POLITECNICO DI TORINO

SCUOLA DI DOTTORATO

Dottorato in Ingegneria Informatica e Sistemi – XX ciclo

Tesi di Dottorato

in

Operational Research

# Models and methods for the City Logistics

The Two-Echelon Capacitated Vehicle Routing Problem

**Jesus Gonzalez Feliu**

| Tutore | Coordinatore del corso di dottorato |
|---|---|
| prof. Roberto Tadei | prof. Pietro Laface |

Marzo 2008

# Contents

# Aknowledgements

In Spain there is a common sentence that says something like this ”it is common in well-born people to know how to say thanks in a good way and sincerely” (*Es de bien nacido ser agradecido*. The final result of this PhD. Thesis is also in a way because of the technical, emotional, moral or alternative help of all the people I'll name below. I will certainly forget some people (I'm doing it quickly and my memory is not as precise as a swiss clock) but be sure that I remember well all the people that helped me in different ways during these three years.

First of all, my parents, Jesús and Natalia, who know me well and have always be on my side. They helped me to take the difficult decision of going abroad to study when I was 18, and since this moment (and also before) they were supporting me during all the time. Even in the hardest moments their wisdom was guiding me. Grcies pare! *Gràcies mare! Ha costat però com sempre m'haveu dit al final he acabat la tesis. Gràcies de tot cor!*

Another person who lived my humor changes and the daily problems I was dealing with is my love, Ruby. She was supporting me, even if she would like me to do something different, and this is really important for me. Thanks a lot, my love. *Je t'aime, mon bijou! Gracias, amapolita mía, eres maravilosa!*

Also my grandparents have been present, even if only my grandmother Natalia is still alive. Seh asks me every time I phone her when I am going in Spain. She misses me a lot and I miss also her so much. My grandmother Rosario died during the period I was preparing to go to Canada, and my grandfather Manuel, her husband, few months before making the access test to enter the doctorate studies. My grandfather Salvador is more than ten years following me even if he is not present, and he is like a father for me.

Continuing with "family", I have to mention Ruby's mother, Belinda, and Ruby's younger brothers, Blu and Dany, who are like a family for me. Also Lorenzo and Carla (Badellino family), who are more than the owners of the apartment I passed all these years. They are like my Italian parents and they made a lot for us (the people that lives and lived in this apartment). Also my flat mates: Vishal, who was at the place all the time, Valentin, my first flat mate alongside Vishal and who is like a big brother for me, even if we don't communicate so frequently, Manel, Arnaud, Vinod Pandey, Ji and Branislav. These years at the apartment were so nice, and I liked to live with all of you.

Professionally, I would'nt follow this doctorate course without the help and motivation of Prof. Bruno Dalla Chiara (DITIC - Politecnico di Torino). He introduced me to my professor, Prof. Roberto Tadei (DAUIN - Politecnico di Torino), who helped me and gave me the possibility to make this thesis. Also Guido Perboli, Marco Ghirardi, Simone Palamara, Gigi Arcaro and Luciano Governali, people who were in the OR group of DAUIN - Politecnico di Torino at the time, were very gentle with me. Other people I aknowledge are Prof. Federico della Croce (Politecnico di Torino), Andrea Grosso (Università degli studi di Bologna) and Giuliana Carrello (Politecnico di Milano), whose advises at the PhD. Thesis dissertation were very interesting and I tried to apply the most I could in the time that was not a lot. Special thanks to Francesco Paolo Deflorio (DITIC - Politecnico di Torino) who is more than a colleague, he is a friend to me.

Prof. Daniele Vigo (Università degli studi di Bologna) helped me at the beginning to choose a good and interesting subject, and the 2E-CVRP is also his creation. He gave me very important comments to improve the models and his ideas made the research advance well. Prof. Teodor Gabriel Crainic (CIRRELT and UCAM) invited me at CIRRELT - Universit de Montral and helped me so much. It was not the best period of my life, and I found in him understanding and help. Thanks a lot! Another professors I want to aknowledge is prof. Nicoletta Ricciardi (Università La Sapienza di Roma), prof. Alfredo Marin (Universidad de Murcia), prof. Paolo Toth (Università degli studi di Bologna) and prof. Martine Labbé (Université libre de Bruxelles).

The people from the now disappeared true and old labdar were my first reseach colleagues. Ernesto is now a friend to me, even if the is too busy, but it's normal. He's a great worker, and he has to take care about Alicia and the little Camilo. Michelangelo, Max, Ashutosh, Ashoka, Prassana (sorry for the names, guys), Andy, Luisa, Alessandro, Attilio ... did I forget someone? Also Marta, Jorge and Javier, who came only for some months, and Andrea, who I saw only for some weeks. Then, the ACLAB people: Freddy, Lorenzo, Ilario, Filiberto, Christian, Ottavia, Jose and of course a special mention to Salva, who came in September and with whom I can speak for hours and hours.

I also followed some Master students: Simone Amico, who did a huge job, which will not be considered by "pure" operational research people but which I liked a lot, Simone Favat, who I could'nt follow so well because I went to Canada, Emanuele, Simona and Lorraine (the two girls, during my research period in Canada, but I followed them only in the preliminary steps because of the "method approach" which was different).

The people I met in Canada, who were a lot and I had good moments: Benoit, Simon, Endong (my three office mates), Ehud (my meal mate), Claudio, Nicole, Tolga, Gunes, the French Benoit, Roberta, Marie-Eve, Nicolas, Bea and his boyfriend, Manuel, Francesco and plenty of nice guys and girls (a least 8) I forgot their names. A special mention to Solange and Nicolas, my Canadian family, and to the BIG Rusty Priske I could'nt meet although he lives in Ottawa. No, I did not forget my great friends Massimo (my Sard brother) and Yolanda (and also her family: Alina, Mirko and Tania). Another special mention to the MYTH prof. Mike A. Florian I could see in person at CIRRELT.

I don't forget my Spanish mates, but I think some will not appear here, but it's only for a temporary lack of memory (sorry if this person is you, I didn't make it conciously): Carlos Peris, Sam, María, Toni, Ana, Pipe, Inma, Deb, Javi, Arantxa, Manolo Carda, Diego (Jefe), Rosa Oya, Víctor Lázaro, Llúcia, Marta, Mireia, Silvia, Lucero, A-chanita, Miki, Héctor, Isa, Carlos (Bisbal), Jordi, Blizzard, Iwa, and a big etc of names (sorry, I know the list is much bigger, don't be angry, please).

Last, but not least, the people from VITESSE s.r.l. who game me time and patience in the last weeks of the PhD. Thesis writting: Pauline, Els, Lorenzo, Christian, Sandro, Sophie, Raffaella, Patrizia, Gigi and Gianluca.

A special mention to Pecas, my "little" dog-friend, who died in spring 2008 (so few before presenting the PhD. Thesis).

And thanks to all the people that I didn't cite explicitly but that helped me a lot (and sorry for don't remembering your names in this precise moment). Thank you also, dear reader, for taking a while to consult this PhD. Thesis.

**Jesús González Feliu**

# Introduction

The transportation of goods constitutes one of the main activities that influences economy and society, as it assures a vital link between suppliers and customers. Since 1970 transport activity has more than doubled in the European Union (EU). Between 1970 and 2003, freight transport has raised about +185%, whereas passenger transport also presents an important increasing trend (+145%). Today road transport is dominant over other modes of transport, with a market share of 45% for the transportation of goods and 87% for passenger transport, as shown in the European Union's *Energy and Transportation Report 2000-2004* ( (44) ). Moreover, the transportation sector represents about 10% of the European Union GDP , and it is also a major source of employment, with more than 10 millon workers in the EU (2003). However, transportation costs are also important. In the European Commission Energy & Transportation report 2000-2004, the transportation costs were estimated at around  210 billions. The EU Commission prospects that between now and 2020 the demand due to the transportation of goods in the EU member states will increase by 70%, and by 95% in the ten new member states and that this will raise the transportation costs by EUR 80 billions per year.

Another important negative aspect of transportation is its environmental and social costs. In 1998 the transportation of goods was responsible for 28% of gas emissions. Between 2003 and 2010 this share is likely to increase to 50%. Environmental costs can be represented in function of the quantities of polluting substances that each vehicle produces, so estimating the different traffic flows an compositions, these costs can be estimated. Social costs quantification is however more complex and sometimes these aspects are not directly transformed into a numeric value in transport costs evaluation.

These problems are amplified in cities. In fact, urban activities related to freight transportation are extremely important, for most of the activities thet take place in a city require transportation of commodities. But freight transportation can be disturbing in some urban areas. The infrastructures used by freight transportation vehicles (streets, roadways and parking areas) are usually the same as those used by private and public vehicles for the transportation of people. In the last years, both passengers and freight traffic flows have increased, and in some cities this trend has traduced into situations of congestion and environmental nuisances, like air pollution, noise, and other diseases related to this problem, for example increase of waiting times or difficulty in finding a parking place. Freight transportation vehicles are usually big (usually more than 3.5 t.), their flows are not decreasing but growing, and more strict quality policies are required for some categories of goods.

In the last decades, urban freight distribution have been studied in order to deal with these aspects, and a new discipline, the city logistics, has been developed to decrease traffic, pollution, noise and other diseases related to freight transport in urban areas. In Europe, some real applications have been developed, most of them proposing an alternative transportation system which realizes a urban freight distribution service, although in some cases combined with a specific normative. These systems are usually based on one or more Urban Freight Distribution Centers (UDC), which are intermediary platforms where freight, arriving from different locations outside the city, is organized into smaller and less polluting vehicles that will be able to satisfy each request of goods in some urban areas. Most of these experiences have been planned to deal with a specific problem, although they present many similarities between them. However, no standardization has been made, and the vehicle trips and crew scheduling organization is made following traditional techniques. Even when optimization procedures are used, the entire transportation system, which involves more then one stage (called echelons) is not considered, but each echelon is optimized separately, losing a system view.

Moreover, in other transportation applications, multi-echelon distribution policies are applied. However, literature in this field refers to supply chain and inventory problems, without studying the global costs of the transportation system to optimize. In fact, in real applications, no global system view is used in cost optimization, except in some very specific distribution examples which will be furtherly presented.

The main contribution of this work is the introduction of a family or route optimization problems which deal with multi-echelon transportation systems, the *Multi-Echelon Vehicle Routing Problems.* were the routing and the freight management are explicitly considered at the different levels. Moreover, we introduce and study one of the simplest types of Multi-Echelon Vehicle Routing Problem, the *Two-Echelon Capacitated Vehicle Routing Problem* (2E-CVRP). In 2E-CVRP, the freight delivery from the depot to the customers is managed by shipping the freight through intermediate depots. Thus, the transportation network is decomposed into two levels: the 1st level connecting the depot to intermediate depots and the 2nd one connecting the intermediate depots to the customers. The objective is the minimization of the total transportation cost of the vehicles involved at both levels. Constraints on the maximum capacity of the vehicles and the intermediate depots are considered, while the timing of the deliveries is ignored.

This PhD. Thesis is organized as follows. In chapter 1, the main city logistics concepts and applications will be presented, and, from these experiences, a group of guidelines for the planning of urban freight distribution systems are described.

The second chapter will be dedicated to general freight transport strategies, focusing on multi-echelon distribution. The generalities of multi-echelon transport, and a brief survey of scientifical studies on multi-echelon inventory systems and supply chain are also presented. Then, the main transportation applications which follow these strategies will serve as examples to define the organization of multi-echelon distribution systems. Because the main problematic arises on transportation cost optimization, a survey of transportation costs optimization is presented in chapter 3, presenting the vehicle routing problem and its main variants, as well as the few works that have been realized in multi-echelon distribution systems decision problems.

In chapter 4 we give a general description of Multi-Echelon Vehicle Routing Problems, which is the generalization to N-echelon of the well-known Vehicle Routing Problem family. We will then focus on two-echelon problems, and in chapter 5 we give a general description of Multi-Echelon Vehicle Routing Problems, presenting the main definitions and the most common variants that can be considered. Section 5.2 is dedicated to the introduction of the decision problem (the 2E-CVRP) and present two mathematical models to represent the route optimization problems in a basic two-echelon distribution system. The first model derives from multi-commodity network design and will be called flow-model, whereas the second one derives from the classical TSP and VRP formulations, and will be noted as TSP-like model.

In order to evaluate the models, four sets of test instances for the 2E-CVRP are introduced in chapter 6, where the models are tested, focusing on the flow model, for which computational results are discussed, showing the behavior and limits of these models.

Finally, guidelines for solving medium instances of 2E-CVRP will be proposed in chapter 7. In this chapter, set covering and set partitioning route formulations will be presented. The first method will be described, presenting its limitations after the first calculation tests, which are related to the importance of the connexion between the two levels and the difficulty to generate a set of compatible columns. For this reason, we also present the guidelines to realize a method that do not separate both levels in the column generation step, but generates a set of first and second level columns which are connected.

# Chapter 1

# An overview on City Logistics

Traditionally, freight transportation planning and decisions have been realized by the operating companies and usually not by the public authorities. In the last years, public authorities have begun to develop some ideas to manage and contain the levels of congestion, air pollution and noise, but most of them are political or normative decisions made to solve emergency situations. Some of the most common measures taken by the authorities in different countries are restrictive policies such as regulation on parking, street access, hours of operations, etc.

In several countries, surveys and data collection activities have been undertaken, and some studies give elements of freight transportation analysis and organization for urban areas (43; 56; 49; 82; 112; 75; 11; 61) . These efforts are aimed at a better understanding and quantifying these phenomena and represent a first step in the development of this young discipline. Several European cities have developed transportation policies for urban freight distribution and some city logistics services have been started; nowadays, some of them are successful and operative (e.g., Monaco, Zurich, as well as in a number of cities in Germany, the Netherlands, France, and Italy). These experiences are usually realized by local authorities and companies, and only in a few cases coordinating efforts between different experiences have been made, resulting into normative and financial aid policies. However, no national or international guidelines and policies for projecting city logistics services have been elaborated, even if many similarities can be found between the different projects and experiences already started in EU countries.

The concepts of City Logistics (92; 64; 101) are being developed to answer to this need. The main objectives of the City Logistics studies are the following:

- Reduce congestion and increase mobility of freight transportation services in urban areas.
- Reduce pollution and noise; Contribute reaching the Kyoto targets; Improve living conditions of city inhabitants.
- Avoid penalizing the city center commercial activities such as not to "empty" them.

## 1.1 City Logistics studies and experiences

Before the 80ś, the urban traffic due to freight transportation did not constitute a problem, and it was not managed by public administrations. The only type of interventions from the municipalities were provisional measures to deal with emergencies. Most of them, which had some common ideas, were restrictive measures, like pricing strategies for circulation in the city centers (which have become a permanent measure in some big cities like London), regulation of parking areas and limited traffic zones. In the 80ś, some route optimization methods and the introduction of incentive measures have become an alternative, an also a complement, to the classical strategies. With urban traffic increasing, and the raise of congestion not only in big but also in medium cities, some public administrations have affronted the problematic of urban freight distribution, that was managed traditionally only by the transportation carriers. In the 90ś and the beginning of the $XXI^th$ century, with the contribution of some public administrations and in some cases, with European Union financing funds, some cities started to study how to organize urban freight distribution in order to decrease traffic and pollution derived from this transportation sector.

One of the first studies that try to define some common policies for urban freight transportation management has been developed by the COST network and the European Commission (42) . *COST Action 321: Urban Goods* Transport is a European study made by 12 countries (Denmark, France, Finland, Germany, Greece, Italy, The Netherlands, Slovenia, Spain, Sweden, Switzerland, United Kingdom). The aim of this action was the study of the design and operation of innovative measures to improve the environmental performance of freight transport in urban areas, analyzing how the air pollution, noise and energy consumption were to be reduced by optimizing the use of freight transportation vehicles in urban areas through the application of modern logistical devices and appropriate administrative measures.

Different cities involved on the COST 321 action have proposed and tested pilot studies whose results have been used to widen the knowledge of the effects and acceptability of the measures, to prepare the way for the introduction of appropriate measures in Europe as a whole, as well as to increase public awareness of the problems caused by traffic of vehicles used in urban freight distribution and the need for international co-operation in this field. Administrative measures and logistical methods employed in the operation of truck fleets have been examined to see which contributions are the best to reduce the environmental impact. Measures and methods have been examined for economic efficiency and environmental benefits in demonstration projects, taking into account the direct and indirect effects on traffic flow and the location of commercial activities. Three main areas of possible action have been determined. These possible intervention areas are transport efficiency, infrastructure and technology. Related to them, some recommendations have been formulated:

- It is important to note the main differences between cities and to recognize the related contexts. There is not an optimal common policy to improve urban freight transport issues, but rather a set of common goals and a global need for additional and more comprehensive information and documentation.
- Individual measures 'per se' are not able to address the complexity of the problematic of the urban freight distribution and of the many interactions between sector activities and policies.
- Innovative intermodal interfaces could play a major role in the overall improvement of the sector, through a rationalization of demand, and the reduction of negative impacts.
- The technical improvement of vehicles and their fuel is also a promising direction of improvement.

