W^e$ in period $t \in T$
$IC_{t,j,k}$	Cost of installing capacity level $k \in K_L \cup K_W$ in location $j \in L \cup W$ in period $t \in T$
$OC_{t,j}$	Fixed cost of operating existing facility $j \in L^e \cup W^e$ in period $t \in T$
$PC_{t,s,r}$	Cost of procuring one unit of raw material $r \in R$ from supplier $s \in S$ in period $t \in T$
$MC_{t,\ell,p}$	Cost of manufacturing one unit of product $p \in P$ at plant $\ell \in L$ in period $t \in T$
$TC_{t,o,d,i,m}$	Cost of shipping of one unit of item $i \in R \cup P$ using transportation mode $m \in M$ from origin $o \in S \cup L \cup W$ to destination $d \in L \cup W \cup C$ in period $t \in T$
$EC_{t,w,p}$	Cost of purchasing one unit of product $p \in P$ from an external source by warehouse $w \in W$ in period $t \in T$

Table 3.2: Fixed and variable costs.

Logistics costs include procurement, production, outsourcing, and distribution costs. The latter depend on the choice of the transportation modes for moving raw materials and end products through the network. The available transportation modes differ with respect to their variable costs and capacities. For example, railroad and road freight transport may be viable options with known trade-offs (cost, service time, environmental impact). External costs concern the purchase of end products from other business firms. The consolidation of outsourced products takes place at the warehouses. A pre-specified percentage of the total customer demand for each end product sets an upper bound on the total quantity that can be acquired from an external source. This option may be attractive when the cost of setting up a new facility to process given items is higher than the cost of outsourcing them.

Table 3.3 introduces additional input parameters. We assume that the available capacity levels are sorted in non-decreasing order of their sizes. For plant locations $\ell \in L$ this means that $Q_{\ell,1} < Q_{\ell,2} < \ldots < Q_{\ell,|K_L|}$. Similar conditions hold for warehouse locations. As distinct technologies may be used by different plants to manufacture a given product $p \in P$, we consider a production consumption factor $\mu_{\ell,p}$ for every plant location $\ell \in L$. Moreover, the quantity of raw materials required to manufacture one unit of a specific product may also differ among the various plants. Plant-dependent BOMs are specified by the parameters $a_{\ell,r,p}$. In contrast, the usage of warehouse capacity by an end product does not depend on the warehouse location, thus a consumption factor γ_p is assumed for every $p \in P$.

Constant α_j is used to specify a minimum throughput level at facility j, the latter being defined by multiplying α_j by the capacity installed in location $j \in L \cup W$. In this way, it will be guaranteed that facilities are operated at least at a meaningful level.

All decisions are implemented at the beginning of each time period. As indicated in Table 3.4, strategic decisions on facility location and capacity acquisition are ruled by binary variables, while tactical logistics decisions are described by continuous variables. The statuses of *new* facilities (i.e. plants, warehouses) over the time horizon are controlled by the variables $y^n_{t,j}$. If a new plant or warehouse is established at the beginning of period t in site $j \in L^n \cup W^n$ then $y^n_{t,j} = 1$ and $y^n_{\tau,j} = 0$ for all other periods $\tau \in T, \tau \neq t$. Regarding the *existing* facilities, if facility $j \in L^e \cup W^e$ ceases to operate at the beginning of period t then $y^e_{t,j} = 1$ and $y^e_{\tau,j} = 0$ for all periods $\tau \in T, \tau \neq t$. Observe that if a *new* facility j is available in period t then $\sum_{\tau=1}^{t} y^n_{\tau,j} = 1$. Similarly,

Symbol	Description
$a_{\ell,r,p}$	Number of units of raw material $r \in R$ required to manufacture one unit of product $p \in P$ in plant $\ell \in L$
$QS_{t,s,r}$	Capacity of supplier $s \in S$ for raw material $r \in R$ in period $t \in T$
$\mu_{\ell,p}$	Production capacity usage by one unit of product $p \in P$ in plant $\ell \in L$
γ_p	Handling capacity usage by one unit of product $p \in P$ in a warehouse
Q^e_j	Capacity of existing facility $j \in L^e \cup W^e$ available at the beginning of the time horizon
$Q_{j,k}$	Capacity of level $k \in K_L \cup K_W$ that can be installed in facility $j \in L \cup W$
$\overline{Q}_j$	Maximum overall capacity of facility $j \in L \cup W$ in each time period
$QM_{t,o,d,m}$	Capacity of transportation mode $m \in M$ from origin $o \in S \cup L \cup W$ to destination $d \in L \cup W \cup C$ in period $t \in T$
$\sigma_{i,m}$	Capacity utilization factor by one unit of item $i \in R \cup P$ in transportation mode $m \in M$
$d_{t,c,p}$	Demand of customer zone $c \in C$ for product $p \in P$ in period $t \in T$
$\lambda_{t,p}$	Fraction of the total demand for product $p \in P$ in period $t \in T$ that can be delivered directly from plants to customers; $0 \leq \lambda_{t,p} < 1$
$\beta_{t,p}$	Fraction of the total demand for product $p \in P$ in period $t \in T$ that can be supplied by an external source; $0 \leq \beta_{t,p} < 1$
α_j	Number between zero and one for a facility $j \in L \cup W$ (to set a minimum throughput level at the facility)

Table 3.3: Additional input parameters.

if an *existing* facility j is operated in period t then $\sum_{\tau=1}^{t} y^e_{\tau,j} = 0$.

3.2 Network redesign constraints

In this section, we describe in detail the specific constraints that compose our LRD problem.

Symbol	Description
$y^n_{t,j}$	1 if a new facility is established in location $j \in L^n \cup W^n$ in period $t \in T$, 0 otherwise
$y^e_{t,j}$	1 if an existing facility in location $j \in L^e \cup W^e$ is closed at the beginning of period $t \in T$, 0 otherwise
$u_{t,j,k}$	1 if capacity level $k \in K_L \cup K_W$ is installed in location $j \in L \cup W$ in period $t \in T$, 0 otherwise
$x_{t,o,d,i,m}$	Quantity of item $i \in R \cup P$ shipped in period $t \in T$ from origin o to destination d, $(o,\ d) \in OD$, using transportation mode $m \in M$
$z_{t,w,p}$	Quantity of product $p \in P$ provided by an external source to warehouse $w \in W$ in period $t \in T$

Table 3.4: Decision variables.

Supplier-related constraints

Supplier selection and procurement of raw materials are ruled by the following conditions:

$$\sum_{\ell \in L} \sum_{m \in M} x_{t,s,\ell,r,m} \leq QS_{t,s,r} \qquad t \in T,\ s \in S,\ r \in R \tag{3.1}$$

$$\sum_{s \in S} \sum_{m \in M} x_{t,s,\ell,r,m} = \sum_{p \in P} a_{\ell,r,p} \left(\sum_{w \in W} \sum_{m \in M} x_{t,\ell,w,p,m} + \sum_{c \in C} \sum_{m \in M} x_{t,\ell,c,p,m} \right) \qquad t \in T,\ \ell \in L,\ r \in R \tag{3.2}$$

Constraints (3.1) limit the quantity of each raw material provided by a supplier. Equalities (3.2) ensure that the total amount of raw material r purchased by a plant is used to manufacture end products. Observe that the expression within brackets on the right-hand side defines the total quantity of product p manufactured in time period t in plant ℓ. End products can be distributed to warehouses or delivered directly to customer zones.

Facility location and capacity acquisition constraints

The following constraints impose the required conditions for setting up new facilities and closing existing facilities over the planning horizon. In addition, they also rule the installation of capacity and its expansion. For notational convenience, let us denote $K_j = K_L$ if $j \in L$ and $K_j = K_W$ if $j \in W$.

$$\sum_{t \in T} y^n_{t,j} \leq 1 \qquad j \in L^n \cup W^n \tag{3.3}$$

$$y^n_{t,j} \leq \sum_{k \in K_j} u_{t,j,k} \leq \sum_{\tau=1}^{t} y^n_{\tau,j} \qquad t \in T,\ j \in L^n \cup W^n \tag{3.4}$$

$$\sum_{k \in K_j} u_{t,j,k} \leq 1 - \sum_{\tau=1}^{|T|} y^e_{\tau,j} \qquad t \in T,\ j \in L^e \cup W^e \tag{3.5}$$

$$\sum_{k \in K_j} Q_{j,k} \sum_{t \in T} u_{t,j,k} \leq \overline{Q}_j \sum_{t \in T} y^n_{t,j} \qquad j \in L^n \cup W^n \tag{3.6}$$

$$Q^e_j \left(1 - \sum_{t \in T} y^e_{t,j}\right) + \sum_{k \in K_j} Q_{j,k} \sum_{t \in T} u_{t,j,k} \leq \overline{Q}_j \left(1 - \sum_{t \in T} y^e_{t,j}\right) j \in L^e \cup W^e \tag{3.7}$$