In line with the increasing awareness of urban environmental problems, it seemed then necessary to define common guidelines and measures apt to reduce the negative impacts of freight transportation within cities. Actually, one of the highest difficulties became the search of a consensus between the different actors for establishing such guidelines. A final rapport enumerating the different experiences and a group of measures was published by the European Commission (43) , but many classifications and groups of similar measures were proposed. Instead of finding a common pattern, the action provided a detailed list of possible cases and actions.

### 1.1.1 The European experiences

In Europe we can find many medium cities that started city logistics projects (43; 84) . Most of them have organizing and infrastructural similarities, and in some cases, similar strategies were followed at national levels in some countries. National and regional initiatives are however few, and they arise on normative or financial aspects more than on service organization or transport planning. We will briefly present some examples of European cities with operative city logistics service and show these common points.

#### ***Germany***

Germany was one of the first countries that developed City Logistics systems, which started in many cases in the early 90ś. Most of these cities are small or medium (less than 1.000.000 inhabitants), although we can find an operative system in Berlin. Their common point is the creation of a Urban Distribution Center (UDC), which is built and maintained with the help of public subventions (in some cases, the service is totally financed with public funds). The UDC is mainly located in a peripherical area or in the proximity of a main highway. A urban freight distribution service is created, and brings the freight from the UDC to its final destination. Another common characteristic is that in the initial phase, the participation to the project by transportation carriers was voluntary, and in many cases the municipalities have not created a specific normative to incentive the usage of such urban freight distribution services. The common points of the German projects are:

- The need of coordinating and optimizing the vehicle load. The average vehicle load with the City Logistics services is 70-80

- The high degree of privatization and the voluntary collaboration between private enterprises.

- The usage of small vehicles in urban areas, reducing the number of vehicles (55% in average) and also the transportation costs (20-30%).

In Kassel, 10 transportation carriers have agreed to collaborate and have subcontracted the urban transport of goods to one logistics operator. Every day this operator receives the information relative to the freight it have to be delivered in the urban area and visit the origin locations of the freight. Then, this freight is transported to a UDC, where itś reorganized into smaller vehicles (7,5 T.) in function to the destination. From the UDC, these vehicles travel to destination of each freight request. Another example can be found in Bremen, which service has been financed with public funds at 40% (the remaining 60% has been given by private investors). In this case, the transportation carriers bring the freight to the UDC, where it is organized into smaller vehicles. The transportation from the UDC to freight destination is realized by only one operator, the same that manages the UDC. The adhesion to the service is voluntary, and there are currently 12 transportation carriers that collaborate, 9 of them regularly. A similar approach have been realized in the city of Essen, which uses the same organization as the city of Bremen. Special attention can be given to Freiburg, where the transportation carriers have been divided in 4 groups. Each group combines their expeditions to decrease costs and pollution. The particularity of this system is that each group works separately, and no UDC has been projected.

In the city of Berlin we can find one of the first City Logistics services, which was started in 1993 with one UDC and currently coordinates the expedition of 10 carriers in 2 UDC. There are two logistic operators, whose role is to collect all the goods that have to be delivered in the urban area and bring them to one UDC, where itś aggregated into small vehicles and distributed to destination.

***The Netherlands***

In the Netherlands, the creation of UDC with voluntary participation is not enough to solve the traffic problem. Without regularization policies, the efficiency of these systems is highly limited. We can observe one of the first normative policies for freight traffic regulation. This is based on two concepts already presented: restrictive and incentive policies. In a first time, a restrictive policy obliges transportation carriers to enter in determinate zones at defined hours of the day. Some carriers are habilitated for distributing in other periods of the day (incentive policy) but for this they have to obtain a "Urban Distribution Permission" which is delivered only to carriers that respect a number of criteria. We can observe two main examples: Amsterdam and Utrecht.

In Amsterdam, a first UDC was created in 1996, and a first service was adopted by a group of private companies alongside the Commerce Chamber and the municipality. Nowadays, the service has evolved. In a first phase, restrictive measures were applied to urban freight traffic, in terms of dimension, weight and pollution emissions of the freight distribution vehicles (creation of Limited Traffic Zones (LTZ), which cannot be entered by freight transportation vehicles without permission). The number of permissions is also limited. In a second phase, an incentive policy strategy was developed to regroup transportation carriers for aggregating the deliveries. In a third time, 9 CDU have been identified and developed and a transportation network is now operating in the city of Amsterdam.

A similar system is operating in Utrecht, but it is not working well because of the low number of carriers that have obtained permissions (only 2), which causes most of the carriers to use the traditional freight transportation planning.

***France***

In France, the government, being conscious of the problem the freight distribution in urban areas can become, has created a division of the Ministry of Transportation for Urban Freight Distribution (75) . In the 90ś, this division has promoted different studies for determining measures and policies to regulate the urban freight distribution. Some of these studies have evolved into projects, and nowadays some cities have operating city logistics services.

In 1998, an innovative project was presented in La Rochelle by the Chamber of Commerce, the municipality, the Association of Commercial and Artisan Activities and the transportation carries, with public municipal funds. A platform (UDC) has been developed in the peripheries of the city, and a private company (which election is periodically made by public competition) is managing the service. The service is made by nine (electric vehicles, two of them having a controlled temperature system, in a similar way to the projects already described, and a transport planning software consent to optimize the freight distribution. Some restrictive measures for freight distribution vehicles with weight superior to 3,5 T. have been created to incentive the usage of the service.

The case of Paris is quite different. In a fist time, no common distribution service had been created, but, instead, some restrictive policies (for some highways in the most populated hours of the days, but LTZ or specific time periods for freight distribution have not been created) had been combining to some incentive policies, like reserved parking areas . In January 2007, the city of Paris adopted a new normative for the freight transportation and distribution. It consists on a group of restrictive measures for the biggest and most polluting vehicles, and the regularization of the parking areas in order to facilitate the freight distribution (time limit for parking, exclusivity of some zones for specific vehicle categories, etc). In 2003, in order to serve some small streets, pedestrian areas, or very congested zones, a service of urban freight distribution made by small vehicles (non motorized tricycles) has been created, as an alternative, in competition with the traditional services. A company manages the service, and receives the different requests. Freight arrives to a logistic platform (which can be considered as small UDC), where the tricycles are loaded. This service is useful for small distances in the city center, and can be combined with the system of habilitated parking and maneuver areas for freight distribution. The service was a success and in 2005 it has expanded, increasing the number of vehicles and employees of the company to satisfy the increasing demand.

***Italy***

In Italy, we can find some city logistics systems projects. Nowadays, few of them are operating and are financially auto-sufficient. In other cities, similar projects were started, but only in medium cities are operative and fulfilling their objectives. In Rome, the dimension and the characteristics of the city are delaying the definition of a common freight transportation system for the city center areas, and in Milan, only a initiative from the Public Transportation company is operating, but currently this initiative does not have big repercussions. Other cities, like Naples and Turin, have not started such systems.

A national association was born in 2004 to deal with the urban distribution problems. The Italian City Logistics Association realized some events to compile and study the city logistics projects in Italy. In 2006, in its 2nd annual conference, the problem of finding national guidelines and policies was the main subject. After analyzing different experiences, it was considered necessary to establish general guidelines for projecting and planning city logistics services. The importance of an analogue figure to the existing mobility manager but specialized in freight transportation (the proposed name was logistics manager) for medium and big cities was evidenced. It was also determined that separate measures do not constitute a sustainable and competitive city logistics system, remarking that some measures already adopted by Italian cities are only provisional and will not reach in long term planning without combining them to other type of measures. These measures can be organized in four groups:

- Normative policies, which can be restrictive or incentive.
- Information and communication tools.
- Infrastructural, technological or civil engineering contributions.
- Partnership between public and private enterprises.

Some of the experiences presented in Italian City Logistics events are projects which are not operative. We present three operative city logistics services and a regional initiative whose projects are in an explorative phase or where the first tests are taking place.
The oldest operating system for urban freight distribution in Italy can be found at Padova. The particularity is that the system is operative and auto-sufficient without strong restrictive polices. Instead, the municipality has proposed high incentive plicies, which makes the usage of the service advantageous for small and medium transportation carriers, those which are predominant in Italy, and then constitute a big percentage of freight traffic in urban areas. At Pavoda

there is a Logistic Center near the city, so the UDC has been realized inside this platform, which implies lower costs with respect to the project entirely new infrastructures. The service works as follows: the transportation carriers which want to use the services bring their freight to the UDC. There, goods are consolidated into electrical vehicles, which travel to the LTZ zone of the city. The service is working so well that it has been extended to other zones of the city.

Another successful case can be observed in Vicenza, where the service is called VELOCE. The system is similar to Cityporto, in operative terms, but in normative ones there are some differences. The LTZ accesses are very restrictive and from may 2005 only the city logistics services are allowed to enter in all the city zones. The transportation carriers must bring their freight to the UDC, controlled by a municipal service which sub-contracts the transportation service to a cooperative transportation company. In the UDC, goods are consolidated into GPL vehicles, which travel to the LTZ zone of the city. The service is auto-sufficient but some big transportation carriers do not agree with the highly restrictive policies.

In Milan, a urban freight distribution service, CITYPLUS, was born. This is a particular case, being a division of the city Public Transportation Company, ATM. They use two distribution centers (which are located at two ATM bus depots) and operate as a private enterprise in competition with the other transportation carriers. They have the advantage of using the advantages that ATM has (reserved highways and parking areas, permissions, etc.), but there is not a specific normative of the city to incentive or oblige other carriers to use this service.

In 2005-2006, the region of Emilia-Romagna started a project which aim is to propose guidelines and a normative on freight distribution in urban areas. This region proposed a first "program agreement for the air quality", in different zones and for each of them specific funds have been established to finance city logistics projects on cities with more than 50.000 inhabitants. These projects are in an explorative phase.

***Other experiences***

In other European cities we can find different or similar approaches in order to regulate the urban freight distribution. We can divide these examples into two groups: cities where only normative policies were applied and cities where a freight distribution system for city centers and/or other urban areas were developed. In the first case we can find cities like Munich (Germany) or Copenhagen (Denmark), where different restrictive policies were applied to consent the freight distribution only in some hours of the day and to reduce the number of commercial vehicles entering the city center. In Copenhagen, different certificates are needed to enter some zones of the city (freight transportation and private automobile traffic). In the second case we can find the example of Malaga (Spain), which is similar to La Rochelle experience. We will present other two cases: the freight distribution service in the city of Montecarlo (Monaco), which uses the UDC solutions, and the case of London (United Kingdom) which only applies normative measures to regulate the traffic in the city center.

Montecarlo (Principate of Monaco) was the first city in Europe that adopted a urban freight distribution strategy based in a UDC combined to a normative policy which limited the entrance to the city for some categories of vehicles. Currently, vehicles with weight higher or equal to 3.,5 T. must deposit their load in the UDC. A private company will be in charge of distributing the freight from the UDC to final destination. In 1999 the service customers were decreasing in number, but in the last years the volume of freight using the service have increased considerably.

London has been one of the first cities in Europe to introduce road pricing measures to limit the traffic in the city center. This strategy was combined to the reduction of parking areas, and the raise of parking tariffs, and was realized for both private automobile and freight distribution traffic. The traffic in the Inner London area (the most restrictive zone of the city in terms of road pricing and parking price) has decreased in the last years in 15%. Traffic diseases and delays have been reduced on 30%. It has become popular and very extended to use the night period for freight distribution, although the normative policies for freight distribution is of the competence of each neighborhood.

### 1.1.2 The unsuccessful projects

In Basel (Switzerland) a system that presents many similarities to the urban freight organization systems of German cities was developed. The transportation carriers are collaborating voluntary, and a UDC have been realized. Three logistic operators realize the urban freight transportation service, in the same way as already presented, using ecological commercial vehicles (3,5 T.), and although theoretically the project was advantageous, it have been ended. The results in terms of traffic and pollution decreasing were good, but most of the transportation carriers did not participated in the initiative.

The case of Genova (Italy) is also particular. The historical center is one of the biggest in Europe, and the geographical situation (the city is situated alongside the ligurian sea, and the city center is located on a hill) makes the city center quite inaccessible for commercial vehicles, for the streets are small and the number of commercial activities of small-medium dimensions is big. To deal with this problem, and the congestion and pollution problems derived from it, the municipality created a "hub" (with a structure and organization was similar to the UDC presented in other examples), and instead of restricting the freight distribution in terms of time restrictions, the first road pricing system in Italy was developed. This was not specific of freight transportation but also extended to private automobile traffic. This system was experimentally tested in 2003. This experience's first results were very positive, and it seemed the system should work well. In 2004, the project became an operating service extended to other areas of the city. The project, which was financed by the Italian Ministry of Environment, was finished after 19 months of activity, as the service stopped for the lack of financial support. The system was not auto-sufficient, so the project was concluded in February 2005. This case evidences the importance of obtaining a system which can be auto-financed at least for covering the operational costs, as well as the importance of tactical and operational strategies.

## 1.2 Important factors that can be considered in decisions related to City Logistics

To evaluate the convenience of a solution for urban freight distribution we can consider different aspects. These aspects can be easily calculated and will be able to serve as a confrontation indicator for different solutions. Before determining such aspects we need to identify different factors that can be used to compare different distribution strategies, on different points of view and for the available data. In this section we will present some of these factors, which have been determined from the experiences and studies presented in sections 1 and 2, and from their conclusions.

### 1.2.1 Aspects related to monetary costs

One of the main factors that is considered for urban freight distribution decisions is the transportation cost. This cost (which can be calculated in km, hours or in a monetary value) correspond to the total cost of the vehicles traveling from the origin to the destinations, then returning to the starting depot. In vehicle routing problems, these costs are considered in terms of distance (or time). Note that these values can be also used to calculate other indicators, such as the pollution or in some cases the noise. However, this is not the only cost that has to be considered in transportation problems. The cost of infrastructural and technological solutions realized for building the transportation system can be also important, depending on the funds that the public authorities dispose for a City Logistics system. These costs are generally not refunded, and it can be important to consider them in some cases. Other costs that could also be considered are the execution costs and the maintenance costs, which have to be assumed by the transportation system in long term optics to assure the system is auto-sufficient and can be operative without the aid of public funds. However, it will depend on the objectives and ideas of the decider to choose which cost aspects are useful to compare the different solutions.

### 1.2.2 Environmental aspects: air and soil pollution

One of the main objectives for the City Logistics systems is to decrease the air pollution. In the last decades, the composition and the variety of fuels has changed, and cars have become more available to people. The changes in living habitudes have raised the usage of the private cars, and the traffic congestion and air pollution are two of the main problems of many European city centers. We can observe different polluting substances, which are directly related to fuel. The first is $CO_2$, which is one of the main sources of global heating. The Kyoto Protocol was adopted in 1997 at the third Conference of the Parties to the UNFCCC (COP 3) in order to achieve "stabilization of greenhouse gas concentrations in the atmosphere at a level that would prevent dangerous anthropogenic interference with the climatic system". This protocol established that the countries which had signed it were to reduce the emission of $CO_2$ by 5% in 2010. The second group of polluting substances is the group of Nitrogen oxides, also known as $NO_x$, which proportions in fuel smokes are variable in the different fuel products. In the last years, plumb and other polluting substances that were very common in fuel have been reduced. Nowadays, new products such as Euro 4 fuels and GPL do not use big quantities of these substances, so they do not constitute a pollution problem for transportation vehicles. The last identified polluting substance is not a gas, but consists of solid particles that are produced by both fuel and tires. In City Logistics, the different types of solutions can influence the air pollution levels in different ways. The organization strategies and the restricting and incentive policies act on the traffic congestion, and consequently on the air pollution. The technological strategies, which arise on the combustion energies for transportation vehicles, can also contribute to reduce the levels of some of these substances. However, not all the "ecological vehicles" are able to reduce in a significant way some of these substances (for example, the reduction of solid particles is a question that have not been studied deeply, and GPL produces less $CO_2$ and $NO_x$ but these emissions are not close to zero). Another problem is the indirect pollution emission (for example, for hydrogen vehicles, the hydrogen creation process requires big quantities of energy which is not always obtained in non polluting ways), but these indirect pollution levels are difficult to estimate. The best solution to reduce pollution seems to be a combination of organization and technological strategies. Some examples have been observed in 2.1, and several systems produce a considerable reduction of both congestion and pollution.

### 1.2.3 Noise measurement and limits

Another factor that has to be considered as an environmental aspect is the traffic noise. At first sight, it seems that noise it's a measurable factor, which can be used to provide objective data. Actually, what is important for human health and for city comfort is not the absolute value of the noise emission but the perception of that noise. Some studies proved that the perception of noise cannot be expressed in absolute values (for example, in dB). The type of noise (frequency, duration), and the nature of the sound respect to the environmental noise can influence the sensation of disturb in each person. Also physiological (illness, weariness, etc.), psychological and social (noises in stations, airports, marketplaces are better tolerated than noises in parks, libraries and other socially considered "quiet" places) factors can modify the perception of the noise in each situation. In this case we can consider noise as a factor that can be used to rank the different solutions, from the less disturbing to the most disturbing, or we can create an indicator which considers not only the quantitative but also the qualitative factors of noise (to differentiate the noise in the day time from the noise at night, or to consider the effects of a general traffic reduction compared to a solution which reduces only the freight transportation traffic.

Another problem of the city centers is the traffic congestion. This can be a measure of the diseases associated to a specific urban freight distribution strategy, and it can be expressed in numbers. For obtaining congestion orders of magnitude for possible solutions the best way is to use traffic prevision models, which will not be defined here. For more details on traffic prevision models see (80) . Another way to estimate traffic for decisions at a strategic planning level can be to evaluate the state-of-the-art in the subject (in this case, the studies resumed in section 2), and, using statistics on similar successful cases, determine average traffic reduction for the considered solutions. It's also important to use the appropriate measures to evaluate the traffic. In this case, the concept of equivalent vehicles seems good for traffic prevision, for it homogenizes the composition of the urban traffic into a comparable measure. A traffic measure will then be defined using this concept, in terms of flow (equivalent vehicles/hour) or in terms of a non-dimensional indicator which considers all what have been presented.

### 1.2.4 Social aspects

Other factors that are to be considered, and could be very useful in some situation, are related to restriction and comfort levels for different categories of people. In city centers, where the main problem is the reduced space and the need for many people to accede or pass through, different categories can be involved in freight transportation problem. We present three of them: transportation carriers, involved commercial activities and other citizens. The first category, the transportation carriers, is often the less considered in the organization of urban freight distribution. However, these carriers are the main actors of urban transportation, and its needs and opinions have to be considered, at least to avoid big conflicts between transportation carriers and public administrations, which can produce other diseases. For this category, restrictive normative policies are not considered as a good solution, but they can be open to alternative solutions as incentive measures or a freight distribution organization which will not affect their economy in a considerable way. The second category, the commercial activities, are the most affected by the freight distribution strategies. For them, freight transportation is necessary to their activity, because their customers will depend on their product offer and availability. They have less instruments to block the system in respect to transportation carriers, and in general these activities are small or medium (big commercial activities have their own transportation service which in general can be compatible with the service provided by the public administrations), so their economy cannot survive without the goods they are proposing. The third group, which is in general the most important for politicians, is the rest of the people, who do not participate directly to the freight transportation but they divide the same transportation network. Trucks blocking a street, problems to park because of freight transportation, and other situations will be considered negative by the usual drivers of city centers. On the other hand, a system which reduces congestion and produces more parking areas, or only the perception of no big commercial vehicles in the city center can be seen as good solutions. Note that all these three indicators are not quantifiable in an empiric way, because they are more related to sociological aspects.

## 1.3 City Logistics strategic planning: Types of strategies and their simulation aspects

As seen in section 1.1, Many city logistics projects are operative in different medium European cities. In big cities, these projects are less common, but some of them can be found. In these examples, different measures were proposed to deal with freight transportation traffic diseases and problems. We can observe that most of the projects follow a similar pattern, which consists in combining different measures to decrease traffic and pollution due to urban freight transport.

Freight transportation planning policies can be applied at different moments and for different objectives. It is then important to distinguish the different planning levels (30) , which are the following:

- Strategic planning involves the highest level of management and requires large capital investments over long term horizon. Strategic decisions determine general development policies and broadly shape the system's operating strategies. It is important to project a good system, which can be operative without big public finance contributions. At this point, infrastructural and technological innovation measures take part.
- Tactical planning aims to ensure, over a medium term horizon, an efficient usage of the existing resources in order to improve the system's performance. It is the phase of control and which evidences the strengths and the weaknesses of the urban freight transportation service.
- Operational planning is performed by local actors in a very dynamic environment. The time factor plays an important role and a detailed representation of the activities, facilities and vehicles is essential.

Tactical and operational planning are also considered while projecting the system at a strategic level. Once the service is operating, daily planning will evidence where the system is not working, and tactical planning will try to solve these problems. When an infrastructural or service organization problem is found, a new strategic planning phase has to be developed to modify the system in order to improve it . These three planning perspectives will influence the type of actions that have to be planned in the whole city logistics system. We will focus on strategic planning, which is the planning level adopted when projecting a new city logistics system.

At strategic level, decisions are related to infrastructural, financial and high level enterprise organization (e.g. directions, guidelines, offered services, number of workers or terminal location). In this section we will describe strategic planning guidelines for urban freight distribution. At this projecting level, the first step is to define the objectives of the service. In city logistics, different objectives can be defined but the directions are only two:

- Reduce congestion and increase mobility of freight transportation services in urban areas, without penalizing the commercial activities of the city centers.
- Reduce air and noise pollution, providing an efficient service to urban commercial activities.

In both objectives, the reduction of noise and pollution diseases will be results of the city logistics application, but normative policies are not always going into similar directions. Reducing congestion pollution is reduced, but reducing pollution with technological solutions will not result into a reduction of pollution.

The fundamental idea that underlines most unsuccessful or not efficient real application initiatives in city logistics, is the fact that each shipment, firm, and vehicle is considered individually. Rather, one should consider that all stakeholders and movements are components of an integrated logistics system. This implies the coordination of shippers, carriers, and movements as well as the consolidation of loads of several customers and carriers into the same ”green” vehicles. The term City Logistics encompasses these ideas and goals and explicitly refers to the optimization of such advanced urban freight transportation systems.

In urban freight distribution strategic planning, some important aspects have to be defined. The first of them is the type of company that will realize the service. In general, the planning and organization of the service is done by the public administration or a company which is controlled by the municipality. Freight transport is usually sub-contracted, but the infrastructures and the vehicles are propriety of the municipality. Another important aspect is financing the transportation system. Two types of funds are usually considered: only public funding and mixed funding (in this case the municipality covers a part of the total funds and the rest are funds coming from banking or financial companies, and in few cases by the transport companies operating in the urban freight distribution system). It can determine the form and objectives of the company which is making the service. The third aspect is to define the role and the guidelines of the transportation service. Once these guidelines are determined, the infrastructures can be defined and projected. A standard city logistics service (called basic City Logistics service) at the strategic level can be defined as follows. Given a urban area and a group of normative policies (which can be adapted and modified to increase the service efficiency), we can identify one or more critical zones. For each zone, one UDC is defined. In medium cities, one UDC is enough (except some particular cases) whereas in bigger cities several UDC can increase the system efficiency). Once the UDC are defined, we need to define the type of service which be proposed. It can be a competition service (different carriers can use the infrastructures, or the City Logistics service is not supported by highly restrictive normative measures), or a monopole service (only one carrier or a group or transport companies which are not in competition between them are allowed to offer the urban freight distribution service). The last but not least common aspect that is also determined at this level is the choice of the vehicle fleet. Given a planning period and a prevision of the freight volumes (with its distribution), the initial fleet of vehicles for the transportation service on the considered period is defined. It's at this that the characteristics of the new vehicles are defined. The trend is to use "ecological" vehicles, or vehicles with a low polluting gas emissions, which in general is limited to three types of vehicles: Euro4 fuel vehicles, GPL vehicles and electric vehicles. Innovation is made in this aspect and in a future new types of vehicles will be proposed, as hybrid commercial vehicles or hydrogen combustion motors.

The service project and planning is more focused on strategic planning, since tactical and operational planning will be done while the service is operating. However, in order to consent these planning operations, a city logistics service has to consider at the strategic level how these planning activities can be done and which tools will be used. It seems important to ensure an efficient service that vehicle routing and crew scheduling have to be planned with the help of decision and optimization tools. The main question is how innovation can be applied at this step or, where it is possible to apply existing standard algorithm and planning tools, to identify and develop standard and unified trip planning tools. In scientific literature, big advances have been made in vehicle routing optimization, which is one of the optimization problems that offers a big quantity of published papers every year. However, most of the commercial tools are based on few simple procedures, and a dominant algorithm is easily identified, with some variants which are based on the same heuristic method: the savings algorithm (22) , developed in the 60's. Only some commercial solvers, which are expensive and used only in universities and some specific fields, but their application to realistic transportation planning is reduced to a few specific cases. Why university research in VRP is not applied in most of real situations is not the objective of this study, but we can see that one of the first algorithms for the VRP, which is considered a base algorithm but is not very performant, can be considered as a dominant design for freight transportation planning tools for the following reasons: its solution is not the best, but the algorithm produces fast results that help to reduce the costs respect to the traditional routing construction methods, the algorithm is cheap and easy to implement (it is considered one of the basic Operational Research algorithms) and some free versions can be found on internet. Another important characteristic is that this algorithm is easy to understand and use, so it is easy to integrate it in a commercial planning tool with familiar interfaces that can be used by people which are not experts on optimization or computer sciences. In the current freight transportation planning methods, this algorithm which is used for more then 40 years can be considered as one standard on vehicle routing applied optimization. Another important point is the coordination between the optimization algorithm and the organization of the enterprise (number of employees, tasks of each employee, importance of calculation times in planning, etc), and a simple fast algorithm is more adaptable to big number of customers and quick changes in the system.

Summarizing, a basic City Logistics system is formed by the following elements:

- First organization aspects: objective of the service and type of company that will plan and control it. At this level, the main factors and problems to be reduced by the system have to be determined and described because the distribution system will be projected to achieve the defined objectives.

- Normative policies, proposed by the municipal authorities. These normative aspects have to be coordinated with the strategic planning concepts, and a successful city logistics solution will not be based only on normative policies, but to present other solutions that sustain and constitute a complement to the presented normative.

- Strategic planning concepts:

  – Infrastructural aspects: In this phase, the infrastructures' usage, alongside to the need of realizing new infrastructures, is evaluated. In real applications, the usage of one or more Urban Distribution Center is common to most City Logistics solutions.

  – Freight distribution network: The distribution system is defined in this step. In general, common city logistics networks are defined as follows: from the suppliers, freight is received on Urban Distribution Centers, which are usually located in peripheric areas of the city. Then, small freight distribution vehicles distribute the freight to its final destination.

  – Vehicles and technological aspects: once the distribution system is defined, it is important to find the adequate technological solution.

  – Transportation planning tools, to optimize routes, to manage vehicles and crews (even in real time situations), or to model the traffic in order to evaluate the different solutions.

From this basic system we can develop more complex organizations but most of the real world operating City Logistics services can be deducted from the proposed standard. The financial aspects are not a standard, as we saw, so they are not considered in the basic system.

To evaluate the performances of a new city logistics service at a strategic planning level, we need simulation and prevision tools. At this level, the service is at the preliminary project level, and there is not enough data to produce very precise and accurate studies.

## 1.4 Simulation of City Logistics strategies

In traffic simulation, main aspects to consider are:

Demand simulation: O/D Matrix definition. Offer simulation: network and arc travel costs construction. Trip planning : vehicle load and customer sequence configuration.

Each strategy type for city logistics planning act on one these aspects in different ways. We present these relations between city logistics strategies and traffic simulation aspects as follows:

- Normative policies:
  - Restrictive:
    - Timetabling for freight distribution: changes on demand (time windows and time period)
    - Road pricing and limitations: network modification
  - Incentive: network modification
    - Information and communication tools: small changes for traffic simulation
    - Infrastructural and civil engineering contributions: network configuration (arcs and nodes characteristics changes)
- Technological innovation in vehicles: not big changes (capacities, speeds)
- Organization an planning optimization: trip planning (transportation cost optimization)

# Chapter 2

# Multi-echelon distribution systems

As we have explained in section 1.3, the strategic level is a fundamental stage in transport planning, because the transportation system, main strategies and other economic, infrastructural and organizational aspects are defined. Moreover, the feasibility of the transportation system project and its adaptation to the different needs and general environment have to be studied. It is therefore important to consider the different motivations which will define the company's main service guidelines (for example characteristics and quantity of freight to transport, customer's needs, distance and network specificities, infrastructural, operational and additional costs, type of service, quality standards and penalties in case of not respecting them). In this chapter, we will describe some of the main transport strategies (see section 2.1), focusing on the importance of multi-echelon distribution. After a brief introduction of the basic concepts on multi-echelon distribution systems (see section 2.2), real case examples will be shown (see section 2.3) and a synthetic survey of the main literature's contributions on multi-echelon systems and their main real applications will be presented (see section 2.4).

## 2.1 Transport strategies

One of the first aspects which takes place in freight transportation strategic planning is the the definition of different shipping strategies. Nowadays, there are several strategies, but they derive from few main aspects:

- Transportation mode: Due to different motivations (characteristics and quantity of freight to transport, distance, costs, urgency, ...) and the offer in terms of transportation modes (road, railway, sea, air ...), a transportation solution will be defined using one or a combination of vehicles. If the transportation from origin to destination is realized using more than one of these modes, the transport is called multimodal or intermodal; otherwise, the strategies are monomodal.

- Vehicle usage: In some road transportation strategies, vehicles are loaded to capacity. This policy is known as truckload (TL). Instead, in other real applications, like in city logistics, most of the vehicles are not full-loaded, so the applied policy is known as Less-Than-Truckload (LTL).

- Hierarchical level: This aspect can be defined using two groups of strategies (direct shipping and multi-echelon distribution).

### 2.1.1 Modality strategies

The mode of transport, also known as means of transport, is the general term used for the different kinds of transport facilities that are often used to transport freight or people, i.e., a combination of networks, vehicles, and operations necessary to transport people or goods from an origin to a destination. The main transport modes for freight are the following:

- Road transportation, which is commonly used to freight transportation by motorized vehicles (trucks, vans and other similar vehicles) using the same roads as private cars and public transport road vehicles.

- Railway transportation, which is more popular for freight in U.S. and Canada, and is used to make medium and long distances.

- Sea transportation. The most common sea transport policy is the container transportation, which is easily combined with container trucks and trains, and also, in the case of transport of vehicles, roll on - roll off sea transport.

- Air transportation: whereas air transport is not one of the cheapest freight transport modes, it has its importance. Air transport is often used by urgent shipping and courier carriers, who use passenger regular and charter flights to send their documents and parcels. However, couriers and parcel service companies also schedule regular cargo flights, and for some products, due to its specificity, urgency or for other reasons, freight cargo regular lines or charter flights can also be used.

- Other modes: inland waterways, animal-powered transport, pipelines, etc.

When a freight quantity has to be transported from an origin to a destination, depending on the nature, the characteristics, the quantity and the urgency of the load, a transportation planner will choose the best solution in terms of modality. Two main strategies can be considered. The first possibility for the transport is to use only one mode. In this case, the transport is considered *monomodal*. Most of the monomodal transport cases are related to road transport (for example, TL carriers, self-made transport for some small commercial activities, local and regional courier and postal distribution, daily press and soft drinks distribution), although examples of monomodal non-road transportation can be found. This is the case of pipeline transportation, ferry and railway transportation services for private cars, or some railway internal transportation in big production factories.

The second possibility is to transport the freight from an origin to a destination using more than one transportation mode. The freight can be repackaged or reorganized, and some labeling, control and security operations can be realized at the trans dock terminals. This transport process, which is composed by two or more different modes is known as *multimodal* (35) . This term is used in a general way, without specifying if the load unit changes in each mode,during the entire transport process. The movement of goods in one and the same loading unit of a road vehicle, which uses successively two or more modes of transportation without handling the goods themselves while changing mode is called *intermodal transport*. By extension, the term *intermodality* has been used to describe a transportation system whereby two or more modes are used to transport the same loading unit or truck in an integrated manner, without loading or unloading, in a door to door transport chain. Moreover, in Europe, the term *combined* transport is used for intermodal transportation systems where the main part of the journey is made by rail, sea or inland waterways, and any initial and/or final legs carried out by road are as short as possible. These terms have been defined by the Intersecretariat Working Group on Transport which members are the *United Nations Economic Commission for Europe*, the

*European Conference of the Ministers of Transport* and the *EUROSTAT*, in a need of unifying transportation terminology (105) .

### 2.1.2 Vehicle usage strategies

In some transportation strategies, vehicles are loaded to capacity. This policy is known as truck-load (TL). Instead, in other real applications, like in city logistics, most of the vehicles are not full-loaded, so the applied policy is known as Less-Than-Truckload (LTL). These two groups of transportation strategies differ in the usage of the vehicle in terms of load (23) . Moreover, LTL carriers are used in real applications to define a specific sector of freight distribution, which companies use an LTL policy. In freight distribution companies, two main sectors are considered: LTL carriers and parcel carriers. A parcel carrier traditionally handles shipments with weights lower than approximately 150 US Pounds (8) . They often handle shipments with a large amount of packages, and their measure unit is the parcel or package. LTL carriers are used to deal with bigger units (like palets or containers), and their rates (per kg) are usually less expensive than parcel carrier's rates. Both types of distribution carriers are similar in the fact that they use the same distribution policy in terms of vehicle usage.