Constraints (3.3) impose that at most one new facility (plant/warehouse) can be established in a potential location over the time horizon. Moreover, once open, new facilities cannot be closed. Constraints (3.4), resp. (3.5), rule the installation of capacity in new, resp. existing, locations. In each time period, at most one capacity level can be selected provided that a facility is already operating in that site. On the other hand, if a new facility is established in a given period then a capacity level must also be acquired in the same period. Moreover, constraints (3.5) guarantee that if an existing facility has its capacity extended then it cannot be closed. Since the location variables $y^e_{t,j}$ are binary, the term $\sum_{\tau=1}^{|T|} y^e_{\tau,j}$, which appears on the right-hand side of (3.5), takes a non-negative integer value. If this term is zero then facility $j \in L^e \cup W^e$ remains in operation over the whole time horizon. As a result, the right-hand side of (3.5) is equal to one, thereby setting an upper bound on the number of capacity expansions that are permitted in each period. If $\sum_{\tau=1}^{|T|} y^e_{\tau,j} = 1$ then facility j was closed in some period τ and so the right-hand side of (3.5) becomes zero. In this case, the facility cannot be expanded in any period (recall that the variables $u_{t,j,k}$ that appear on the left-hand side of (3.5) are binary). Finally, notice that values of $\sum_{\tau=1}^{|T|} y^e_{\tau,j}$ larger than one are not possible as this would yield a negative right-hand side of (3.5). These constraints would become invalid since $u_{t,j,k}$ represent binary variables.

Constraints (3.6), resp. (3.7), set a limit on the capacity expansion of each new, resp. existing, plant or warehouse.

Constraints related to capacity utilization

The next sets of constraints establish minimum and maximum capacity utilization limits on both plants and warehouses. These limits depend on the acquisition of capacity over time. In every location, the capacity available in each period is determined by

$$\begin{aligned} A^n_{t,j} &= \sum_{k \in K_j} Q_{j,k} \sum_{\tau=1}^{t} u_{\tau,j,k} && t \in T,\ j \in L^n \cup W^n \\ A^e_{t,j} &= Q^e_j \left(1 - \sum_{\tau=1}^{t} y^e_{\tau,j}\right) + \sum_{k \in K_j} Q_{j,k} \sum_{\tau=1}^{t} u_{\tau,j,k} && t \in T,\ j \in L^e \cup W^e \end{aligned}$$

The above capacities are used for processing end products at plants and warehouses. For ease of exposition, we introduce two expressions for the total quantity processed by plants and warehouses:

$$
\begin{aligned}
F^L_{t,j} &= \sum_{p\in P} \mu_{j,p} \left(\sum_{w\in W} \sum_{m\in M} x_{t,j,w,p,m} + \sum_{c\in C} \sum_{m\in M} x_{t,j,c,p,m} \right) && t \in T,\ j \in L \\
F^W_{t,j} &= \sum_{p\in P} \gamma_p \sum_{c\in C} \sum_{m\in M} x_{t,j,c,p,m} && t \in T,\ j \in W
\end{aligned}
$$

Constraints (3.8) and (3.9) guarantee that the total quantity of products manufactured by each plant must be within pre-defined lower and upper limits in each time period. Constraints (3.10)–(3.11) impose similar conditions on warehouses. Observe that the lower capacity utilization limits refer to minimum throughputs corresponding to a given percentage of the maximum capacities.

$$\alpha_j\, A^n_{t,j} \le F^L_{t,j} \le A^n_{t,j} \qquad t \in T,\ j \in L^n \tag{3.8}$$

$$\alpha_j\, A^e_{t,j} \le F^L_{t,j} \le A^e_{t,j} \qquad t \in T,\ j \in L^e \tag{3.9}$$

$$\alpha_j\, A^n_{t,j} \le F^W_{t,j} \le A^n_{t,j} \qquad t \in T,\ j \in W^n \tag{3.10}$$

$$\alpha_j\, A^e_{t,j} \le F^W_{t,j} \le A^e_{t,j} \qquad t \in T,\ j \in W^e \tag{3.11}$$

Transportation, outsourcing, and demand-related constraints

$$\sum_{i\in R\cup P} \sigma_{i,m}\, x_{t,o,d,i,m} \leq QM_{t,o,d,m} \qquad t\in T,\ (o,d)\in OD,\ m\in M \quad (3.12)$$

$$\sum_{\ell\in L}\sum_{c\in C}\sum_{m\in M} x_{t,\ell,c,p,m} \leq \lambda_{t,p} \sum_{c\in C} d_{t,c,p} \qquad t\in T,\ p\in P \quad (3.13)$$

$$\sum_{c\in C}\sum_{m\in M} x_{t,w,c,p,m} = \sum_{\ell\in L}\sum_{m\in M} x_{t,\ell,w,p,m} + z_{t,w,p} \qquad t\in T,\ w\in W,\ p\in P \quad (3.14)$$

$$\sum_{w\in W} z_{t,w,p} \leq \beta_{t,p} \sum_{c\in C} d_{t,c,p} \qquad t\in T,\ p\in P \quad (3.15)$$

$$\sum_{\ell\in L}\sum_{m\in M} x_{t,\ell,c,p,m} + \sum_{w\in W}\sum_{m\in M} x_{t,w,c,p,m} = d_{t,c,p} \qquad t\in T,\ c\in C,\ p\in P \quad (3.16)$$

Capacity constraints on individual transportation modes are imposed by (3.12). For supplier-plant links (s,ℓ), these constraints rule the utilization of transportation capacity for raw materials $r \in R$, while for all other origin-destination pairs equalities (3.12) involve end products. Inequalities (3.13) limit the quantity of direct deliveries from plants to customers for every product. Constraints (3.14) guarantee the conservation of product flows for all operated warehouses in each time period. These constraints along with inequalities (3.10) and (3.11) state that outsourced products also use the handling capacity available at warehouses. Constraints (3.15) impose an upper limit on the total outsourced quantity per product type. Finally, constraints (3.16) ensure the satisfaction of all customer demands over the time horizon.

Domains of variables

The following constraints (3.17)-(3.21) represent non-negativity and binary conditions.

$$x_{t,o,d,i,m} \geq 0 \qquad t \in T,\ (o,d) \in OD,\ i \in R \cup P,\ m \in M \tag{3.17}$$

$$z_{t,w,p} \geq 0 \qquad t \in T,\ w \in W,\ p \in P \tag{3.18}$$

$$y^n_{t,j} \in \{0,1\} \qquad t \in T,\ j \in L^n \cup W^n \tag{3.19}$$

$$y^e_{t,j} \in \{0,1\} \qquad t \in T,\ j \in L^e \cup W^e \tag{3.20}$$

$$u_{t,j,k} \in \{0,1\} \qquad t \in T,\ j \in L \cup W,\ k \in K_L \cup K_W \tag{3.21}$$

3.3 Objective function

The objective function (3.22) minimizes the overall strategic and logistics costs. Fixed facility costs are given by the first four components, while variable costs are determined by the remaining terms.

$$\begin{aligned}
\text{Min} \; & \sum_{t \in T} \sum_{j \in L^n \cup W^n} FC_{t,j}\, y^n_{t,j} + \sum_{t \in T} \sum_{j \in L^e \cup W^e} SC_{t,j}\, y^e_{t,j} + \sum_{t \in T} \sum_{j \in L \cup W} \sum_{k \in K_j} IC_{t,j,k}\, u_{t,j,k} + \\
& \sum_{t \in T} \sum_{j \in L^e \cup W^e} OC_{t,j} \left(1 - \sum_{\tau=1}^{t} y^e_{\tau,j} \right) + \sum_{t \in T} \sum_{w \in W} \sum_{p \in P} EC_{t,w,p}\, z_{t,w,p} + \\
& \sum_{t \in T} \sum_{s \in S} \sum_{\ell \in L} \sum_{r \in R} \sum_{m \in M} \left(PC_{t,s,r} + TC_{t,s,\ell,r,m} \right) x_{t,s,\ell,r,m} + \\
& \sum_{t \in T} \sum_{\ell \in L} \sum_{j \in W \cup C} \sum_{p \in P} \sum_{m \in M} \left(MC_{t,\ell,p} + TC_{t,\ell,j,p,m} \right) x_{t,\ell,j,p,m} + \\
& \sum_{t \in T} \sum_{w \in W} \sum_{c \in C} \sum_{p \in P} \sum_{m \in M} TC_{t,w,c,p,m}\, x_{t,w,c,p,m}
\end{aligned} \tag{3.22}$$

The problem of re-designing an existing logistics network is modeled by the objective function (3.22) subject to the constraints (3.1)–(3.21). We remark that for $L^e = \emptyset$ and $W^e = \emptyset$, the formulation reduces to the special case of designing a new network.