These terms are used in road transportation, and are in general associated to the transportation carrier (46) . However, similar strategies are used as well in other monomodal and multimodal transportation systems. Therefore we will extend these vehicle usage terms to any mode of transport, using the road transport notation. A TL carrier, or full Truck-Load (FTL), is defined as a transportation company, usually operating in the road transportation domain, which generally transports an entire trailer-load to a single customer. These carriers normally deliver an empty vehicle to a shipper who will fill it with freight for one destination. After the vehicle is loaded, it travels to its destination. The vehicle is not supposed to be loaded to its maximum capacity, but what defines a TL policy is that all the load that is transported with one vehicle is delivered to only one destination. Due to the fact that vehicle-Load carriers are asked to ship such a variety of items, a vehicle-Load carrier will often be specialized in moving a specific typology of freight. TL transit times are in general directly related to the distance between the transport's origin and its destination (23; 8) .

Less-Than-Truck-Load (LTL) shipping refers to freight transportation policies where different delivery requests have to be realized and their freight quantities, which are usually lower than the vehicle capacity, can be arranged in order to optimize the vehicle loads, serving in this way more than one request using one vehicle in the same transport operation. Travel times in LTL policies are longer than TL travel times, and they are not only related to the distance between the origin and the destination of the transport operation. Instead, an LTL transportation has to be studied in order to optimize the overall system, considering a number of destinations which have to be satisfied, a fleet of vehicles, and different informations which will be used to determine the different travel costs of the network. The main advantage of using an LTL carrier is that a shipment may be transported for a fraction of the cost of hiring an entire truck and trailer for an exclusive shipment. Also, because of LTL configuration characteristics, a number of accessorial services can be available, which are not typically offered by FTL carriers due to the limitation of direct transport policies.

### 2.1.3 Hierarchical level strategies

The third important aspect in terms of transport strategies is the way the freight goes to the final destination. When the freight arrives to final destination without changing vehicle, a *direct shipping* or *single-echelon* strategy is applied, whereas when freight is delived from its origin to its final destination passing through intermediate points, where the freight is unloaded, then loaded into the same or into a different vehicle, we sepak of a *multi-echelon* system.

#### Direct shipping

A direct shipping service refers to a freight transportation service where the freight is transported from its origin to the final destination without passing through intermediary states. We can however consider two types of direct shipping, related to vehicle utilization. Direct shipping is the most common strategy in TL transportation, and in this case, the vehicle goes directly from the origin to the destination, where all the freight is delivered at the same point. The other case of direct shipping transportation is single-echelon distribution, where the vehicle, starting from the origin, visits different destinations but without passing through intermediary states. The freight is not unloaded until the vehicle arrives to its destination and no intermediary operations are made: between the origin and the destinations only a transport is being carried out and the freight corresponding to each request remains in the vehicle until it arrives to destination.

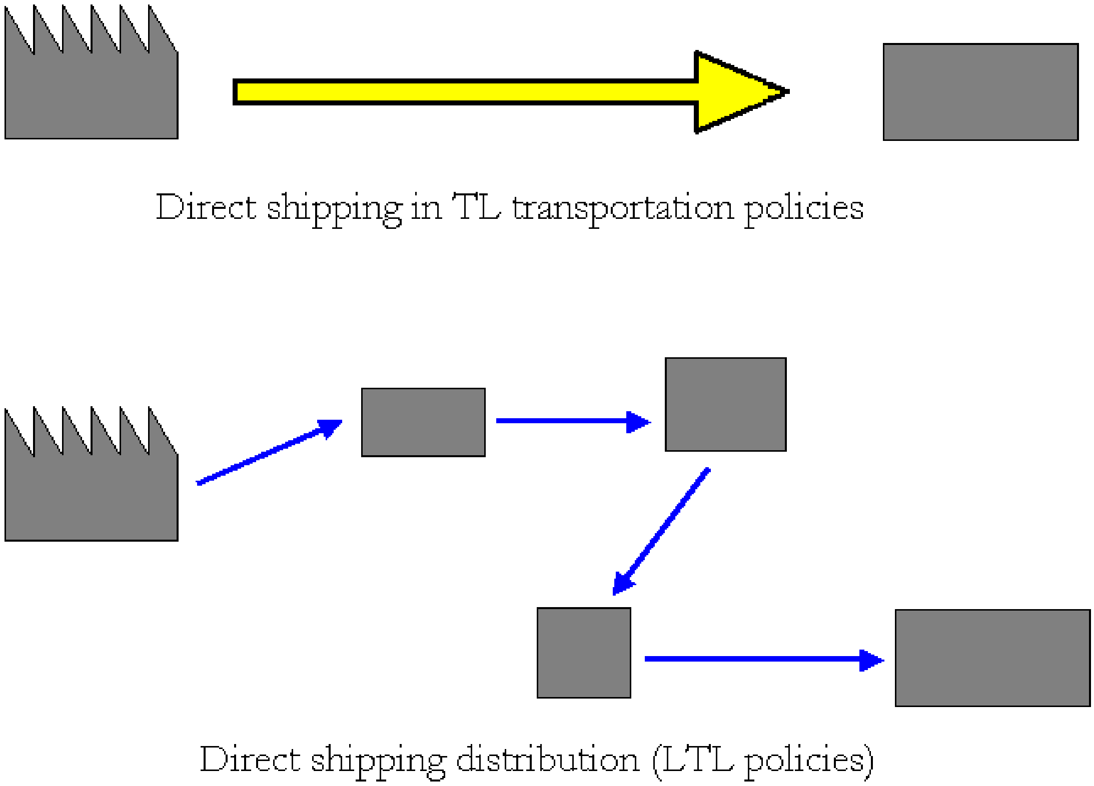


Figure 2.1. Direct shipping on TL and LTL strategies

The advantages of direct shipping are that only the transport is carried out, plus the initial loading and the destination(s) unloading operations. This is traduced in less operations and less cost aspects than those related to multi-echelon systems. Vehicle dimensions are in general related to the transported volumes, although not always the vehicle load is optimized. The main disadvantage is the aggregation level of the freight being lower than those observed in multi-echelon systems. We observe direct shipping in most of TL carriers, some local and regional distribution, and systems where the presence of storage areas and other organizative aspects allow them to separate the whole system into a number of independent (or almost) direct shipping transport operations, plus the operations that connect them (loading, unloading, labeling, etc). Each transport is planned and optimized without considering the other operations, or considering an estimation of their costs and times.

Direct shipping is also used by some commercial activities where the transport of goods is not made by the supplier but by the commercial activity. In general, these small and medium enterprises have a small or medium vehicle, which is used to go (empty) from the commercial activity to the suppliers' location, where it is loaded and then returns to its starting point.

**Multi-echelon transportation**

In transportation, it is not always possible or convenient to deliver the goods directly to the destination. In fact, some transportation systems use intermediary terminals where some operations take place. The different vehicles that belong to these systems stop at some of these points, and in some cases the freight changes vehicle or even mode of transport. Moreover, some additional services, like labeling, packaging, assembling, etc, can be realized at these intermediary terminals. These systems are called multi-echelon (9) , because they are composed by one or more levels, or echelons. Usually, in transport optimization, these systems are decomposed in several single-echelon independent problems, and eventually they are solved separately.

These systems require an initial inversion which can be higher than those corresponding to direct shipping strategies. Additional costs for crew, maintenance and operational organization have also to be considered, because each level also has its workers who are realizing several operations. On the other hand, multi-echelon systems present their advantages, as for example, the possibility of freight aggregation, the usage of regular freight transportation lines with very convenient costs, etc. (80)

## 2.2 Multi-echelon distribution: concepts and definitions

Many real applications follow a multi-echelon strategy (examples of these applications will be shown in section 2.3). This distribution strategy will be defined in this section, and the main concepts and definitions will be presented. The general aims of the process are to ensure an efficient and low-cost operation of the system, respecting different constraints: demand delivered on time, reduction of traffic congestion, etc. Freight has to be delivered from one or more origin points to different destinations. These origins represent industrial facilities or big distribution centers. The destination points are defined by their freight request. Each request is composed by a number of informations, which usually are:

- geographical location of the destination, associated with the request;
- freight quantity to be delivered, known as demand;
- service quality features, such as an indicative delivery time period, measure unit (container, pallet, box, package), retirement mode, and other service policies. To deal with service quality, some penalties can be associated to the non respect of the delivery conditions.

### 2.2.1 Types of multi-echelon systems in freight transportation and distribution industry

Multi-echelon systems are not limited to situations where only the transportation aspect is being considered. In fact, in several industrial fields we find multi-state and multi-level production, distribution and transportation situations that are examples of multi-echelon systems. The main cases are the following:

- Supply chain: a supply chain, also known as logistics network or supply network, is the system of organization, people, technology, activities, information and resources involved in realizing and transporting a product or service from a supplier to a customer. Supply chain activities transform natural resources, raw materials and components into a finished product that is delivered to the customer (9) .
- Multi-echelon inventory systems: this is the technical name given to systems made by one or more factories, a number of storage areas, known as warehouses, and the final destination of freight. Freight requests are not made directly to factories, but to warehouses, which have a stock of freight. These warehouses, which belong in general to distribution enterprises, command freight in big quantities to factories. The aim of these enterprises is not to optimize the transportation cost of the overall system, but to maximize their profit, which can be improved by minimizing the storage costs and the final echelon (from

the warehouse to each destination) transportation costs. These systems involve transportation but each echelon transport can be organized separately because of the system characteristics (9) .

- Multi-echelon transportation systems: a complete transport is considered a multi-echelon system when the vehicle passes through one or more intermediary states where different operations are made. Not all multi-echelon transportation systems follow a distribution logic (for example, multi-echelon TL strategies such as combined transportation rail/road or sea/road where the entire vehicle is loaded respectively into the train or ship, or international TL transportation between countries where customs taxes have to be applied).
- Multi-echelon distribution systems: a distribution system where a multi-echelon logic is applied. Typically, the freight goes from the origin to one or more intermediary points, where different operations are realized, and in some of these points it is unloaded and then loaded into a different vehicle. This is the case of multimodal transportation if we consider it as a system, and in general the distribution sector is using this strategy. In literature, most studies in the supply chain and inventory research fields use the term distribution as the set of operations to which a parcel or freight transport unit is exposed, in the whole distribution process from its origin to its destination. In this study, we refer to distribution as it's defined in transportation science (80) , so without considering .

In all of them, freight must pass through one or more intermediary facilities, which represent intermediary terminals. In supply chain, multiple assembling and production operations take place, and the overall system from the origin (raw products) to the final destination (customer or shop receiving the final product) represents an entire process. In inventory systems, freight can be stored at the intermediary facilities, and in general these points are used as intermediary warehouses where stocks have to be controlled, planned and optimized. In only transportation systems (both TL and LTL strategies) the main activities of these intermediary platforms are related to freight transport or distribution, or to its organization. We will focus on multi-echelon distribution systems, which are those where the connexion between levels which takes place at the intermediary facilities is most difficult to understand and represent in system modeling. At these facilities, several operations take place, offering the basic or some additional services to freight transport. In next section we will describe these activities and show the importance of these intermediary facilities.

### 2.2.2 Intermediary facilities in freight distribution

As presented above, multi-echelon systems are characterized by one or more groups of intermediary stages where various operations can be achieved. In these intermediary facilities, some operations take place, to help the distribution process, reduce costs, give a higher quality service or offer some additional services to vehicle drivers. We will introduce in this section the activities that are more extended in real applications.

One of the most important group of activities that take place at the intermediary platforms is related to trans-doc operations (33) . In most of multi-echelon transportation cases, the main characteristics are related to vehicle changing at least in one intermediary terminal. In these cases, freight is unloaded from the arriving vehicle, then loaded into a different vehicle. This freight can be exposed to package or organization changes, or can change vehicle without submitting changes on the measure unit (i.e., the entire load does nor change nature, form and content in the trans-doc operation).

Other important operations, which are common in many distribution fields, deal with freight reorganization. In some real applications, as for example newspaper or fresh alimentary products distribution, the companies have to deliver products coming from different producers to each destination point (62) . To reduce costs, this freight is reorganized at the intermediary points, where each customer's request is composed by aggregating its demand from each producer, and then the vehicles are loaded.

Another aspect associated to these facilities is the freight storage (9) . Freight can be deposed at the terminals for a small period of time (the necessary to complete the other operations); in these cases, the system can be modeled without considering inventory aspects. When freight is stocked and distributed gradually in function of demand trends and requests, inventory systems can model the whole system. Although in transportation systems production activities are not considered, some additional operations and services can take place at intermediary platforms. For example, labeling, control, package making or the preparation of promotional and special offer products that are not realized by the producers but by the distribution companies.

### 2.2.3 Multi-echelon distribution system organization: definitions and notation unification

As it will be presented in section 2.4, some studies have considered multi-echelon system cost optimization, but the main difficulty of individuating an classing them is that each field uses a different notation and no standard vocabulary has already be proposed. To deal with it, we propose a general definition of a multi-echelon distribution system, presenting the vocabulary and notation which will be followed in this study.

In a multi-echelon distribution system, it is not possible to deliver the freight directly from the origin to the final destination of the request. In fact, freight goes to one or more intermediary facilities, where some of the operations presented above take place. If we define a N-echelon distribution system, N intermediary levels are considered. Each level $k$ has a number of $k$-level intermediary facilities associated to it. The overall transportation network can then be decomposed into N levels:

- the 1st level, which connects the depots to the 1st-level intermediary facilities;
- $N-2$ intermediate levels interconnecting the different intermediary facilities;
- the $N^{th}$ level, where the freight is delivered from the (N-1)th level intermediary facilities to the final destinations.

To deliver the freight, a number of vehicle fleets are defined. Each level usually has its own fleet of vehicles, defined by different characteristics (capacity, dimensions, speed), and can be heterogeneous or homogeneous.

In real applications two main strategies for vehicle disposition at each level can be considered. Given a level, the corresponding vehicles can be associated to a common parking depot, from where they are assigned to each satellite in function of the satellite demand and the vehicle activation costs, which will also depend on the traveling costs from the parking depot to the

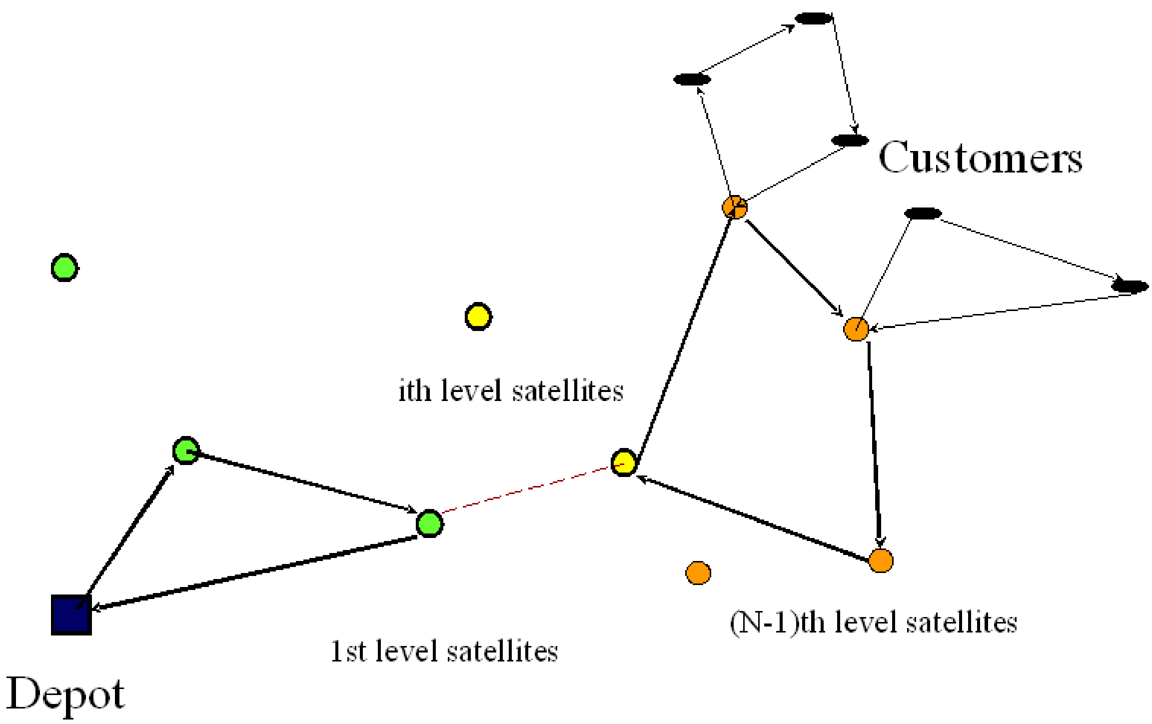


Figure 2.2. Example of a N-echelon distribution network

satellites. Another strategy consist in associating to each satellite a number of vehicles, which will start and end their routes on the considered satellite. In our case we will consider the first strategy, considering similar costs for the assignment of each vehicle to a satellite. In this case, each transportation level has its own fleet in order to deliver the goods. The vehicles assigned to a level cannot be reassigned to another one.