3.4 Model enhancements

There are various ways of enhancing the LRD model. A simple strategy is to multiply the right-hand side of the capacity constraints (3.12) by appropriate sets of facility location variables. For the flow of end products from plants and warehouses to customer zones, this corresponds to replacing inequalities (3.12) by

$$\sum_{p \in P} \sigma_{p,m}\, x_{t,j,c,p,m} \leq QM_{t,j,c,m} \sum_{\tau=1}^{t} y^{n}_{\tau,j} \qquad t \in T,\ j \in L^n \cup W^n,\ c \in C,\ m \in M \tag{12a}$$

$$\sum_{p \in P} \sigma_{p,m}\, x_{t,j,c,p,m} \leq QM_{t,j,c,m} \left(1 - \sum_{\tau=1}^{t} y^{e}_{\tau,j}\right) \quad t \in T,\ j \in L^e \cup W^e,\ c \in C,\ m \in M \tag{12b}$$

Similar transformations can be easily performed for transportation modes linking suppliers to plants. In this case, raw materials cannot be moved from supplier s to plant ℓ with transportation mode m in time period t unless plant ℓ is operated in that period. For plant-warehouse pairs, transportation modes cannot be selected unless the origin plant and the destination warehouse are both operated. In this case, it is necessary to duplicate inequalities (3.12) for these types of transportation links, thereby adding $|T| \cdot |L| \cdot |W| \cdot |M|$ new constraints to the LRD model. We opted to transform inequalities (3.12) without increasing the model size. To this end, for plant-warehouse pairs we extend the right-hand side of (3.12) to ensure that products are only distributed to warehouses in operation.

Chapter 4

Memetic Algorithm

A memetic algorithm (MA) is an evolutionary algorithm that makes use of a local search heuristic. The general idea behind an MA is to combine the advantages of evolutionary operators that identify interesting regions of the search space, with local neighbourhood search that quickly finds good solutions in a small region of the search space.

The primary objective of the developed MA is to set the status and size of facilities (the values of the binary decision variables) in order to obtain feasible solutions to LND/LRD problems. In this algorithm, a solution/an individual is composed by a set of closed and open facilities. The MA includes the following procedures:

1. Create initial population - P(0)

2. Apply local search to P(0)

3. Binary tournament selection

4. Crossover

5. Facility status correction procedure

6. Mutation

7. Capacity correction procedure

8. Apply local search to new population

Two criteria are used to stop the MA procedure. One is by setting a maximum number of generations. Thereby, when the MA achieves the established number of generations the procedure terminates. The other criterion defines a maximum number of generations without improvement of the best solution. Currently, the MA procedure stops if there is not an improvement of the best solution over 60% of generations.

4.1 Genetic algorithm

In this work, the genetic representation is set up of two chromosomes, the S-chromosome that represents the statuses of the facilities (i.e. open/closed) in every period and the F-chromosome that describes the capacity sizes available at the active facilities. The number of chromosomes in each solution is equal to the number of time periods while the number of genes in each chromosome is equal to the number of existing and new facilities. However, in the F-chromosome, each gene has three alleles referring to three different capacity levels k ($k \in K_L \cup K_W$).

For each given choice of the binary variables (S and F-chromosomes), the original problem reduces to solving a linear program (with continuous variables representing the flow of products and the outsourcing quantities). Each sub-problem is solved to optimality with Cplex. Its objective function value represents the solution fitness.

4.1.1 Initial Population

The initial population procedure aims at creating a set of solutions that will be used by the algorithm to generate new individuals in the next generations. In order to create feasible solutions, the following features are considered:

- The total capacity available in time period t (at plants and warehouses) must cover the demand requirements in that period ($t \in T$).
- In every time period t, the capacity installed in each facility j ($j \in L \cup W$) cannot exceed the global capacity limit of that facility.

Although these two features are crucial to ensure the feasibility of a solution, they do not guarantee the quality of the solution with respect to the objective function value. Naturally, the random selection of facilities to be operated does not ensure that the best set of facilities is selected to satisfy the demand. In fact, the initial solutions are typically "expensive" as they represent logistics networks with more capacity than the one needed to cover the demand. In order to readjust the choice of capacity sizes, a local search procedure is applied. Often, this technique helps to improve the quality of the initial solutions.

The initial population, P(0), is initialized with four solutions, independently of the number of time periods or customer zones. Preliminary numerical tests indicated that this seems to be a good choice. Observe that initializing P(0) with more individuals will have a strong impact on the time performance of the algorithm. On the other hand, taking less individuals will have a negative effect on the population diversification.

After each new generation, the new individuals are added to the previous population, originating the new population.

Generation of the initial population:

1. For each period[t], ($t \in T$)

   1.1 Determine the minimum demand requirements to be satisfied by plants and warehouses in period[t] ($t \in T$)

   1.2 Separately for existing plants and warehouses ($j \in L^e \cup W^e$), evaluate available capacity in period[t]

      1.2.1 If available capacity is not enough to satisfy the demand in period[t] then randomly select a facility[j] ($j \in L \cup W$)

      1.2.2 Determine installed capacity at selected facility j until period $t - 1$

      1.2.3 Select the most suitable size k ($k \in K_L \cup K_W$), starting with the smallest size

2. Apply local search (Section 4.2, p. 40)

3. Evaluate solution quality using Cplex to solve the subproblem

4.1.2 Genetic operators

Binary tournament selection

From the pool of N individuals (with N denoting the number of existing parents), N-1 individuals will be selected by applying a binary tournament. This is equivalent to selecting the N-1 best individuals of the pool, but in this procedure, the randomness of the two solutions selected to participate in the binary tournament defines the sequence of parents that will be used in the crossover. For example, combining the winner of the binary tournament [i] with the winner of the binary tournament [i+1] ($i+1 \leq N-1$) will yield one pair of parents for crossover. Hence, in the next genetic operator, the pairs of parents are already established.

Crossover

This genetic operator will act through two random crossover points, OX_j, j $\in \{1,2\}$, as proposed by Michalewicz and Fogel [20]. Thus, three arguments will be created that will influence the generation of the offsprings. The three arguments are as follows:

- $[1; OX_1]$
- $]OX_1; OX_2]$
- $]OX_2; |T|]$

The offspring [i] will be created by combining the genes of the S-chromosome and F-chromosome of the two parents. This offspring will receive new genes from the parent $[i+1]$ by changing the 1^{st} and 3^{rd} arguments given above and keeping the original genes from the 2nd argument belonging to parent [i].

The number of crossovers (q) depends on the number of parents available, however the total number of parents can be an even or odd number. As a result, if the total number of parents is odd, the parent[q] will not have a pair to generate a son. To prevent this situation, the following formula is used:

$$\lfloor q \rfloor = \left\lfloor \frac{\text{Number of parents in current generation} - 0.5}{2} \right\rfloor$$

For example, if the number of available parents is 5, the crossover procedure can only be applied twice. In the first one, the procedure will be applied to parent[1] and parent[2] and in the second time to parent[3] and parent[4].

Crossover procedure:

1. Select randomly crossover points OX_j, j $\in \{1, 2\}$
2. Select two parents, parent[i] and parent [i+1], $i \in \{1, ..., q\}$

   2.1 From $[1, OX_1]$, transfer the genes of the S-chromosome and F-chromosome:

   2.1.1 parent[i] to son[i+1]

   2.1.2 parent[i+1] to son[i]

   2.2 From $]OX_1, OX_2]$, transfer the genes of the S-chromosome and F-chromosome:

   2.2.1 parent[i] to son[i]

   2.2.2 parent[i+1] to son[i+1]

   2.3 From $]OX_2, |T|]$, transfer the genes of the S-chromosome and F-chromosome:

2.3.1 parent[i] to son[i+1]

2.3.2 parent[i+1] to son[i]

3. Call "Facility status correction" procedure

At this stage, it is required to take into account the possibility that a facility is assigned an invalid status. For example, if in the 2nd argument some facilities have installed capacity, then it is necessary to verify if those facilities are open. If this is not the case then a status correction procedure is applied.

Facility status correction procedure:

1. For each solution[i] ($i \in N$), determine in the S-chromosome the open period[t] ($t \in T$) of each facility[j] ($j \in L \cup W$)

   1.1 If gene[t][j] of facility[j] = 1: S-chromosome[τ] : for period[τ] ($\tau \in \{t+1, ..., |T|\}$), set genes $[\tau][j] = 0$

2. For each solution[i] ($i \in N$), analyze the installed capacity size k ($k \in K_L \cup K_W$), in period[t] ($t \in T$) for each facility[j] ($j \in L \cup W$)

   2.1 If facility size[t][j][k] = 1, update installed capacity[t][j].

   2.2 If installed capacity[t][j] > Global capacity of facility[j]

   2.2.1 F-chromosome: from period $[\tau](\tau \in \{t, ..., |T|\})$, set genes $[\tau][j][k] = 0$

   2.2.2 S-chromosome: set genes $[\tau][j] = 0$

Mutation

Mutation is an operator that allows the diversification of the population and is applied to offsprings generated by the crossover procedure. However, not all elements benefit from this operator. With a mutation probability (Pm), as indicated by Mühlenbein et al. [21], $Pm = \frac{1.7}{N \times |T|}$, the mutation probability decreases in each new generation due to adding new individuals to the population. As a result, mutation is more likely to be used in early generations and only when a random number (random mutation $\in [0, 1]$) is less than Pm.

This procedure is applied to facilities by selecting randomly one active facility (current facility) and another that is not operated yet (new facility). After selection, the genes of the S-chromosome and F-chromosome are swapped from the current facility to the new selected facility.

Mutation procedure:

1. Calculate Pm
2. Select son[i] for mutation
3. Generate a random continuous number, random nutation (random mutation $\in$ $[0, 1]$), to decide if mutation is applied to current son[i]. If random mutation $<$ Pm, go to 4. Otherwise, mutation is not applied.
4. Select randomly two genes[j], $j \in \ L \cup W$

   4.1 If 1$^{\text{st}}$ random $j \in L$, then the 2$^{\text{nd}}$ j will also be a plant ($j \in L$). Else, the 2$^{\text{nd}}$ facility j is a warehouse ($j \in W$)

5. Swap S-chromosome and F-chromosome

6. Call "Facility capacity correction" procedure

7. Select son[i+1] and go to 3

After this procedure, it is required to take into account the transferred capacity into the new facility, because the global capacities are different from facility to facility and may transform a feasible solution into an infeasible one. To prevent this from happening, a capacity correction procedure is applied after mutation. This procedure analyzes the installed capacity of a facility during its active time periods and verifies if the global capacity limit is satisfied.

In case of an infeasible solution, the capacity correction procedure will restrict the installed capacity of the facility and in addition, a new facility will be opened to satisfy the demand[t] that otherwise could not be covered.

Facility capacity correction procedure:

1. For each open facility[j] ($j \in L \cup W$) analyze the already installed size k ($k \in K_L \cup K_W$), in period[t] ($t \in T$). Set k = current size

2. For each period, analyse if Installed capacity[t] $<$ Total demand[t] and facility size[t][j][k] $=$ 1

   2.1 For each facility[t][j], regarding the global capacity of facility[j], select the most suitable size k, defined as new size k (current size $<$ new size $\leq k$).

   2.2 If installed capacity[t] with new size $\leq$ Global capacity of facility[j]

      2.2.1 S-chromosome: update the period in which the facility is opened

2.2.2 F-chromosome: for facility[j] in period t, set current size = 0 and new size = 1

3. If Installed capacity[t] < Total demand[t], it is required to open a new facility

3.1 Select randomly a new non-active facility

3.2 Starting with k = 1 (current size), select the most appropriate size (new size) for facility[j], as follows

3.2.1 If facility size[t][j][k] ≤ Global capacity of facility[j], then

3.2.2 S-chromosome: if required, update the period in which the facility is opened

3.2.3 F-chromosome: set current size = 0 and new size = 1

4. If Installed capacity[t] < Total demand[t], return to step 3

4.2 Local search

This procedure tries to improve the fitness of new individuals. To start this procedure two facilities belonging to the same network layer are randomly selected. The first corresponds to an active facility (current facility) and the second one to a non-operated facility (new facility). In order to keep diversity, a facility (new facility) can only be selected once.

After a new facility has been selected, the genes of the S-chromosome and the F-chromosome are swapped. At this stage, due to the different characteristics of the facilities, the sizes and the global capacity of facilities must be taken into account.

As a result of these differences, four possibilities can occur by changing the current size of the new facility:

- The current size installed at facility[j] at period[t] cannot supply demand and a larger size must be selected.
- The current size can supply demand and the global capacity is satisfied. As a result, no changes are required.
- Due to the different capacity sizes of the new facility it is possible to select a smaller size.
- The solution has an excess of capacity and it is possible not to use the selected facility. This may occur when at previous local searches, the surplus of installed capacity is much higher than the required to satisfy demand.

In the current procedure, the term "most feasible new size" is introduced and it means that the selected size is the best option because it does not compromise the global capacity of the facility[j] and also allows the satisfaction of demand[t].

With this procedure, 15% of the neighbours available in the solution are visited. A higher percentage of neighbors visited will lead to an increase of the computational time required to solve the problem and a lower percentage will have a residual impact on the diversification of the solutions. Preliminary tests indicated that 15% is an appropriate value for this parameter. The total number of neighbours of a solution is equal to the number of existing and new facilities. For example, if a solution contains two existing facilities and 10 new facilities, the number of local searches will be equal to $\lceil(2+10)\times 0.15\rceil = 2$.

The purpose of this procedure is to find a facility that is able to satisfy demand[t], but it is less expensive than the "initial facility". However, if there is enough installed capacity to supply demand[t] without the initial/new facility, this facility is removed from the solution.

Local search procedure:

1. Select initial facility: randomly choose a facility that is operated ($j \in L \cup W$)
2. Select new facility: randomly choose a non-operated facility[j']
3. Swap the genes of the S-chromosome and F-chromosome from the initial facility to the new facility
   - 3.1 Determine the current size k ($k \in K_L \cup K_W$) installed
   - 3.2 Select the most feasible new size k or if possible remove current facility[j]
4. If current size $\neq$ new size, evaluate solution using Cplex to solve the subproblem and keep new solution, independently if it improves or not.
5. If current size = new size, apply local search to a 2^{nd} neighbourhood
   - 5.1 To another "new" facility, apply steps 1, 2 and 3
   - 5.2 If new solution has enough capacity to satisfy demand[t] ($t \in T$), the global capacity of facility[j] is satisfied and the new solution value is better than the previous one, keep the new solution. Else, restore initial solution.
6. If the number of local searches $<$ $\lceil$(number of total facilities)$\times 0.15\rceil$, go to 1. Else, stop

Chapter 5

Computational results

In total, 60 test instances were randomly generated with different time periods and different number of customers (50, 75, and 100). Half of the instances belong to the network redesign (LRD) class and the other half to the network design (LND) class. The interested reader is referred to Cortinhal et al. [8] for further details. Six different combinations of the parameters $\lambda_{t,p} = \lambda$ and $\beta_{t,p} = \beta$ were considered as follows: $\lambda_{t,p} \in \{0\%, 10\%\}$, $\beta_{t,p} \in \{0\%, 10\%, 20\%\}$ for every $t \in T$, $p \in P$, thus yielding in total 360 cases. For each parameter combination, a CPU time limit of 8h was set for IBM ILOG Cplex 12.6.0. Both the model and the heuristic were implemented in C++. All experiments were performed on a PC with a 3.5 GHz Intel Core i7-3770K processor, 8 GB RAM and running Windows 7 (64-bit). The heuristic was run with 5 generations. This parameter results from results from running the test instances with different values for the maximum number of generations. A number larger than 5 leads to an increase of the computational time required to create all generations. In contrast, lower quality solutions could be obtained when the number of generations is less than 5. Moreover, the heuristic also stops when there is not an improvement of the best solution over 60% of the generations.

In the next table the sizes of the test instances are presented for both problem variants and $|T| \in \{4, 6\}$. The table gives the average, minimum and maximum number of binary and continuous variables and the number of constraints. Table 5.2 displays a

Problem		$\|T\| = 4$			$\|T\| = 6$		
class	Number of	Avg	Min	Max	Avg	Min	Max
LRD	Binary variables	432	288	560	648	432	840
	Continuous variables	156830	32161	345201	235244	48241	517801
	Constraints	8605	3637	13949	12883	5439	20891
LND	Binary variables	363	240	480	544	360	720
	Continuous variables	129505	26001	289601	194258	39001	434401
	Constraints	8010	3370	13100	11992	5040	19620

Table 5.1: Size of the test instances

classification of the best solutions found by Cplex within the pre-specified time of 8h is given.