## 2.3 Examples of multi-echelon distribution systems

In this section, we will focus on some of the real cases where multi-echelon transportation systems are used, and their main applications. In the past decade multi-echelon systems have been introduced by practitioners in different areas, in order to deal with different situations, in particular to reduce storage costs and the number of shipments. Some of the real transportation systems that follow multi-echelon strategies are:

- *Logistics enterprises and express delivery service companies.* These operators are usually in a multi-echelon system. Their offices are used as intermediate points where the freight that has to be delivered is being organized and the vehicles which will transport the freight to another intermediate point (airport, regional center, etc.) or to the final destination are being composed. (102)

- *Multimodal freight transportation.* In the past decade, the number of intermodal logistics centers in the countries of central and south-west Europe increased. This is a good example of freight distribution involving two or more echelons (90) . In a classical road-train multimodal distribution the freight goes from the producer to a logistic center by road and then it is loaded on a train that goes to another logistics center. The train is unloaded and the freight goes by road to its final destination.

- *Grocery and hypermarkets products distribution.* Large companies use hypermarkets as intermediate storage points to serve smaller stores and supermarkets of the same brand in urban areas.

- *Spare parts distribution in the automotive market.* Some companies uses couriers and other actors to deliver their spare parts. This is the case of FIAT and General Motors, whose spare parts are distributed by TNT (102) from their factories to the garages. Similarly, Bridgestone (17) uses an organization of the distribution system in zones and sub-zones, to decrease the transportation times and reduce the size of the storage areas.

- *E-commerce and home delivery services.* The new possibilities given by the development of e-commerce and the home delivery services offered by some supermarkets and other stores like SEARS (96) , in some large cities imply the presence of intermediate depots used to optimize the delivery process.

- *Newspaper and press distribution.* In Denmark, a comparative study of heuristics for solving a two-echelon newspapers distribution problem was made for two competing newspaper editors who shared printing and distribution facilities for reducing the total costs (62) . In press distribution, it is also common to see distribution companies that receive the publishing products from the editors and distribute them to the selling points. This is also a two-echelon distribution system.

- *City logistics.* As we explained in chapter 1, in the past decade, researchers started to investigate the urban areas as a system, without considering each shipment, firm, and vehicle individually. Rather, one should consider that all stakeholders and movements are components of an integrated logistics system (31) . This implies the coordination of shippers, carriers, and movements as well as the consolidation of loads of several customers and carriers into the same "green" vehicles. The adopted distribution system is typically a two-echelon system.

## 2.4 Literature review for multi-echelon systems

In transport systems planning and development, multi-echelon systems are theorized. In civil engineering, constructive and organization procedures are proposed for projecting intermediary facilities. Moreover, in freight transportation planning, many multi-echelon systems are used and developed, being considered as complex systems where many aspects have to be considered. One of them is the individuation and optimization of the various costs derived from the operational usage of the system. However, in practical cost optimization and management sciences applications to multi-echelon transportation systems, the common strategy is to separate and approximate costs, using the existing tools for travel cost optimization developed for single echelon cases. Only a small number of families of optimization problems consider multi-echelon transportation systems in its overall form instead of approximating their multi-echelon characteristics, and the existing literature related to them is not easy to individuate, for several reasons. The main is that multi-echelon systems are used in different fields, and sometimes applied to specific transportation cases, such as truck-and-trailer distribution, waste collection or applied to a specific freight category which should need dedicated models and methods due to its specificities. Therefore, notations and definitions are usually specific of the considered problem and don't follow standard guidelines, which difficult the individuation of interesting related problems. Another reason is that multi-echelon systems are relatively new in most transportation cost optimization fields, and some aspects are being studied and developed only in current studies. The last, but not least, the difficulty of considering the overall system and the specific aspects of some transportation cases make multi-echelon transportation cost optimization a challenging field, but it also shows the difficulty of developing theories and standard models and methods.

Several fields in operations research and management science, such as location and network design, supply chain, inventory or travel cost optimization, have considered the need of study the entire system without using approximations that derive from splitting the overall system into a number of cases which can be treated separately. In this section we will present the many contributions to cost optimization literature related to two main fields. The first of them is supply chain and inventory optimization (which will be presented in section 2.4.1), where transportation cost is not considered as a system but only the last stager whereas in the others are approximated, or it is supposed that the overall transportation costs can be approximated as the sum of each system best costs without considering the aspects related to the connexion between the various stages. The second one is facility location and network design (which are introduced in section 2.4.2. These approaches don't optimize the transportation costs in tactical or operational operations but considers the multi-echelon nature of the system in strategic planning.

### 2.4.1 Supply chain and inventory systems optimization

As described in section 2.1.3, supply chain are multi-echelon systems, by definition. Supply chain optimization is a field where many studies have been made. A complete review on the terms and main studies up to the late 90's was made by Beamon (9) . The author defined the supply chain as an integrated manufacturing process wherein raw materials are converted into final products, then delivered to customers. At its highest level, a supply chain is comprised of two basic, integrated processes: the Production Planning and Inventory Control Process, and the Distribution and Logistics Process. These Processes, illustrated below in Figure 2.3, provide the basic framework for the conversion and movement of raw materials into final products. Other supply chain surveys (54; 41) , completed Beamon's work until works presented in 2007.

The Production Planning and Inventory Control Process encompasses the manufacturing and storage sub-processes, and their interface(s). More specifically, production planning describes the design and management of the entire manufacturing process (including raw material scheduling and acquisition, manufacturing process design and scheduling, and material handling

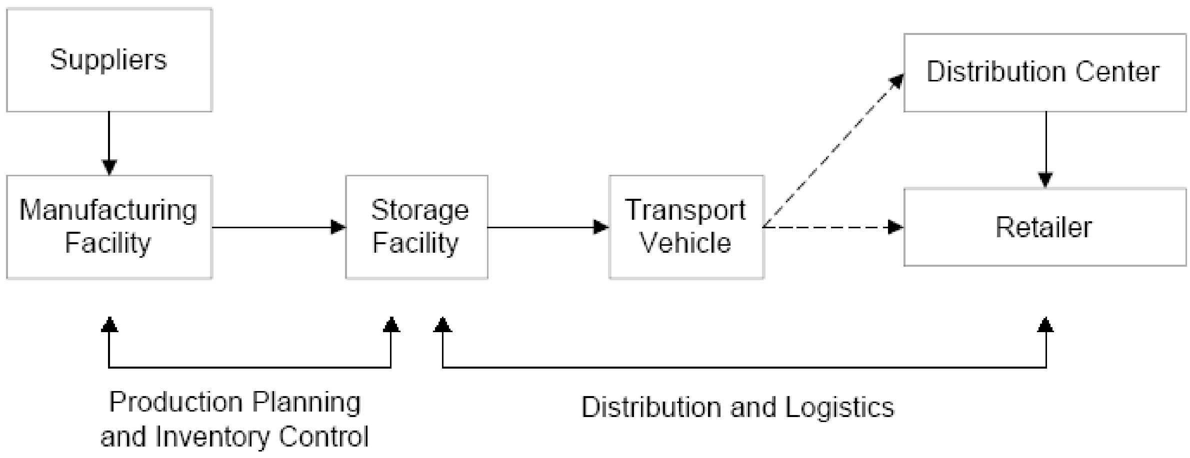


Figure 2.3. Example of Supply Chain process

design and control). Inventory control describes the design and management of the storage policies and procedures for raw materials, work-in-process inventories, and usually, final products.

The Distribution and Logistics Process determines how products are retrieved and transported from the warehouse to retailers. These products may be transported to retailers directly, or may first be moved to distribution facilities, which, in turn, transport products to retailers. This process includes the management of inventory retrieval, transportation, and final product delivery. These processes interact with one another to produce an integrated supply chain. The design and management of these processes determine the extent to which the supply chain works as a unit to meet required performance objectives. Note that in most supply chain models, a single echelon LTL strategy is considered to define the transportation system at this process.

The supply chain illustrated by Figure 2.3 consists of five stages. Dullaert et al. (41) divide also the existing models into two categories, depending on the aim of the model: the first category contain supply chain design models and the second one simulation and planning tools. Generally, models of both categories can be divided into four groups, related to the modeling approach used in them (9) . This modeling approach is driven by the nature of the inputs and the objective of the study. The four groups are:

- deterministic analytical models, in which the variables are known and specified;
- stochastic analytical models, where at least one of the variables is unknown, and is assumed to follow a particular probability distribution;

- economic models;

- simulation models.

In these models, facility location can be modeled using location-allocation problems, and the distribution process is usually considered to be an inventory system, since the storage and inventory costs are in general bigger and more difficult to decrease than transportation costs. We will present in this section the main contributions on integrated production-distribution planning models and algorithms. The facility location supply chain and design problems are respectively location-allocation and network design problems, and will be considered in the considered sections. In the surveys presented below (9; 54; 41) we observe that the distribution process usually follows a single-echelon policy or, in the cases of multiple echelon, the transportation cost optimization is not made in a tactical or operational planning approach, i.e. these costs are not optimized in a whole system where each echelon travel costs depend on the precedent and/or following echelons' travel costs. We will not enter in the detail but we will present the general works in this field.

Arntzen et al. (5) develop a mixed integer programming model to represent a generalized supply chain process. The so-called Global Supply Chain Model (GSCM), that can accommodate multiple products, facilities, stages (also known as echelons), time periods, and transportation modes. More specifically, the GSCM minimizes a composite function of activity days and total cost of production, inventory, material handling, overhead, and transportation costs. These costs have a fixed and a variable component. The model requires, as input, bills of materials, demand volumes, costs and taxes, and activity day requirements and provides, as output: the number and location of distribution centers, the customer-distribution center assignment, the number of echelons (amount of vertical integration), and the product-plant assignment.Other deterministic models follow similar approaches, and are usually considered particular cases of the GSCM. Most common approaches are two or three phase procedures, where in a first phase facilities are located, in a second customers are assigned to facilities and in some cases a third phase minimizes the last stage transportation costs.

Jayaraman and Ross (63) describe the PLOT design system (Production, Logistics, Outbound, Transportation), characterized by one central manufacturing site, multiple distribution center and cross-docking sites and customer zones with demand for multiple items. The PLOT design system is a sequential approach using two different models. The first is a strategic model in which decisions on opening or closing warehouses and cross-docks, and on assigning customer zones to cross-docks and cross-docks to warehouses for each commodity are made. The second model in the PLOT design system is an operational model in which the optimal flow of goods is determined through the network proposed by the first model. A simulated annealing algorithm is presented that solves both models simultaneously. Computational experiments with up to 5 warehouses, 15 cross docks, 75 customer zones and 3 commodities are reported. Within one second, these problem instances can be solved to within 5% from optimality.

Some planning models consider multi-echelon distribution systems in the entire supply chain process. Svoronos and Zipkin considered multi-echelon, distribution-type supply chain systems (100) . In this distribution type each facility has at most one direct predecessor, but any number of direct successors. They also assume a base stock, one-for-one replenishment policy for each facility, and that demands for each facility follow an independent Poisson process. The authors obtain steady-state approximations for the average inventory level and average number of outstanding backorders at each location for any choice of base stock level. Finally, using these approximations, an optimization model is proposed to determine the minimum-cost base stock level. Other multi-echelon problems applied to supply chain will be described in next section, since the distribution phase is in general modeled as an inventory system.

Other recent contributions on inventory systems can be found in (34; 99; 107) . These problems do not use an explicit routing approach for the different levels because their aims are not to optimize the transportation costs but to estimate and optimize the inventory levels. Therefore, storage and inventory costs are supposed more important than transportation costs. Some cases consider the possibility of optimizing the transportation costs, but because of the possibility of storage of the different items, there is not a notion of a system of dependent echelons (in terms of transportation cost) in the overall costs optimization operation, but the different levels are considered as independent in terms of transportation strategies. This is not a multi-echelon transportation system, but a distribution system which considers a number of related transportation systems in its overall process.

### 2.4.2 Multi-echelon facility location problems

Other multi-echelon systems have been studied in operations research related to freight transportation, mainly in facility location, network design and location-routing. The main difficulty of listing the literature contributions is in that there is not a standard terminology and authors often use synonymous terms to the considered distribution systems which are not always evident. Some studies deal with the location of intermediary facilities for a multi-echelon distribution systems (90; 31) , for real or realistic applications. These problems can be generalized into the two-echelon facility location problem.

Hinojosa et al. (59) deal with a multi-period two-echelon multi-commodity capacitated location problem. For each period, a decision is to be taken on which plants and warehouses to open or close and on the amounts of the different products to be shipped from the plants to the warehouses and on to the customers. A mixed integer programming formulation is given and a repair heuristic for obtaining feasible solutions from the lower bounds from a Lagrangean relaxation approach is developed. Computational experiments show that this approach is acceptable for small and medium-sized problem instances.

Marin (72) developed a mixed integer formulation based on twice-indexed transportation variables for the two-echelon uncapacitated location problem, followed by an an analysis of several Lagrangian relaxations to determine good lower bounds on its optimal value.

An application of a two-echelon distribution network is due to Crainic, Ricciardi and Storchi and is related to the city logistics area (31) . They developed a two-tier freight distribution system for congested urban areas, using small intermediate platforms, called satellites, as intermediate points for the freight distribution. This system is developed for a specific case study and a generalization of such a system has not yet been formulated.

A more complex variant of the same problem is the multi-commodity multi-plant capacitated facility location problem proposed by Pirkul et al. (85) . In this problems, the optimum set of plants and warehouses has to be chosen from a potential set and plan production capacities, warehouse capacities and quantities shipped so that the total operating costs of the distribution network are minimized. The entire system presents multiple stages, and the authors present a computational study to investigate the value of coordinating production and distribution planning, with a MIP and a heuristic based on Lagrangian Relaxation.

Ambrosino and Scutell (3) offer a mathematical programming formulation for a number of static and dynamic scenarios based on the general multi-echelon location problem discussed above. To explore the computational complexity of the models, linear programming approaches are used to find the optimal solution or at least provide lower bounds for problem instances based on a real-life case. Computational testing is limited to locating distribution and transshipment points and assigning large customers and customer zones to these distribution facilities. As such, the multiple echelon approach and routing considerations discussed in the earlier sections are not explored in the computational experiments. The optimal solution could only be found for the smallest problem instance, involving possible locations for 2 distribution centers, 5 transshipment points, 5 large customers and 25 customer zones. As the problem instances become larger, the gap between the best integer solution found, within a time limit of several days for the large instances, and the MIP lower bound provided by CPLEX increases rapidly up to more than 45%. As a result, heuristic approaches seem to be more appropriate, even for the scaled down problems used for the computational experiments.

# Chapter 3

# Transportation costs optimization: an overview on vehicle routing

In the last years of the $XX^th$ century and the beginning of the $XXI^th$, the new markets and the globalization have increased the need of reducing the final product costs to compete with the concurrence. Transportation is an important factor that constitutes up to 40% of the final product cost. For these reasons, the usage of *optimization tools* has increased in the last decades. Most of these tools are based on *operations research* and *mathematical programming* techniques. One of the most common decision problems in transport planning is vehicle routing, which can be applied in all levels with different objectives. At strategic level we find location and network design as the main examples of combinatorial optimization aspects of this phase. Routing and scheduling optimization can be common to tactic and operational planning. In general, static problems are used in tactical operations, whereas dynamic aspects are usually characterizing operational planning optimization tools.

One of the most studied combinatorial optimization problems, which present many real applications in tactical and operational planning strategies, is the Vehicle Routing Problem (VRP), which concern the freight distribution following a monomodal single-echelon LTL policy. This problem has become a central problem in the fields of logistics and freight transportation and the utilization of computerized methods for transportation has often resulted in significant savings ranging from 5% to 20% in the total costs, as reported in (103). Usually, in real world applications, the problem can be formulated in different ways to represent different constraints and situations. In this chapter, we will present the basic concepts of VRP as well as its most studied variants and their applications. In a first time, we will briefly describe the decision problem, based on the main bibliographic references, which will also be presented. Then, the different VRP variants and its applications will be presented. Because other interesting combinatorial problems are direct or indirectly related to vehicle routing optimization, we will also present an overview on the problems which concepts and results can be interesting for multi-echelon vehicle routing optimization. In next section, the vehicle routing problem and its variants will be presented. In section 3.3 we will introduce the Location Routing Problem.

## 3.1 Problem families and general classification

In tactical and operational planning cost optimization, two main aspects are considered: the costs derived from the vehicle usage, which are related to the distance traveled by each vehicle, and the costs of the transportation system operators. These costs depend on the employees scheduling, notably on the vehicle crew organization. Note that in most transportation companies the drivers earn an established salary which does not depend on the total time traveled.

Although most of these applications have as main objective to minimize the transport costs of the entire system, a wide variety of objective functions can be also considered for these problems. The most common ones are the following :

- Duration of a route: total time a vehicle needs to execute the route. This includes travel times, waiting times, loading and unloading times and break times.
- Completion time of a route: the time that service is completed at the last location.
- Travel time of a route: total time spent on actual traveling between two locations.
- Route length: total distance traveled between different locations.
- Client inconvenience: difference between actual arriving time at a customer and desired arriving time. In this case, in some constraints can be added to represent, as a high cost raise, the situations where a customer is not delivered respecting the service conditions. The most common application is the definition of soft time windows.