	Cplex	Network	
	solutions	LND model	LRD model
$\|T\| = 4$	Optimal	64.44%	74.44%
	Non-Optimal	35.56%	25.56%
$\|T\| = 6$	Optimal	46.67%	67.78%
	Non-Optimal	53.33%	32.22%

Table 5.2: Cplex performance

5.1 Memetic Algorithm versus Cplex

In Figures 5-1 and 5-2, the performance of the heuristic is compared to Cplex. In particular, the objective function values and the CPU time of the best solutions identified by the heuristic and Cplex are compared as follows:

$$Gap = \frac{\text{Heuristic performance} - \text{Cplex performance}}{\text{Cplex performance}}$$

The analysis is focused on two parameters, the average objective value and the average CPU time. The first one is obtained by stratifying the gaps into five groups and for each group the average value gap is calculated. A similar procedure is applied to obtain the average gap time, but in this case the data is split into two groups, the best time of the heuristic versus Cplex (Best) and worst time of heuristic versus Cplex (Worst).

The analysis is conducted by comparing the LND problem class with the LRD problem class in order to allow an easier interpretation and comparison of each network performance.

LND instances require the opening of more new facilities than LRD instances and as a result, more facilities must be selected to satisfy demand. However, due to the absence of existing facilities in the LND problem class, the time required by the heuristic to analyse capacity restrictions is smaller. Consequently, the CPU time performance of the heuristic is better in the LND class than in the LRD class. However, it is harder for the heuristic to find the better combination of individuals that will constitute a solution in the LND instances. As a result, the objective value performance is better in the LRD class.

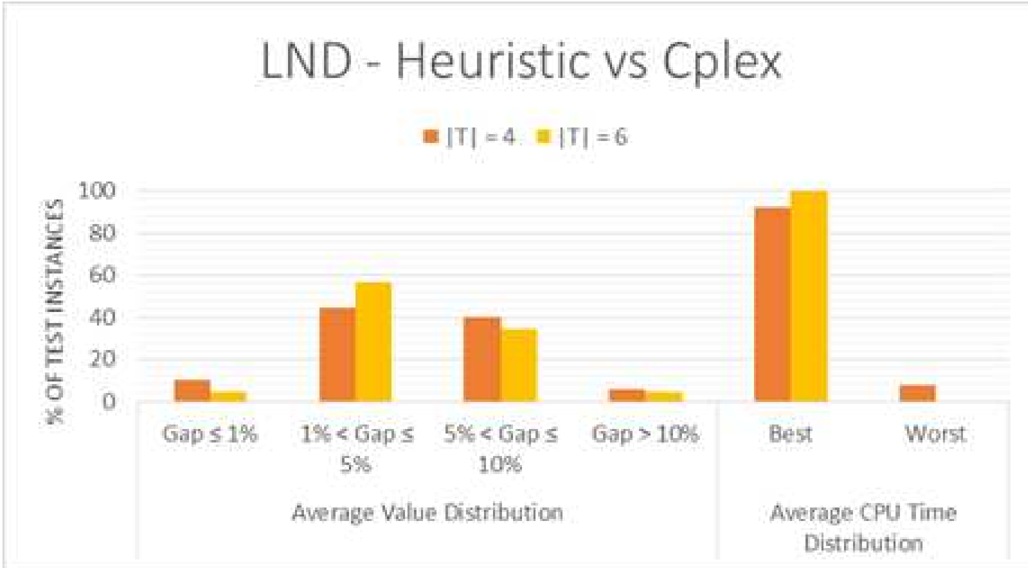

Figure 5-1: LND Model - Heuristic and Cplex performance

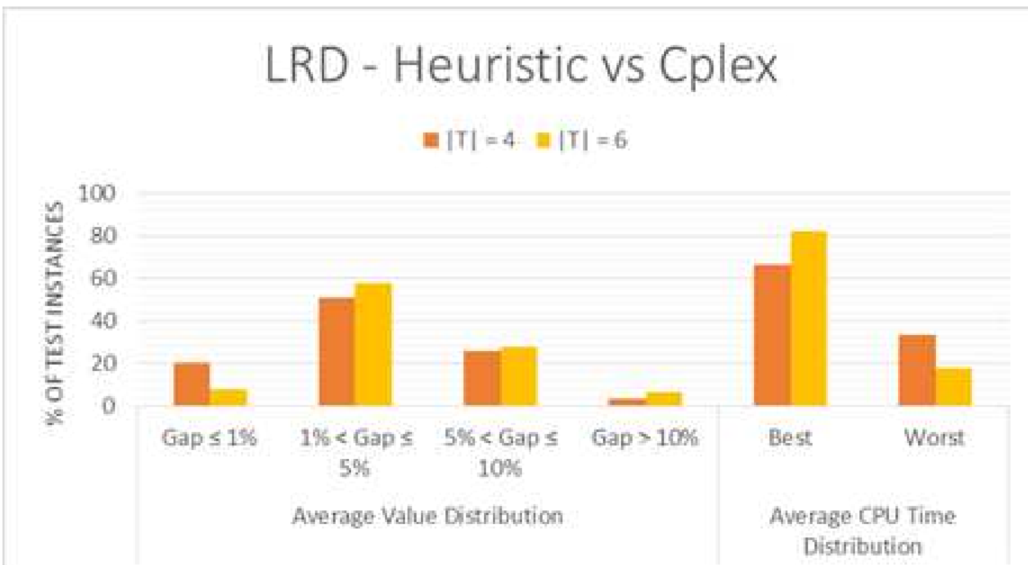

Figure 5-2: LRD Model - Heuristic and Cplex performance

5.2 Performance behaviour of different requirements of outsourcing (β) and direct shipments to clients (λ)

Starting the analysis with LND test instances with 4 time periods, it is possible to verify that different levels of β and λ do not have a significant difference with respect to the objective function value and the CPU time performance. Once a new logistics network is built from the beginning (LND), the gap between the installed capacity and the demand is under the influence of the selected facilities. This means, that more efficient selections will lead to smaller gaps. By analysing Figure 5-3, we verify that the smaller average gap is 3.91% for $\beta = 0\%$ and $\lambda = 10\%$ (at most 10% of the production is sent directly to customer zones). In contrast, the biggest average gap is 5.17% for $\beta = 20\%$ and $\lambda = 10\%$. In this case, at most 20% of the demand is outsourced and 10% of the manufactured products is sent directly to customer zones. There is a significant decrease of capacity handled by plants but not by warehouses. Due to the maximum outsourcing level (20%), less plants are required but the heuristic is not sensitive to minimize the gap between the installed capacity and the required capacity to supply demand. Typically, the installed capacity is much higher than the required capacity, leading to expensive solutions. Figure 5-4 shows that a slightly

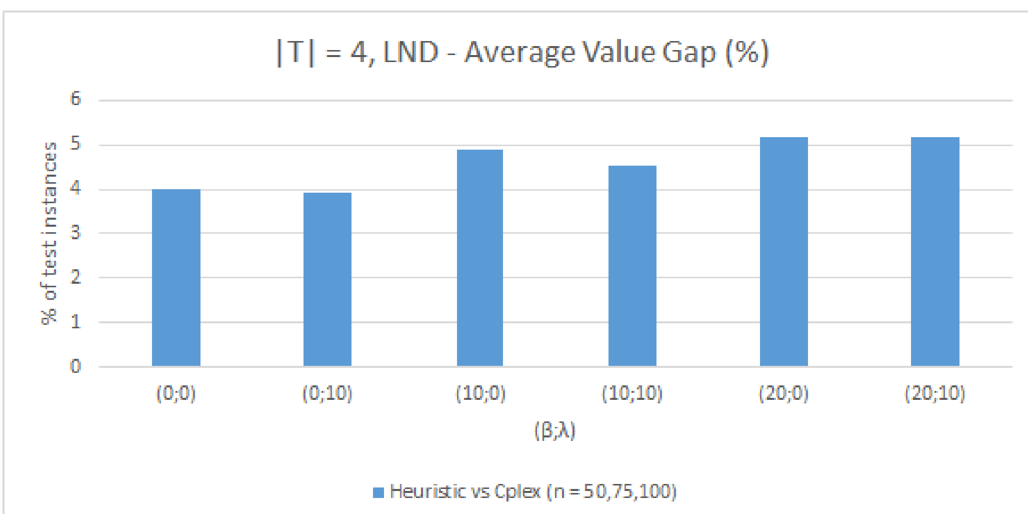

Figure 5-3: LND Model - Value performance, $|T| = 4$

different behaviour is observed regarding the CPU time. This KPI (Key Performance

Indicator) has a better performance for instances with $\beta \in \{0\%,\ 20\%\}$ and $\lambda = 10\%$. When at most 10% of the demand can be shipped directly from plants to customers, less warehouse capacity is necessary and less time is consumed to select a warehouse and a proper size to install. Hence, the time required by the heuristic to find a feasible set of warehouses to satisfy the demand is smaller.

Analysing the transition from $\beta = \{0\%,\ 20\%\}$ and $\lambda = 0\%$ to $\beta = \{0\%,\ 20\%\}$ and $\lambda = 10\%$, it is possible to observe a decrease of performance. Despite 10% of the total demand being shipped directly to customer zones, once again the heuristic is not able to select the most appropriate facilities in order to minimize the gap between the installed capacity and the required capacity to supply demand. The exception is the pair $\beta = \lambda = 10\%$ because the quantity sent by plants, without direct shippment to customers, is equal to the quantity received by warehouses. In this scenario, at most 10% of the plants' production is outsourced and 10% of the demand is shipped directly to customer zones. For $\beta = 10\%$ and $\lambda = 0\%$, the average time gap is -65.56% comparatively to -70.18% for $\beta = 10\%$ and $\lambda = 10\%$. This particular case has a performance similar to the case with $\beta = \lambda = 0\%$ (-71.26%). The performance of the heuristic in the LRD class differs significantly from the LND class. In the latter case, there is no installed capacity in the network at the beginning of the time horizon. In contrast, in the LRD network, at the starting period there is already installed capacity and often, less facilities must be opened to satisfy demand.

Comparing the test instances $\beta \in \{0\%,\ 10\%\}$ and $\lambda = 0\%$ with $\beta \in \{0\%,\ 10\%\}$ and $\lambda = 10\%$, normally it leads to an increase of the average gap, because with $\lambda = 10\%$, the warehouses will handle less 10% of the plants' production and solutions

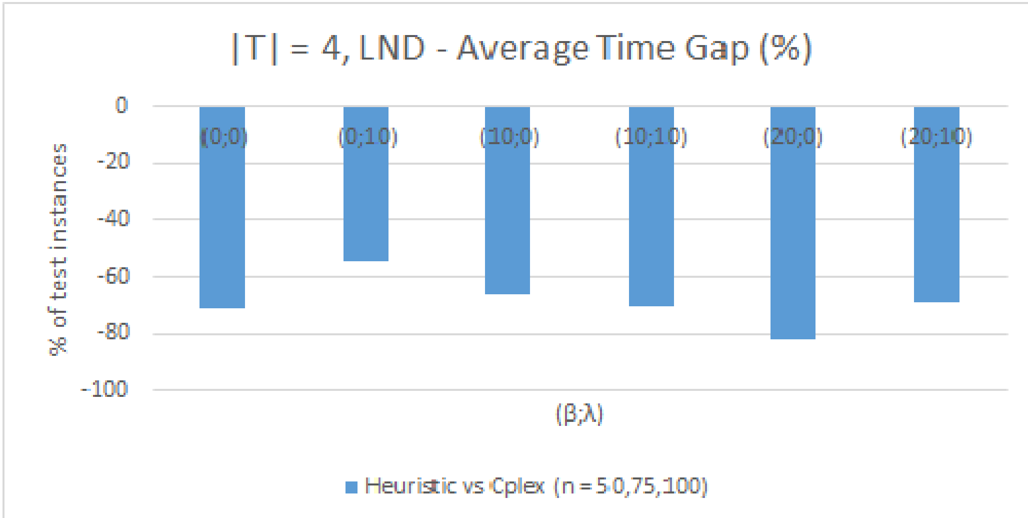

Figure 5-4: LND Model - CPU Time performance, $|T| = 4$

capacities must be adjusted in order to turn solution more efficient in terms of costs. This production is shipped directly to customer zones, but the heuristic is not able to select the most appropriate warehouses to supply demand. As a result, more expensive solutions are created. This is the same reason as with the LRD network, but because there is already installed capacity, less random selections must be done, leading to smaller average value gaps compared to an LND network. In the LRD test instances, the higher average value gap is 5.01% for $\beta = 10\%$ and $\lambda = 10\%$. In contrast, the best average value gap is 1.63% for $\beta = 20\%$ and $\lambda = 10\%$.

The heuristic performance greatly improves for $\beta = 20\%$ and $\lambda \in \{0\%, 10\%\}$ because less quantity must be handled by plants, as at most 20% of the demand requirements can be outsourced. Additionally, when 10% of the plants' production is shipped directly to customer zones, less facilities (warehouses) must be opened. In Figure 5-6 it is possible to analyse the average time gaps of LRD test instances for $|\mathrm{T}| = 4$. For $\beta = \{10\%,\ 20\%\}$ and $\lambda \in \{0\%,\ 10\%\}$ with exception of $\beta = \lambda = 10\%$,

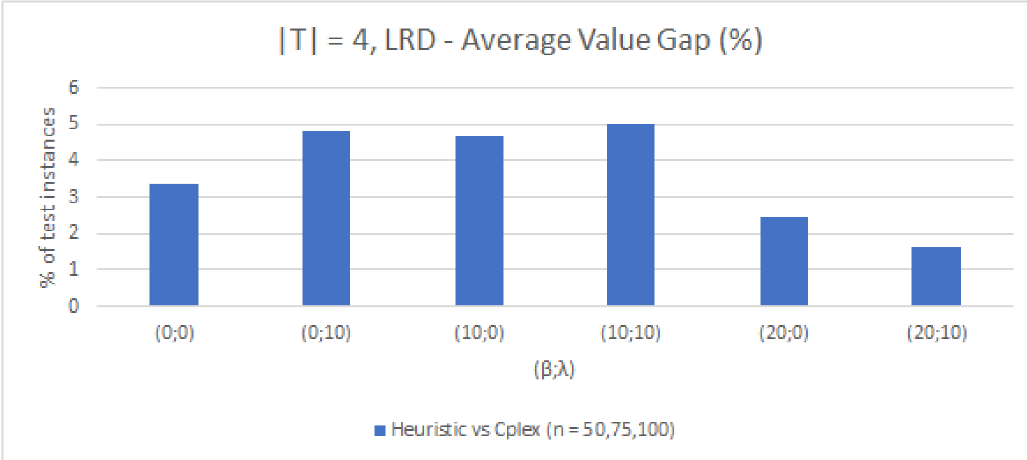

Figure 5-5: LRD Model - Value performance, $|T| = 4$

Cplex is able to obtain the optimal solution faster than the heuristic. In two cases ($\beta = \{10\%,\ 20\%\}$), less quantity must be handled by plants and consequently, the number of required plants is smaller. The same logic is valid when $\lambda = 10\%$, because this implies that direct shipments to customer zones are possible. In the LRD test instances, Cplex is able to take advantage of existing facilities to obtain less expensive solutions. Hence, better solutions are generated by Cplex.

Next, Figures 5-7 to 5-10 are representative of the heuristic performance for |T| = 6 and both problem classes. It is possible to observe a significant difference of the heuristic performance compared to the instances with |T| = 4.

In Figure 5-7, the average value gaps are slightly higher compared to |T| = 4. For |T| = 6 and the LND test instances, the heuristic must select more facilities to supply demand, and so more plants and warehouses must be opened. The heuristic is not sensitive in order to minimize the gap between the total installed capacity in time period t and the required capacity in the same period. Therefore, the heuristic is not

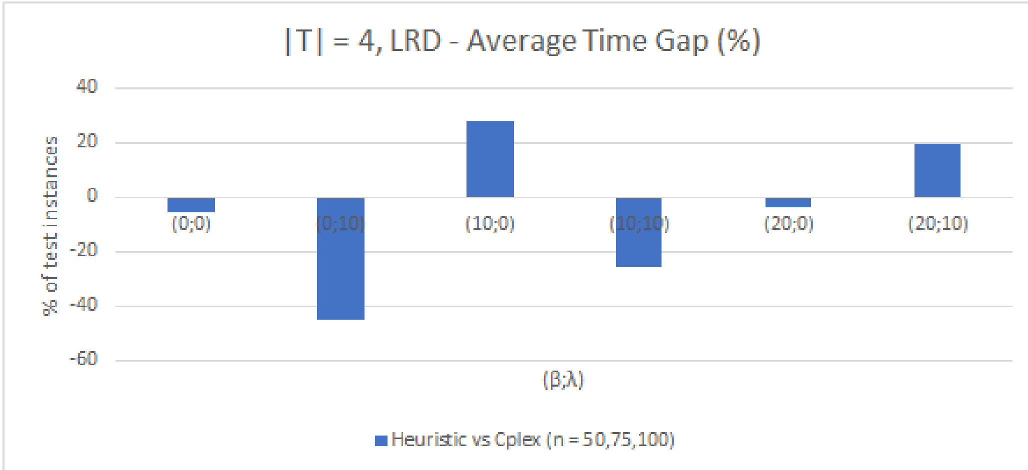

Figure 5-6: LRD Model - CPU Time performance, $|T| = 4$

able to select appropriate facilities and this is once again the cause of these positive gaps.