- Number of vehicles. Normally this function is used as a secondary objective. On some problems in which the costs of vehicles and drivers are the most expensive part of the system costs, the number of vehicles can be the main objective function to minimize. When the number of vehicles is considered along with the total cost, it is common to associate a big fixed cost to vehicle usage, and can be done when the number $m$ of available vehicles at the depot is greater than the minimum number of vehicles required, $m_min$ . Examples of this problems can be found in (57) .

- Profit: in this function there are considered costs and revenues associated with freight transportation. Profit is the difference between revenues and costs.

## 3.2 Vehicle Routing Problems

Vehicle Routing is one of the most challenging and studied combinatorial optimization fields. In some market sectors, transportation costs constitute a high percentage of the value added of goods, so using efficient optimization tools to reduce these costs will result on considerable overall costs reduction, as seen above. Many surveys and studies on this problem and its variants have been realized in the last decades. In this section we will present the VRP, giving the basic concepts (see section 3.2.1) and presenting the main method and algorithm contributions in this area (see sections 3.2.2 and 3.2.2). Then, a brief survey on the most studied variants and other interesting problems which are obtained from the basic VRP will be shown (see section 3.2.4).

### 3.2.1 Definition and basic versions

The Vehicle Routing Problem (VRP) is the generic name given to a whole class of combinatorial optimization problems in which a set of routes for a fleet of *vehicles* based at one or several *depots* must be determined for a number of geographically dispersed points, called *customers*. These vehicles are operated by a set of crews, known as *drivers*, and are traveling to customers using an appropriate *road network*. In particular, the solution of a VRP is obtained by the determination of a set of *routes*, each performed by a single vehicle that starts and ends at its own depot, such that each customer's requirement is fulfilled, all the *operational constraints* are satisfied, and the overall *transportation cost* is minimized. For a detailed definition of the problem and the several models used to defin the basic versions, see (103) .

If we consider the vehicle fleet, two different cases can be considered. The first, where it is supposed that each vehicle have an unlimited capacity (or this capacity does not become a restrictive parameter to be considered), the problem is called Uncapacitated VRP (UVRP), also called multi-TSP. In basic Traveling Salesman Problem (TSP), a graph $G$ with one depot and a number of customers $n$ is given. Given two nodes of the graph $i$ and $j$, the cost $c_{ij}$ represent the cost of going from node $i$ to node $j$. The objective of TSP is to minimize the total travel costs of the round-trip route that, starting and finishing at the depot, visits each customer exactly once. Note that in basic TSP no capacity constraints are considered. A variant of this problem is obtained considering that more than one routes can be considered, without changing the other characteristics of the problem. This problem is known as Multi-TSP, i.e. a version of the Traveling Salesman Problem where more than one salesmen are considered.

The basic problem of the vehicle routing family is then the Capacitated VRP (CVRP, also noted in some cases as VRP). In this version, all customers have to be visited, the demands are known, all vehicles are identical and they all belong to the same depot. The only imposed constraints, other the network and route configuration limitations, are related to the capacity of the vehicles. The objective is to minimize the total travel cost, which is the sum of each route's cost. Using graph notation the CVRP may be described as follows: let $G = (V,A)$ be a complete graph, where V = {0; ... ; n} is the set of the entire graph nodes, where node 0 represents the depot and nodes {1; ... ; n} the customers. Consider two connected nodes i and j; node i is connected to node j by an arc ij. If $c_{ij} \neq c_{ji}$, the graph is directed and the problem is called asymmetric capacitated VRP (ACVRP); otherwise the problem is called symmetric capacitated VRP (SCVRP) and $c_{ij} = c_{ji}$. It is a common assumption that the cost matrix satisfies the triangle inequality,that is

$$c_{ik} + c_{kj} \leq cij \forall i,j,k \in V \tag{3.1}$$

It is also often assumed that each customer is associated with a point in the plane and the cost $c_{ij}$ is equal to the Euclidean distance between the two points.In this case the distance matrix is symmetric and satisfies the triangle inequality. The resulting problem is called Euclidean

SCVRP.

A set of $m$ identical vehicles, each of them with an identical capacity K, is available at the depot. A known nonnegative delivery demand $d_i$ is associated with each customer $i$. In classical CVRP, each customer must be visited by one and only one vehicle. In general, $K \geq d_i \forall i = 1...n$, but in some cases we can find a customer with a demand greater than the vehicle capacity. This situation, in the basic case where each customer must be served by one vehicle, can be modeled in the following way. Consider a customer $i$ with a demand $d_i \geq K$. This customer is represented as a number of $\lceil d_i/K \rceil$ identical customers; in this case, $\lfloor d_i/K \rfloor$ with full capacity are serving node $i$ using a TL policy, whereas another vehicle will include in its total load the remaining quantity of goods having to be delivered to this point.

The CVRP requires the computation of at most K routes with minimum cost. Each route is a vehicle tour such that:

1. each tour starts and ends at the depot;

2. each customer is visited once;

3. the sum of the demands of the customers visited in a tour does not exceed the vehicle capacity Q.

The CVRP is known to be NP-hard in the strong sense because the well known Traveling Salesman Problem (TSP) arises as a special case, as seen above, when we consider only one vehicle of unlimited capacity. Because of this, mathematical models for CVRP are not able to solve medium instances in low times. In literature, this problem have been one of the most studied, and recent studies have developed good methods. Exact algorithms can solve relatively small instances and their computational effort is highly variable (25) . For this reason, exact methods are mainly used to determine optimal solutions of the test instances, while heuristic methods are used in practical applications, and bigger instances. In the next two sections, we will present these two groups of methods, with some examples and applications.

### 3.2.2 Exact methods for the CVRP

In this section, the most commonly used exact methods for solving Vehicle Routing Problems are presented, showing the most representative studies and their limits. As the name suggests, these approaches propose to calculate every feasible solution until the best is reached and proved. Because the number of variable combinations is extremely high even for small instances, these methods propose rules to exclude parts of the solution space, in order to find the best solution by reducing it. Once a solution is found, these methods find better solutions until they prove no solution with lower cost than the current can be found.

One of the first and most successful approaches to find an exact solution to CVRP is Branch and Bound. This method uses a divide and conquer strategy to partition the solution space into subproblems and then optimizes individually over each subproblem. Using branch and bound, we initially examine the entire solution space S. In the processing or bounding phase, the problem is relaxed. By doing it, we admit solutions that are not feasible for the not-relaxed problem. The solution to the relaxed problem is in fact a lower bound on the optimal solution. If the solution to this relaxation is a member of S, the optimum is found. Otherwise, we identify n subsets of S. Each of them (which is subproblem or child of S), is added to the candidate subproblems list (which await processing). This is called branching. To continue the algorithm, we select one of the candidate subproblems and process it. There are four possible results:

- If a feasible solution is found, and it is better than the current one, we replace it with the new solution and continue.
- If the subproblem has no solution, it is discarded.
- Otherwise, we compare the lower bound for the subproblem to the global upper bound, which is given by the value of the best feasible solution to the not-relaxed problem encountered thus far. If it is greater than or equal to our current upper bound,the subproblem can be discarded.
- Finally, if the subproblem cannot be discarded, branching is required, i.e. the children of the considered subproblem are added to the list of active candidates. We continue in

this way until the list of candidate subproblems is empty, at which point our current best solution is, in fact, optimal.

Another method which follows a similar procedure (branching) but with a different approach to exclude solution subspaces is Branch and Cut. This approach, similarly to Branch and Bound, starts by relaxing the decision problem, and solving this relaxation. When an optimal solution is found, and this solution is not belonging to the solution space of the non-relaxed problem, a cutting plane algorithm is used to find further constraints which are satisfied by all feasible points but violated by the current solution. If such an inequality is found, it is added to the linear program. This process is repeated until either an integer solution is found (which is then known to be optimal) or until no more cutting planes are found.

The philosophy of Branch and Price is similar to that of branch and cut except that the procedure focuses on column generation rather than row generation. In fact pricing and cutting are complementary procedures for tightening an LP relaxation. Then to check the optimality of an LP solution a subproblem called the pricing problem which is a separation problem for the dual LP is solved to try to identify columns to enter the basis. If such columns are found the LP is reoptimized. Branching occurs when no columns price out to enter the basis and the LP solution does not satisfy the integrality conditions. Branch and Price, which also is a generalization of Branch and Bound with LP relaxations, allows column generation to be applied throughout the Branch and Bound tree. Although Column Generation bounds have been obtained with other models, the most common formulations derive from set covering and set partitioning models for CVRP where variables represent the usage of one entire route. In this way, each column represents a possible route, and the Column Generation method is used to produce lower bounds to the considered CVRP model. Consider the linear problem P which is the relaxation (in general linear) of the CVRP to solve. Consider a problem P' (Master Problem) which corresponds to the case of P where only few columns (at least one feasible solution) are considered. This problem P' is solved, and the reduced costs of it are calculated. If no negative reduced cost is found, the problem is solve to optimality, and the solution obtained is a lower bound to the non-relaxed CVRP formulation. Otherwise, from the reduced costs calculation (for example, using a Shortest Path Problem with Resource Constraints procedure) we can obtain a number of columns to add. This procedure is called column generator. These columns are added to P' then this method is repeated until no negative reduced costs are found.

For a detailed survey of the exact methods see (29; 87; 66) .

### 3.2.3 Heuristic methods for the CVRP

The VRP is a hard combinatorial problem, and exact algorithms can only solve relatively small instances with in general high computational times. In some real applications, for example the case of real-time operational planning, there are no high time margins to realize the route planning, and it is more important to produce a near-optimal solution but quickly than to prove that the solution found is optimal. In this optic heuristics have been studied, and in the recent years powerful algorithms have been proposed for solving the basic VRP and also its variants.

The number of heuristic algorithms for solving VRP and its variants is very high, and their classification can be a difficult task due to large number of fields and descriptions needed to account for the diversity and intricacy of the different concepts involved in those algorithms, being much more difficult at high level of details. However, a classification system that concentrates on the essential ideas can be quite instructive and useful to choose the most appropriate method to a considered VRP application.

At macro-level, two main groups o heuristics can be defined: classical heuristics and metaheuristics (69) . Methods belonging to the first group perform a relatively limited exploration of the search space and typically produce good quality solutions within small computational times. In metaheuristics, the emphasis is on performing a deep exploration of the most promising regions of the solution space, combining sophisticated neighborhood search rules, memory structures and recombinations of solutions. The quality of solutions produced these methods is much higher than that obtained by classical heuristics, but the computational time increases, in some cases considerably. Moreover, the procedures are usually depending on the context and on parameters which must be finely tuned. In a sense, metaheuristics are no more than sophisticated improvement procedures (69) , which are complex and need a not easy parameter setting phase; for these reason, some metaheuristics obtain high quality solutions for the specific situations they have been developed, whereas classical heuristics are more flexible. Most commercial packages and trip planning tools used in the transport industry and the public administrations are using classical heuristics. We will present both groups, providing examples of studies and methods on the main heuristics.

**Classical Heuristics**

These heuristics can be broadly classified into three categories: Constructive heuristics, two-phase methods an improvement procedures. In general, most heuristics combine some of these methods, in general by a first solution construction and an improvement phase. Indeed, the procedures belonging to the two first categories are able to quickly find a good feasible solution, whereas the methods in the third category are used to improve the solution obtained after the precedent phases.

*Constructive heuristics*, which in general are integrated in more complex methods and used to find a good initial solution, gradually build a feasible solution while keeping and eye on the solution cost. Two main techniques are used in these methods: merging existing routes using a *savings* criterion, and gradually assigning customers to vehicle routes in function of their *insertion cost*.

One of the first and perhaps the most widely known heuristic algorithm was proposed by Clarke and Wright (22) , and it is based on the notion of savings. When two routes (0, ..., i, 0) and (0, j, ..., 0) can be feasibly merged into a single route, a distance saving $s_{ij} = c_{i0} + c_{0j} - c_{ij}$ is defined. The algorithm works in a two step mode where in the first step the savings are computed, creating $n$ routes (0,$i$,0)$\forall i \in \{1,...,n\}$. These savings are then ordered in a non-increasing fashion. In the second step the route are merged. Two versions of the savings algorithm can be considered, related to the route merging procedure used in this step. In the parallel version, given a saving $c_{ij}$, it is determined whereas there exist two routes, each of them containing respectively one arc (i, 0) and one arc (0, j), that can feasibly be merged. If so, both routes are merged into a route containing the arc (i, j). The sequential version considers in turn each route (0, i, ..., j, 0). The first saving $s_{ki}$ or $s_jl$ that can feasibly be used to merge the current route with another containing an arc (k, 0) or (0,l). The two considered routes are merged. If no feasibly merge exists, the next route is considered, repeating the same operations. This procedure is applied until no route merge is feasible. There is great variability in the numerical results reported for the savings algorithm (69) , and authors often do not mention whether the parallel or the sequential version is considered. Although this method is one of the first which were proposed, and the solution obtained is not considered to be very performant (for instances with less than 50 customers, the solutions obtained with this method are higher than the optimum in near 3%), this method is, even nowadays, one of the most used and reproduced in commercial tools, because of its semplicity of construction and modification, and the low memory space and computational times of the algorithm.

The savings algorithm has been improved computing the saving $s_{pq}$ obtained by merging routes p and q in the following way: $s_{pq} = t(S_p) + t(S_q) - t(S_p \cup S_q)$, where $S_k$ is the vertex set of route k and $t(S_k)$ is the length of an optimal TSP solution on $S_k$. A max-weight matching problem over the sets $S_k$ is solved using the $s_{pq}$ as matching weights, and the routes corresponding to optimal matches are merged maintaining feasibility. Several variants of this algorithm are possible, one of which approximates the values $t(S_k)$ instead of computing them exactly. The main versions of this approach can be found at (2; 36; 108)

The *insertion heuristics* build the routes by inserting each customer into a route. For these, they define the insertion cost, which correspond to the cost of inserting a customer k in a route between customers i and j. Some parameters are used in the definition of insertion costs, and have to be calibrated. Given a customer k and a route, each insertion cost for adding k into the route between each couple of customers i and j. The possibility which presents the lowest insertion cost is applied, and k is inserted into the route. Different methods and insertion cost definitions can be used, and the insertion can be done using sequential, parallel, o a two-phase procedure which combines a sequential procedure followed by a parallel procedure. In general, these heuristics use improvement procedures after a route is completed. For more information and the two main methods, see (69) .

*Two phase methods* are those methods where the construction of the solution is obtained by using a clustering phase and a route construction phase, where classical TSP heuristics are used. Two classes of methods are defined: *cluster-first route-second* methods and *route-first cluster-second* procedures.

There are several types of route-first cluster second methods. All of them are two-phase algorithm where the first phase builds customers which represent vehicle routes then the second phase solves a TSP on each cluster to optimize each route. These methods differ in the clustering phase. The simplest family of methods, known as elementary clustering methods, perform a single clustering of the customers set and then determines a vehicle route on each cluster by solving a TSP. The most common methods of this type are the sweep algorithm (53; 109; 110) , where each customer is represented by its polar coordinates in the clustering step, and ranked in increasing order of angle. Then, clusters are made starting by the first customer in the ranking: customers are assigned to the same vehicle until the route capacity or maximum distance length are not exceeded, then the same procedure is followed until each customer is assigned to a vehicle. Other elementary clustering methods clustering phase consists on solving combinatorial optimization problems instead of using a geometric method.The Fisher and Jaikumar Algorithm (47) , which determines 'a priori' a number of seed customers to initialize each cluster, then it solves a Generalized Assignment Problem (GAP) considering allocation costs (which are defined in the algorithm), customer weights and vehicle capacity. In Bramel and Simchi-Levi Algorithm (12) , the first phase consists on determining the seed customers by solving a capacitated location problem and the remaining customers are added to routes in the second phase, using an insertion algorithm.

Other types of cluster-first route-second methods are the Petal algorithms (48) , which consist on an extension of the sweep algorithm where a number of routes, called petals, are generated. Then, the final selection is obtained by solving a partitioning problem. Another family of methods it Truncated Branch-and-Bound (20) , which simplify a Branch-and-Bound procedure. In this procedure, all branches but one are discarded in the route selection, so the resulting tree consists on a single branch at each level. However, a limited tree could be constructed by keeping a few promising routes at each level. For a more detailed survey on these two families of methods see (69) .

The improvement algorithms are methods that need an initial solution. Starting from these solution, operating either on each vehicle route taken separately or on several routes at a time. In the first case, any improvement method for the TSP can be applied, whereas the second case procedures exploit the multi-route structure of the VRP. Most single-route procedures can be described in terms of Lin's $\lambda$-opt mechanisms (71) . In these procedures, $\lambda$ edges are removed from the tour, and the remaining segments are reconnected in all possible ways. If any profitable reconnection (the first or the best) is identified, it is implemented. This algorithm stops at a local minimum when no further improvements can be obtained. This basic procedure has been modified and developed, resulting on other improvement methods as the Or-opt or the 4-opt*. For a more detailed survey on these methods see (69) and (25) .