Regarding the CPU time performance for |T| = 6, there is a significant improvement

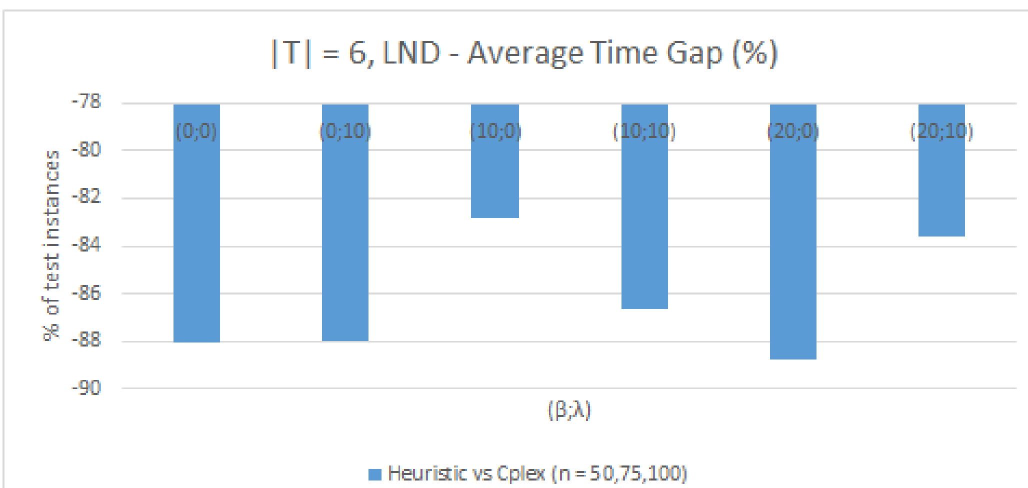

Figure 5-7: LND Model - Value performance, $|T| = 6$

of the heuristic performance. In the "worst" case the heuristic was able to find a solution in 82.83% of the instances (β = 10% and λ = 0%) faster than Cplex and at the "better" this value increases to 88.73% (β = 20% and λ = 0%). With |T|

= 6, more data must be analysed and Cplex spends more time to identify feasible solutions that are on average 4.63% (β = 10% and λ = 0%) and 5.91% (β = 20% and λ = 0%) cheaper than the heuristic solutions. Cplex takes more time to identify a feasible solution. However, it is able to do a better selection of the facilities that must operate, thus minimizing the total cost. In contrast, the heuristic is faster at obtaining a feasible solution but the selection of the facilities that must be opened is not as good as with Cplex. Hence, the heuristic solutions are more expensive.

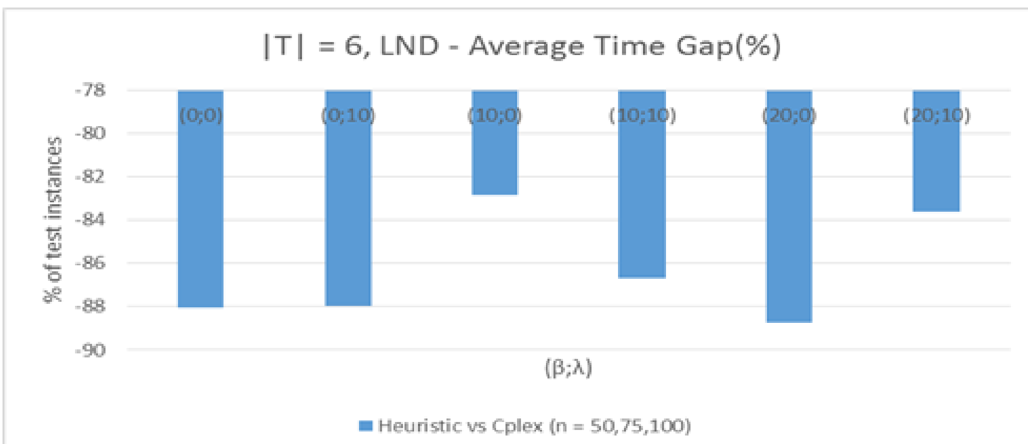

Figure 5-8: LND Model - CPU Time performance, $|T| = 6$

In Figures 5-9 and 5-10, the performance of the LRD class is presented for |T| = 6. Again, there is a significant difference of the heuristic behaviour compared to |T| = 4. The average value gaps are higher, with a minimum gap of 3.74% (β = 10% and λ = 0%) and a maximum gap of 5.73% (β = 20% and λ = 10%).

In contrast to |T| = 4, with |T| = 6 the heuristic is able to perform better than Cplex with respect to the average time gaps. In the "worst" case, the heuristic was 20.78% (β = 10% and λ = 0%) faster than Cplex to find a feasible solution that was on

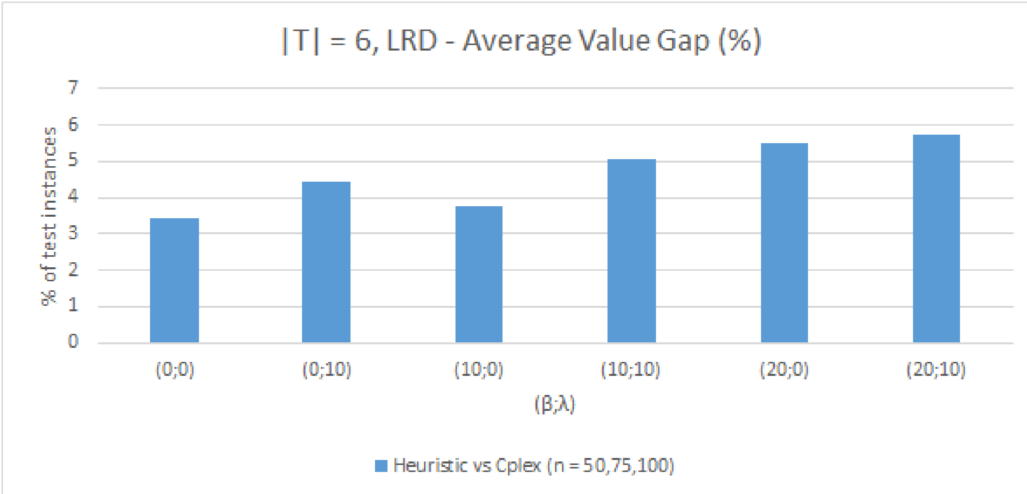

Figure 5-9: LRD Model - Value performance, $|T| = 6$

average 3.74% ($\beta = 10\%$ and $\lambda = 0\%$) more expensive. In the "better" performance case, the heuristic identifies a feasible solution 63.27% ($\beta = 0\%$ and $\lambda = 10\%$) faster than Cplex with a cost that is 4.44% higher (on average). The heuristic is able to

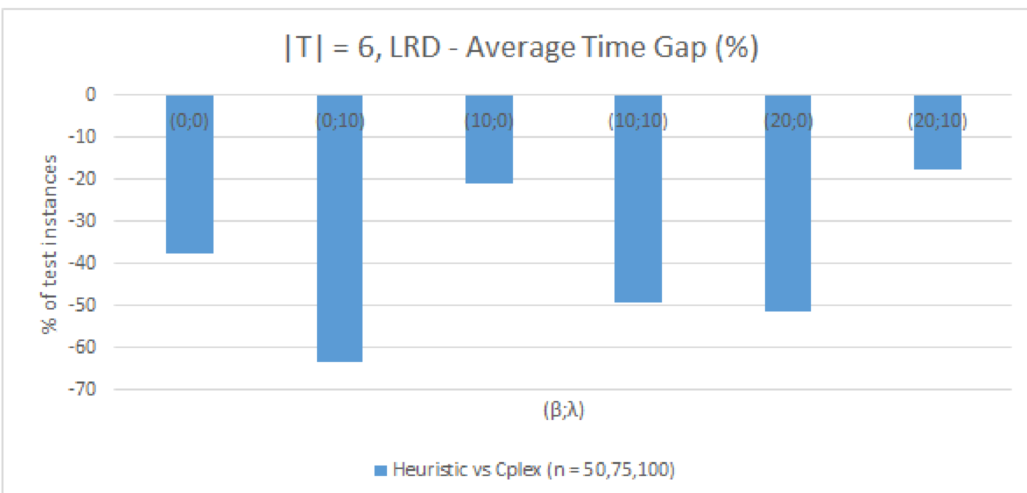

Figure 5-10: LRD Model - CPU Time performance, $|T| = 6$

find a feasible solution faster than Cplex. However, this comes at a cost with respect to the quality of the solution obtained, which is lower than the one identified by

Cplex. When a capacity size must be installed, the heuristic can determine the most appropriate capacity level to select.

However, the same does not happen with respect to the selection of facilities. When a facility must be selected, it is randomly chosen and therefore a capacity size is selected. But, each facility has different characteristics, and when more capacity is required, the heuristic is not able to decide if it is better to expand capacity or if it is necessary to open a new facility. These decisions are taken randomly and as seen with the performance analysis, they influence the quality of a solution.

5.2.1 Suggestions for improvements

Positive aspects of the heuristic are the CPU time performance and the capability to produce homogeneous solutions for different lengths of the planning horizon. However, it is necessary to improve the quality of the solutions. Therefore, in this section possible inefficiencies in the current version of the heuristic are discussed.

The following aspects were identified that could possibly increase the heuristic efficiency:

- Similar probability of selecting facilities in LRD problems. Due to the great number of new facilities in relation to existing facilities, the probability of selecting an existing facility is smaller than the probability of selecting a new facility. As a consequence, it is more difficult to take advantage of existing facilities to perform a capacity extension.

- Optimize the selection of facilities in order to select the "most feasible" facility.

The "most feasible" facility refers to the choice of a facility that minimizes the gap between the total capacity available in time period t and the total demand requirements in the same period. At the same time, the global capacity limit of the facility cannot be exceeded.

- Allow the exchange to different facility types. Currently, any actions must be taken within the same group of facilities (example: if local search is applied to a new facility, the process will be conducted with new facilities).

Chapter 6

Conclusion

The amount of available information is each day bigger and bigger, what requires a great capacity to analyze data in order to better understand clients' needs and keep ahead of competitors. In order to become competitive, a company must be capable of making decisions ahead of its competitors. Thus, time is a valuable asset in decision making. Therefore, the MA has a clear advantage in relation to Cplex, since more solutions can be provided to a decision maker within a given time frame. For example, for an LND problem ($|T| \in \{4, 6\}$), the heuristic is able to provide a solution 77.30% more faster that is on average 4.76% more expensive than Cplex solution. For LRD instances, the heuristic is able to provide 22.57% faster solutions with a total cost that is on average 4.13% higher than that obtained by Cplex.

If a company is not competitive or if it is losing competiveness, the cost associated to this position is not quantified, however, it can have an important impact on the future of the company. As a result, despite of the good performance of Cplex to find optimal solutions or solutions that are less expensive than the heuristic, Cplex reached the time limit of 8h (CPU) without finding the optimal solution in 44.44% of the LND instances ($|T| \in \{4, 6\}$). For LRD instances, Cplex performed is better. However, in 28.89% of these instances Cplex did not obtain an optimal solution after 8h (CPU).

Taking benefit of the heuristic time performance, some changes can be applied to increase the quality of heuristic solutions without having a significant impact on the CPU time performance.

The proposed MA procedure is not very sensitive, all binary decisions are taken randomly (except the selection of capacity levels). If the process becomes more aware of the type of facility decision, the heuristic will be able to generate better individuals in four or five generations. Turning the heuristic more sensitive means that the selection of a facility would not depend on randomness, but it could identify the points in the process that influence the heuristic performance and turn them less dependent on randomness.

Some possibilities for improving the MA were proposed. For instance, the selection between existing and new facilities should be more efficient, to decide in each period what is the most advantageous to a company, choose between expanding the capacity of an existing facility or opening a new facility. If there is available capacity at existing facilities then it should be used because less expensive solutions will be generated. In order to keep diversification in the population, the heuristic may select a set of best options and afterwards select one of them randomly.

Facilities have different characteristics, as a consequence if the heuristic is able to select a facility that is able to supply demand but it is also the cheapest option available, better solutions will be obtained. Once again, diversification can be maintained, if a set of facilities is selected and one is randomly chosen.

Bibliography

[1] A. Antunes and D. Peeters. On solving complex multi-period location models using simulated annealing. *European Journal of Operational Research*, 130:190–201, 2001.

[2] R. Babazadeh, J. Razmi, and R. Ghodsi. Supply chain network design problem for a new market opportunity in an agile manufacturing system. *Journal of Industrial Engineering International*, 8:1–8, 2012.

[3] H. Badri, M. Bashiri, and T.H. Hejazi. Integrated strategic and tactical planning in a supply chain network design with a heuristic solution method. *Computers & Operations Research*, 40:1143–1154, 2013.

[4] R.H. Ballou. Unresolved issues in supply chain network design. *Information Systems Frontiers*, 3:417–426, 2001.

[5] M. Bashiri and H. Badri. A dynamic model for expansion planning of multi echelon multi commodity supply chain. *International Journal of Engineering and Technology*, 2:85–93, 2010.

[6] M. Bashiri, H. Badri, and J. Talebi. A new approach to tactical and strategic planning in production-distribution networks. *Applied Mathematical Modelling*, 36:1703–1717, 2012.

[7] J.F. Cordeau, F. Pasin, and M.M. Solomom. An integrated model for logistics network design. *Annals of Operations Research*, 144:59–82, 2006.

[8] M.J. Cortinhal, M.J. Lopes, and M.T. Melo. Redesigning a three-echelon logistics network over multiple time periods with transportation mode selection and outsourcing opportunities. Technical reports on Logistics of the Saarland Business School 7, Saarland University of Applied Sciences, Available at `http://www.htwsaar.de/wiwi/ForschungundWissenstransfer/publikationen/dateien/arbeitspapierelogistiknr7.pdf`, 2014.

[9] J. Dias, M.E. Captivo, and J. Clímaco. A memetic algorithm for multi-objective dynamic location problems. *Journal of Global Optimization*, 42:222–252, 2008.

[10] S. Elhedhli and J.L. Goffin. Efficient production-distribution system design. *Management Science*, 51:1151–1164, 2005.

[11] E. Eskigun, R. Uzsoy, P.V. Preckel, G. Beaujon, S. Krishnan, and J.D. Tew. Outbound supply chain network design with mode selection, lead times and capacitated vehicle distribution centers. *European Journal of Operational Research*, 165:182–206, 2005.

[12] A.M. Geoffrion and G.W. Graves. Multicommodity distribution system design by benders decomposition. *Management Science*, 20:822–844, 1974.

[13] E. Gourdin and O. Klopfenstein. Multi-period capacitated location with modular equipments. *Computers & Operations Research*, 35:661–682, 2008.

[14] T.P. Harrison. Principles for the strategic design of supply chains. In T.P. Harrison, H.L. Lee, and J.J. Neale, editors, *The practice of Supply Chain Management: Where theory and application converge*, chapter 1, pages 3–12. Springer, New York, 2004.

[15] Y. Hinojosa, J. Kalcsics, S. Nickel, J. Puerto, and S. Velten. Dynamic supply chain with inventory. *Computers & Operations Research*, 35:371–391, 2008.

[16] M.T. Melo, S. Nickel, and F. Saldanha da Gama. Dynamic multi-commodity capacitated facility location: a mathematical modeling framework for strategic supply chain planning. *Computers & Operations Research*, 33:181–208, 2006.

[17] M.T. Melo, S. Nickel, and F. Saldanha da Gama. Facility location and supply chain management: a review. *European Journal of Operational Research*, 196:401–412, 2009.

[18] M.T. Melo, S. Nickel, and F. Saldanha da Gama. A tabu search heuristic for redesigning a multi-echelon supply chain network over a planning period. *International Journal of Production Economics*, 136:218–230, 2012.

[19] M.T. Melo, S. Nickel, and F. Saldanha da Gama. An efficient heuristic approach for a multi-period logistics network redesign problem. *TOP*, 22:80–108, 2014.

[20] Z. Michalewicz and D. B. Fogel. Traditional methods - part I. In *How to Solve It: Modern Heuristics*, chapter 3, pages 55–82. 2$^{\text{nd}}$ Ed. Springer, New York, 2004.

[21] H. Mühlenbein. Genetic algorithms. In E. Aarts and J. K. Lenstra, editors, *Local search in combinatorial optimization*, chapter 6, pages 137–172. Wiley-Interscience, New York, 2003.

[22] J. Mula, D. Peidro, M. Díaz-Madronero, and E. Vicens. Mathematical programming models for supply chain production and transport planning. *European Journal of Operational Research*, 204:377–390, 2010.

[23] E. Olivares-Benitez, J.L. González-Velarde, and R.Z. Ríos-Mercado. A supply chain design problem with facility location and bi-objective transportation choices. *TOP*, 20:729–753, 2012.

[24] H. Sadjady and H. Davoudpour. Two-echelon, multi-commodity supply chain network design with mode selection, lead times and inventory costs. *Computers & Operations Research*, 39:1345–1354, 2012.

[25] P.N. Thanh, N. Bostel, and O. Péton. A dynamic model for facility location in the design of complex supply chains. *International Journal of Production Economics*, 113:678–693, 2008.

[26] P.N. Thanh, N. Bostel, and O. Péton. A DC programming heuristic applied to the logistics network design problem. *International Journal of Production Economics*, 135:94–105, 2012.

[27] P.N. Thanh, O. Péton, and N. Bostel. A linear relaxation-based heuristic approach for logistics network design. *Computers & Industrial Engineering*, 59:964–975, 2010.