Most multi-route improvement methods are based on customer or edge exchange. Van Breedam (16) classified the improvement operations as string cross, string exchanges, string relocation and string mix. String cross (SC) is the operation where two strings (or chains) of vertices (customers) are exchanged by crossing two edges of two different routes. When two strings of at most k vertices are exchanged between two routes we have string exchange (SE). A string relocation takes place when a string of at most k vertices is moved from one route to another. A string mix is defined when both SE and SR are run then the best movement between the solutions of these two procedures is selected.

**Metaheuristics**

In recent years several metaheuristics have been developed for the VRP. These are general solution procedures that explore the solution space to identify good solutions and often include some of the standard route construction and improvement heuristics described above, as a part of the overall procedure. In a major departure from classical heuristic approaches, these methods allow deteriorating and even infeasible intermediary solution in the course of the search process. The best known metaheuristics developed for the VRP typically identify better local optima than classical heuristics, but their calculation times are also higher.

We can distinguish several main types of metaheuristics: improvement-based heuristics (simulated annealing, deterministic annealing, tabu search), population mechanisms (genetic algorithms, variable neighborhood search) and learning mechanisms (neural networks, ant colonies approaches). Hybrid algorithms (obtained by combining two or more of these types) or evolutionary and memetic algorithms (which use a Genetic Algorithm to individuate the areas of the solution space where a good solution can be found and then finding an improved solution in these subspaces using another algorithm, in general tabu search) can also be considered as another type of metaheuristics.

Simulated annealing (SA) methods are iterative neighborhood search procedures where, in each iteration $t$, a solution $x$ is drawn randomly in $N(x_t)$. $Iff(x) \leq f(x_t)$, then $x_{t+1}$ is set equal to $x$; otherwise: $x_{t+1}$ will be equal to $x$ with a probability $p_t$ and to $x_t + 1$ with a probability $1 - p_t$, where $p_t$ is a decreasing function of $t$ and of $f(x) - f(x_t)$. This probability is usually a function of $\Theta_t$, which denotes the temperature at iteration $t$, and has to be defined using a ruel called cooling scheme. Some of the best known SA methods are those of Robusté, Daganzo and Souleyette (91) and Alfa, Heragu and Chen (1) , which are early metaheuristics, and the Osman Simulated Annealing Algorithm (81) , which improves considerably the efficiency of the algorithm by using a better starting solution, a richer neighborhood search mechanism and a more sophisticated cooling scheme.

Deterministic annealing (DA) operates in a similar way to SA, except in the fact that a deterministic rule is used for the acceptance of a move. Two standard implementations ate threshold-accepting (40) and record-to-record travel (39) . Golden et al. (55) and Li et al. (70) applied a record-to-record travel procedure to large scale VRP instances.

Tabu search methods are iterative procedures where sequences of solutions are examinated as in simulated annealing. Given an iteration t, the next neighbor of the current solution $x_t$ is considered, and possible solutions are searched in this neighborhood. To avoid cycling, the solutions which were examinated recently are forbidden (or tabu) for a number of iterations. To alleviate time and memory requirements, it is customary to to record an attribute of tabu solutions rather than the solutions themselves. The basic tabu search mechanism can be enhanced by several computational features. The most used are diversification and intensification. Over the last 15 years, tabu search has been applied to the VRP, becoming one of the most used metaheuristics for this family of problems. Most known algorithms of this type are Osman's Tabu Search (81) , Taburoute (50) , the Granular algorithm of Toth and Vigo (104) and the Unified Tabu Search Algorithm (UTSA). The Granular Tabu Search Algorithm (GTSA) is based on the idea that removing the nodes unlikely to appear in an optimal solution could considerably reduce the neighborhood size and thus the computational time. The UTSA was developed by Cordeau et al. (26) as a unified tool to solve periodic and multi-depot VRPs, and it has been extended to other VRP variants (27; 28) , becoming one of the base tabu search algorithms used in the last years in many vehicle routing applications and in some researches related to realistic case studies. It possesses some of the features of Taburoute but it uses only one initial solution and fixed tabu durations. The last versions of UTSA are able to obtain solutions for some of the most known VRP instances which are close to the optimum (0,56%) without any increas in computational time respect to the first versions of this algorithm. For a more detailed survey on tabu search algorithms see (25; 51) .

A genetic algorithm (GA) is a randomized global search technique that solves a problem by initiating processes observed during natural evolution, and it is based on population mechanisms. A pure GA is a generic problem-solving method that uses little heuristic information about the problem domain. Basically, this method evolves a population of bitstrings (which, in the case of VRP, represent a sequence of customers), called chromosomes, where each of one represents a solution to a particular instance. This evolution takes place through the application of operators that reproduce natural phenomena observed in nature. In general, the algorithm is an iterative procedure which works as follows: starting from a randomly generated initial population of chromosomes $X^t$, at each iteration the first three steps are repeated k times, then the fourth step is applied:

- Reproduction step: two parent chromosomes are selected.
- Recombination: two offspring chromosomes are obtained from the parents using a crossover operator.
- Mutation: A random mutation is applied to each offspring, with a small probability.
- Generation replacement: A new population of chromosomes is created by removing the 2k worst solutions in $X^t$ and replaced by the 2k new offspring.

In these algorithms, two parameters are defined: $T$, the number of generations and $k$, the number of selections per generation. The best solution produced over the $T$ generations is the final result of this algorithm. Pure genetic algorithms are not always performant, although some methods have been developed with good results (89; 95) . Other population-based approaches are the Hybrid and Evolutionary Algorithms (83; 86; 73) , which combine local search with population search methods, like genetic algorithms. Hybrid algorithms are obtained by a combination of a GA and a local search procedure. An evolutionary algorithm (EA) is a subset of evolutionary computation, a generic population-based metaheuristic optimization algorithm. An EA uses some mechanisms inspired by biological evolution: reproduction, mutation, recombination, and selection.

A memetic algorithm (MA) is another method which derives from GA, and it uses a population of agents to search the best solution for a problem, using a given fitness function which ranks the goodness of the solutions. The agents examinate candidate solutions throughout the search space, using knowledge about the problem to improve the solutions, and cooperating and competing among themselves. Cooperation means that cooperating agents give rise to new agents which share characteristics from them, while competition is achieved by selection pressure over the population of agents. Although this description seems very similar to conventional GA, MA take a qualitatively different approach. In general, most MA are defined by hybridizing a GA with a local search (LS) technique, using LS as an additional step within the GA. In many cases, MA are briefly resumed as the equation MA = GA + LS. This family of algorithms can be considered as a part of the hybrid algorithms, but not all the hybrid methods are memetic. An example of this method applied to VRP is the algorithm proposed by Moscato and Cotta (76) .

Other techniques which are sometimes used in VRP are the learning mechanisms, such as Ant colonies algorithms and neural networks. Ant systems methods are inspired for analogy with real ant colonies foraging for food, which mark their path to food by an aromatic substance called pheromone. Ant systems methods reproduce artificially these phenomena, associating objective functions with the quality of food sources and recording the values in an adaptative memory which represents the pheromone trails. Neural networks are computational models composed of units that are richly interconnected through weighted connections, like neurons in the human brain. These units, called neurons, can learn from experience in order to produce better solutions. We will not expand on these methods; for more information on both ant colonies and neural networks see (51) .

### 3.2.4 Main VRP variants

In real world applications, because of the diversity and complexity of the systems and situations that can be found, the basic CVRP represents only a few number of cases, sometimes in a simplified way. However, it becomes the basic version from which several variants can be obtained. In function on the situation and the important parameters that define the system, one or more groups or constraints can be added or modified to represent the different. In other cases, the objective will be also modified. In this section we will present the most known variants, classifying them into groups and citing the main studies which can be found in literature.

Other variant arise when heterogeneity is considered. The most common problem with heterogeneity is obtained when the vehicles have different capacities. In this case, a set of $m$ vehicles is considered, each of them with a capacity $K_l; l = 1...m$. When the maximum distance that each vehicle can cover is considered a limitation, the problem is known as Distance Constrained VRP (DVRP), while when both the groups of constraints are considered, the problem is named Distance Constrained Capacitated VRP (DCVRP). This distance can be expressed in terms of geographical distance or travel time, being this second representation the most studied in this variant. Moreover each vehicle may be associated with a different maximum travel time $T_l; l = 1,...,m$ (see Laporte et al. (68) ).

Other variants can be obtained for network features reasons, or when the service offered is different from the basic problem. As shown in figure 3.1, the main network features are related to depots and intermediary facilities, whereas main service features are related to service quality in terms of delivery time at customers (VRPTW), and to the possibility of realize deliveries in the same vehicle route without passing through the depot (VRPPD).

In the case where a less-than-truckload policy with vehicle trips serving several customers is applied only at the second level, the problem is close to a multi-depot VRP. However, since the most critical decisions are related to which satellites will be used and in assigning each customer to a satellite, more pertinent methods will be found in multi-depot Location Routing Problems

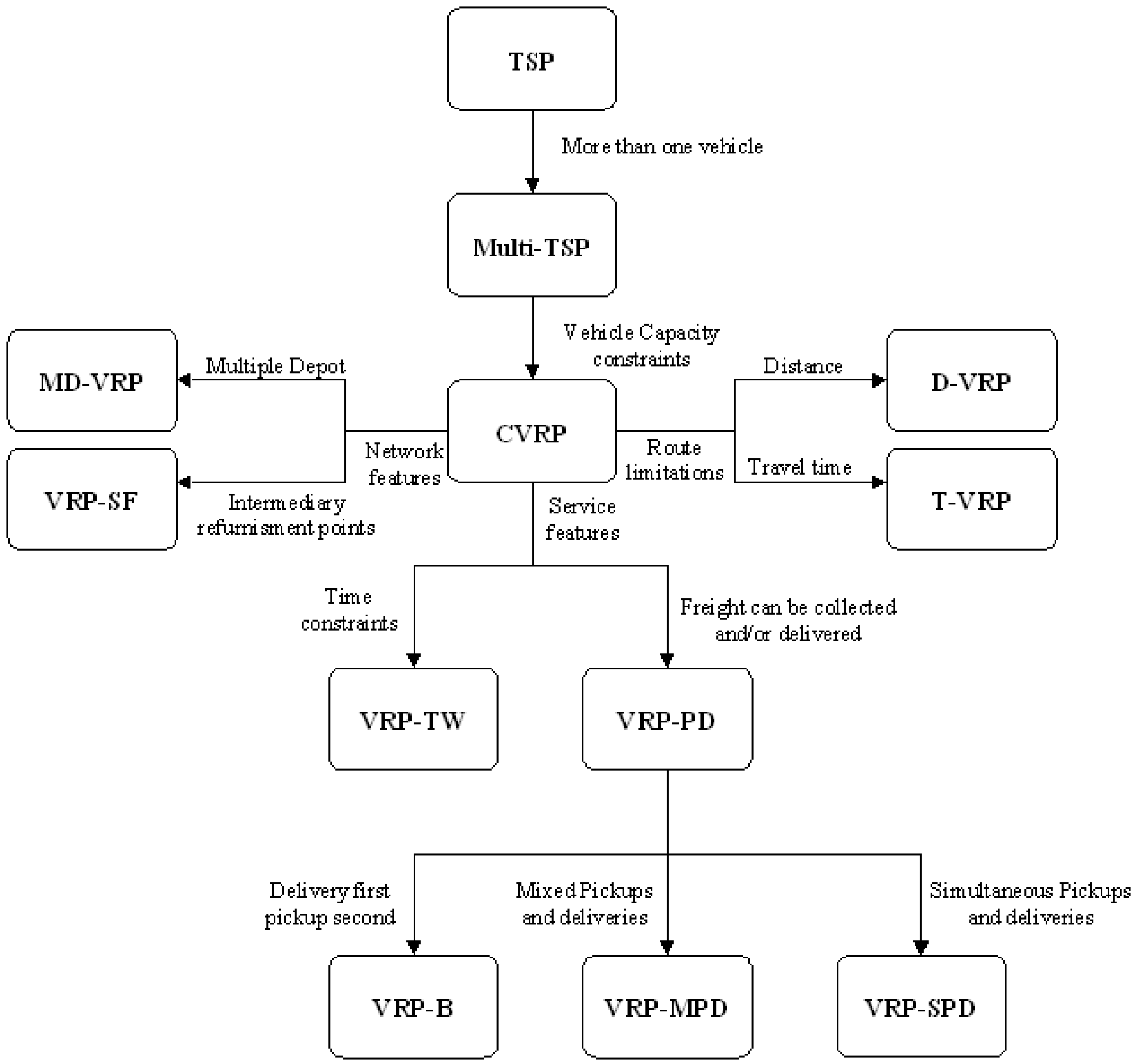


Figure 3.1. Main VRP variants and their relation to the basic version

(LRP). In these problems, the location of the distribution centers and the routing problem are not solved as two separate problems but are considered as a more complex problem (65) . For a more detailed survey of LRP, see (74) and (78) . Although many of these studies refer to direct shipping strategies (i.e., single echelon), some heuristics have been developed for the multi-echelon problem (62) .

As presented in chapter 2, satellites are intermediate facilities where several operations can be made. We note that one variant of VRP considers satellites facilities. In this variant, known as VRP with Satellites facilities (VRPSF), the network includes facilities that are used to replenish vehicles during a route. When possible, satellite replenishment allows the drivers to continue the deliveries without necessarily returning to the central depot. This situation arises primarily in the distribution of fuels and some other retail applications; the satellites are not used as depots to reduce the transportation costs (32; 4; 7) . This is not a multi-echelon transportation, but is more close to multi-depot variants.

Time constraints, which are important in some real applications (e.g. express courier carriers, postal services, newspaper distribution, e-commerce, etc), have been studied on VRP-TW problems. A Time window is defined as the time interval inside which a vehicle can arrive to a destination to satisfy a request (24) . Two types of time window constraints can be defined:

- Hard time windows, which are defined as a strict constraint, in which there is no possibility for a vehicle to arrive to destination after the upper time limit. It is usually also impossible to arrive to destination before the lower time limit, but in some cases this possibility is considered, allowing the possibility of stopping the vehicle at destination until the lower time limit is reached.

- Soft time windows, which are defined in the objective function, and represented by an increasing cost penalty if the vehicle arrives to destination outside the time window interval. This representation of the time windows can be applied to many real applications, in which the request can be satisfied, in some conditions, even if time constraint is not strictly respected.

For a detailed survey on this class of problems see (24; 14; 15) .

Another important group of problems is defined when customers are not only receiving freight, but some quantity of goods must be also collected there. This group of problems presents many cases, because of the different possibility to consider the pickups and the deliveries in route organization. Savelsberg and Sol formulated in (93) a general problem for transportation services with pickups and deliveries: the General Pickup and Delivery Problem (GPDP). In the GPDP a number of routes has to be constructed to satisfy transportation requests. A fleet of vehicles is available to operate the routes. Each vehicle has a given capacity, a start location and an end location. Each transportation request specifies the quantity of load to be transported, the locations of pick up and the locations of delivery. Each load is transported from its set of origins to its set of destination without transshipment. Three groups of problems are special cases of the GPDP, and are:

- The pickup and delivery problem (PDP). In this problem, each transportation request specifies a single origin and a single destination and all vehicles depart from and return to a central depot.
- The dial-a-ride problem (DARP) is a particular case of the PDP applied to on demand collective passenger transportation service. In this case, the requests are usually passengers instead of freight quantities (and vehicle capacities are expressed in number o passengers), and time factor is usually important. The problem presents additional constraints representing the insertion a new transportation request on a vehicle when it arrives, which often happens in a dynamic and real time environment.
- The vehicle routing problem with Pickups and Deliveries (VRPPD) is a PDP in which either all the origins of the deliveries and the destinations of the pickups are located at the depot.

If we consider VRP with Pickups and Deliveries, different policies can be applied to manage pickup and delivery requests (77) . The most commonly used strategy is to schedule the deliveries first, and after the last delivery request is satisfied, to proceed with the pickup requests. This strategy is known as Delivery First, Pickup Second. We can also distinguish between the cases where deliveries and pickups are scheduled and optimized as two separate activities, which can be modeled as two independent VRP, and the cases in which the vehicle goes directly from the last destination to which freight was delivered to the first destination on which freight have to been picked-up. In this second group of cases, the vehicle, after delivering all the freight, have to go to customers having freight to be collected and then it will return to the depot. The decision problems that model these cases are known as VRP with backhauls (VRP-B). Another strategy is to combine pickups and deliveries, and the vehicle, after delivering some requests, if there is place to pickup freight, can go to a destination and pickup other requests. This case is known as VRP Mixed Pickups and Deliveries (VRP-MPD). The last case corresponds to the situation in which at a destination the corresponding freight is delivered, then, in the same vehicle, other freight is picked up from the same destination. This strategy, often used in newspaper distribution, and in some fresh products, is known as Simultaneous Pickups and Deliveries (and the decision problem associated to it is noted as VRP-SPD.

The other strategy is to allow each demand to split. In this case, visits to each customer are allowed. This problem is called VRP with Split Delivery (VRPSD)(see Dror et al. (38) ).

Many variants are also obtained by combining two or more groups of constraints belonging to the problems presented above, for example time constraints, multi-depot and/or satellite facilities, etc.

## 3.3 Location Routing Problems

Other transportation cost optimization problems which are close to VRP and present multi-echelon versions are Location Routing Problems (LRP). In its most general form, the LRP is defined as follows. Given a network G composed by two types of nodes: facilities (where freight is allocated or temporarily stored, they have similar characteristics to VRP depots but are in general capacitated) and customers (defined in the same way as in classical VRP variants by their position and their requested demand) and one or more fleets of vehicles, each of them defined by its capacity. The costs of this transportation system are associated both to vehicle routes (travel costs) and to facilities (allocation, activation and facility usage costs). The LRP seeks to minimize total cost by simultaneously selecting a subset of candidate facilities and constructing a set of delivery routes that satisfy a number of constraints. The basic ones, analogously to basic VRP, are the following:

- customer demands are satisfied without exceeding vehicle capacities;
- facility capacities are not exceeded;
- each route begins and ends at the same facility.
- the number of active vehicles do not exceed the specified limits;

In these problems, the location of the distribution centers and the routing problem are not solved as two separate problems but are considered as a more complex problem (65) . Analogously to VRP, many variants are defined for these problems. Various classification schemes are available in literature to categorize either pure VRP or LRP problems. The most complete of them is provided by Min et al. (74) . This classification defines the main aspects related to the problem perspective. These aspects are:

- Hierarchical level: as already described in chapter 2, two main strategies are defined. Single stage problems represent the direct shipping strategy, and multiple stages problems deal with multi-echelon distribution strategies.

- Nature of demand/supply: in general, demand requests are made in advance, so the freight quantities are determined before the transportation system is optimized. In these cases, the decision problems are deterministic. However, in some real cases and for some freight categories, customers are defining the freight quantities of their request at the time of the arrival of the supplier's vehicle. In this case, decision problems are based on statistics and uncertainty modeling, and are noted as stochastic LRPs.

- Number of facilities: in general, most of LRP variants deal with multiple facilities. However, in some specific applications, single facility problems can be also defined.

- Vehicle fleet characteristics: as explained in section 3.2.1, the overall vehicle fleet can be homogeneous (all the vehicles have the same characteristics) or heterogeneous (each vehicle has characteristics). A particular version of the heterogeneous vehicle fleet is the multiple vehicle fleet LRP, where a number of homogeneous vehicle fleets, each of them characterized by a different capacity, are defined. This capacity can also be limited (capacitated vehicles) or unlimited (uncapacitated vehicles).

- Facility characteristics: in the same way as for vehicle characteristics, the facilities can be capacitated or uncapacitated. In general, facility capacities are heterogeneous.

- Planning horizon: in general, two planning horizon strategies are used in real applications. Single period problems represent the cases the distribution planning is made for one single specific configuration of requests (e.g. trip planning for a single day). If this configuration is defined not for a single moment but for a period of time (e.g. weekly planning where each day has a different request configuration).

- Time windows: as defined in section 3.2.4, time constraints of TW type can be considered or not. If they are not considered, no deadlines are defined. In the case they are considered they can be soft TW, which represent loose deadlines (when one of them is not respected, costs related to the freight delivery to the considered customer increase) or hard TW, which represent strict deadlines that must be strictly respected.

- Objective function: different objectives can be considered in the function to optimize (see section 3.2.1). Variants of LRP can be single objective, when only one aspect is considered in the objective function, or multiple objective, where more than one aspects are combined in the function to optimize.

Most of the LRP deal with single stage, multiple facility LRP. Detailed survey of works on single echelon LRP have been made by Min et al. (74) and Nagy and Sahli (78) . Some multiple facilities LRP consider a central pre-located depot to which all facilities to be located are connected, but no tour planning from the depot to the rest of facilities is involved. Although in both works some multiple stage LRP are described, the small number of multi-echelon LRP studies and the fact they are recent and they are sometimes presented using different terminologies. We will use the term multi-echelon to refer to multiple stage problems, and the VRP terminology for a better understanding of the following studies. We will focus on multi-echelon LRP, although a brief survey on main single-echelon LRP will also be presented.

Exact methods have been developed for a small number of LRP models that are derived from two-index flow formulations for the vehicle routing problem (VRP). We will introduce the main studies on exact methods for multiple facilities LRP, which can be extended to multi-echelon transportation. Laporte et al. (67) solve a multiple facilities problem in which at most p facilities are located by adapting Miliotis REVERSE algorithm for the TSP. The largest problems solved have seven candidate facilities and 40 customers. Laporte et al. (79) solve a multiple facilities capacitated LRP using a constraint relaxation method. In their work, the largest problem solved to optimality has eight candidate facilities and 20 customers.

In the field of the studies about heuristic methods for the single-echelon multiple facilities LRP, Wu et al. (111) consider a multiple facilities location-routing problem with a heterogeneous vehicle fleet in which the number of available vehicles is limited. The problem is decomposed into a location-allocation problem and a vehicle routing problem, both to be solved using a simulated annealing metaheuristic in which a tabu list is used to prevent cycling. Tests on problem instances up to 150 nodes show that this approach outperforms previous approaches on traditional multiple facilities location routing problems with a heterogeneous fleet and unlimited number of vehicles.

In general, two-echelon LRP follow the concepts defined by Jacobsen and Madsen (62) . The problem consists of determining the location of intermediary facilities (considering that the starting depot is already determined, in the case of single starting depot cases), allocating the customers to transfer points and designing both fist-echelon and second-echelon routes (called by Jacobsen and Madsen respectively primary and secondary tours). The two-echelon LRP is an extension of the multiple facility LRP, as already said.

The first application of a two-echelon distribution system with the minimization of the total transportation cost as objective function can be found in (62) . This problem was known firstly as *two-level LRP*, but generalizing the term two-echelon is used to distinguish these problems for bi-level (or multi-level) programming in combinatorial optimization. In this study, a comparison of several fast heuristics for solving a real case application where two newspaper editors combine their resources in terms of printing and distribution in order to decrease the overall costs. Newspapers are delivered from the factory to transfer points, which must be chosen from a set of possible facilities, and then other vehicles distribute them from these transfer points to customers. The authors propose three fast heuristics and compare them. The first is the Three Tour Heuristic, which is based on the observation that if the last arc of each route, the problem becomes similar to a Steiner Tree Problem. This tree is constructed by a greedy one-arc-at-a-time procedure. The other two heuristics, which are sequential, combine heuristics for both VRP and Location-Allocation problem. The ALA-SAV heuristic is a three stage procedure composed from the Alternate Location Allocation (ALA) of Rapp and Cooper (88) and the Savings algorithm (SAV) of Clarke and Wright (22) . The third heuristic (SAV-DROP) is also a three stage procedure composed from the Clarke and Wright Savings algorithm and the DROP method of Feldman et al. (45) .

Another real application for a two-echelon distribution system was proposed by several authors for raw milk collection at farmyards (18; 106; 94) . The milk is collected at storage tanks on the farmyards every or every other day and must be transported to dairy plants. Some of the farmyards cannot be visited by a lorry-trailer combination because of space restrictions. The first two studies allow only one transshipment location per trailer and a fixed lorry-trailer assignment. In both cases a heuristic sequential solution approach is adopted. This procedure is a 3-stage procedure, where in a first phase customers are grouped into clusters; then, the determination of one transshipment location per trailer takes place; the third and last stage is the routing phase.

The *road-train routing problem*, introduced by Semet and Taillard (98) , also known as *truck-and-trailer routing problem* (TTRP). This problem concerns defining a route for a road-train, which is a vehicle composed by a truck and a trailer (both with space for freight loading). Some of the roads are not accessible by the entire convoy, but only by the truck. In these cases, the trailer is detached and left at a customer's location (called a "root") while the truck visits a subset of customers, returning to pickup the trailer. In a way, this problem can be represented as a two-echelon distribution system, using the LRP notation. The intermediary facilities become the customers where the trailer is parked while the truck visits a group of customers. The difference from Jacobsen et al. (62) is that is that in this case some customers can be served directly by the primary tour. A two-stage algorithm is proposed by the authors. An initial solution is obtained by a sequential algorithm and improved by tabu search, where customers are reallocated. This method do not distinguish between locational and routing moves.

We find in literature several heuristics applied to the same problem proposed by Semet and Taillard. Semet (97) proposed a clustering first routing second solution method. First customers are allocated to roots then the resulting routing problems are solved via Lagrangian relaxation. Gerdessen (52) assumes that all customers have unit demand and each trailer is parked exactly once. Initial solutions are found using a number of sequential heuristics. These are then improved by a selection of VRP improvement heuristics. Chao (19) developed a two-stage algorithm where in the first phase an initial solution is obtained with a cluster first route second heuristic and the second phase improves the initial solution using a tabu search algorithm with customer reallocation moves.

Scheuerer (94) presents two new construction heuristics: a clustering-based sequential insertion procedure and an adaptation of the well-known sweep algorithm by Gillett and Miller (53) and a tabu search improvement procedure. Moreover, the author adapts these procedures to the multi-depot and the multiperiod version of the problem. Hoff and Lokketangen (60) have presented a case study for milk collection in Norway. The problem they consider is essentially a multi-depot, multi-period TTRP with heterogeneous vehicles and without trailer customers. They propose a sophisticated tabu search algorithm for solving their problem and report successful solution of real-world instances, improving on the existing tour plans used by their industry partner.

The most complex and general multi-echelon LRP is defined by Ambrosino and Scutella (3) . Although the general purpose of the paper is to present general model for multi-echelon network design problems which represent real network planning cases, a multi-echelon LRP can derive from the general formulation.

# Chapter 4

# The Multi-Echelon Vehicle Routing Problem

Multi-Echelon distribution systems optimization is not developed considering the overall system transportation costs. In general, in order to simplify, vehicle routing problems are developed for the single-echelon strategy, and in multi-echelon systems transportation is approximated in all levels or only one of them is considered and the rest are approximated. In this chapter we will consider a multi-echelon distribution system, whose physical interpretation is presented in 2.2. In this chapter we will present a new family of combinatorial optimization problems which will focus on optimizing the overall transportation cost considereing the entire system in the optimization process, without simplifying into a sum of independent smaller sub-problems.

In section 4.1 we will present the N-Echelon VRP, the basic vehicle routing optimization problem for a N-Echelon distribution system. In this section, the main concepts and notation for the new family of problems, represented by the N-E VRP, will be defined. Then, in section 4.2 a brief description of the basic NE-VRP variants, based on the classification presented in section 3.2.1 will be proposed.

## 4.1 N-Echelon VRP: Basic definitions ans problem description

The Multi-echelon VRP can be defined in different ways and can represent different real multi-echelon distribution systems. In this section we will present the basic problem, which will be noted as N-Echelon VRP.

Consider a N-Echelon distribution system, with a single origin, called depot, and a number of final destination, called customers. Using combinatorial optimization terminology, and concretely VRP notation, we define a graph G which represents the transportation network. The nodes of graph G are divided into N+1 different groups.

The first group is composed by the origin of the transportation system. This depot is associated to a vertex of the road graph, where the depot is located. Other characteristics of the depot, which in general are not specified, are its capacity, in terms of freight quantity which is available at the depot, and eventual fixes costs. The available vehicles can be also associated to the depot, but this aspect will be deeply developed when defining the vehicle fleets. Analogously to the single-echelon well known VRP basic problem, we consider one depot. This depot is considered as a level 0 node.

A second group of nodes, which defines the $N^t h$ level, contains the final destinations where freight must be delivered or collected. These nodes are known as customers, which typical characteristics are:

- vertex of the road graph where the customer is located;
- freight quantity, known as demand, which must be collected or delivered at the customer;
- periods of the day during which the customer can be served, which are known as time windows;

- times required to deliver or collect the freight at the customer location, which correspond to loading and unloading operations, and in some cases, some administrative operations that have to be realized at the moment of the freight's delivery or pickup;

- subset of vehicles which can be used to serve the customer, due to its characteristics (dimension or weight), the customer's characteristics (distance, accesses and parking areas) or the network configuration (road characteristics, temporal limitations, or other events which can limit the access of some vehicles).

Then, N-2 groups of intermediary nodes are defined. For each level k, a group of nodes known as k-level satellites, represent the different facilities where freight which has to be delivered to a customer arrives for intermediary operations. Typical characteristics of each k-level satellite are the vertex of the road graph where the k-level satellite is located, eventual costs for loading and unloading operations, satellite activation or freight waiting penalties, and, analogously to the customers, the subset of vehicles which can be used to serve it.

Arcs connect these nodes in the following way. The depot is connected to 1st-level satellites, and these satellites between them, by 1st-level arcs. Then, each 1st-level satellite is connected to 2nd-level satellites, which are also connected between them, by 2nd-level arcs. Generalizing, for each level $2 \leq k < N$, a set of k-level arcs connect (k-1)th-level satellites to k-level satellites, and also each k-level satellite with the other satellites belonging to the same level. Finally, (N-1)th level satellites are connected with customers by N-level arcs, which also connect customers between them. To each arc connecting node $i$ to node $j$, a travel cost $c_ij$ is associated. As for classical VRP problems, these costs can be symmetric ($c_ij = c_ji$) or asymmetric.

N-1 independent fleets of vehicles are defined, each of them associated to a level. The delivery from the depot to the customers is managed by rerouting and consolidating the freight through different intermediate satellites. The general goal of the process is to ensure an efficient and low-cost operation of the system, while the demand is delivered on time and the total cost of the traffic on the overall transportation network is minimized. For each level fleet, its size and vehicles' composition can be fixed and defined according to the requirements of the satellites and/or customers, and to the network characteristics. Each vehicle is defined by a number of characteristics, which the most significant can be:

- Home depot of the vehicle, which is the place where the vehicle is parked when not active and its global service's starting and ending point. Note that not always this information is important, and in some cases costs related to vehicle travels from its home depot to the starting point of the route and the returning costs to this depot are not considered in the transportation costs optimization.
- Level and starting point of a route, which is a depot or a satellite. Each vehicle is assigned to a level $k \geq 2$, and will operate only between a (k-1)-level node (which can be fixed or variable) and k-level nodes. We can also find some real applications where vehicles are not associated to a level but to each satellite.
- Capacity of the vehicle, which is defined as the maximum freight quantity which can be loaded into the vehicle. Usual capacity measures are weight and volume, but in some cases it can also be expressed in number of pallets or number of containers.
- Possible subdivision of the vehicle into compartments, each of them characterized by its capacity and by the types of goods that can be carried.
- Devices available for the loading and unloading operations.
- Costs associated with the vehicle utilization, which can be fixed or variable, and correspond to the different costs associated with the utilization of the vehicle. The fixed costs are known as activation costs, and are counted once each time a vehicle enters in service. The variable costs are expressed related to another measure unit (per time unit, per distance unit, per route, per level, etc.).

Sometimes, it is not possible to serve each customer, due to different limitations (for instance vehicle or crew availability, quality standards, incompatibility with availability times for customers, road restrictions, overbooking, bad planning and other unexpected events). In these cases, not all the customers will receive all the amount of freight they requested (or not all the freight will be collected from each customer), in order to visit all customers serving them the best considering the unexpected limitations. If not, another possibility is to not serve a subset of customers. To deal with these situations, different priorities, or penalties associated with the partial or total lack of service, can be assigned to each customer.

One of the main characteristic of VRP variants is that each route performed to serve customers start and end at the same depot. In ME-VRP, considering a level $k \geq 2$, each k-level route starts and ends at the same point belonging to level k-1, and serves only k-level nodes, to return to its origin. For each level k, each (k-1)-level node is characterized by the number and types of vehicles associated with it and by the global freight quantity it can deal with. In some real applications, the customers are a priori partitioned among the depots, and the vehicles have to return to their starting point at the end of each route. In these cases, the overall NE-VRP can be decomposed into several independent problems. If this policy is applied to each level, the overall problem is decomposed into a number of single-echelon problem, i.e. classical VRP problems, for which many techniques have already been studied. In these cases, a system approach is not necessary to optimize them. For these reasons, these easily decomposable problems will not enter in the objectives of this work.

From these aspects, we can define the basic N-Echelon Vehicle Routing Problem. Considering the N-echelon graph defined above, it can be decomposed into $k \geq 2$ levels:

- First level is formed by , which connects the depots to the 1st-level satellites;
- $k - 2$ intermediate levels interconnecting the satellites;
- the last level, where the freight is delivered from the satellites to the customers.

To this graph, a N-echelon transportation system is associated. We can make the following considerations. N-1 fleets of vehicles (one per level) are defined. Each vehicle fleet is associated to a level, and have specific characteristics. Each vehicle belonging to this fleet brings freight from a k-1 level node to a subset of K level nodes. Travel times, and then the costs derived of vehicle travels, are static and a priori determined.

Demand associated to customers is deterministic, and known in advance. We consider that the depot is uncapacitated and all the freight required by customers is available. Other aspects which define the satellites is that inventory and stochastic aspects are considered for satellite demands. These demands are however not known in advance, and constitute one of the variable aspects to consider in optimization. No long period storage is allowed at satellites. Freight can remain at each satellite for some time, but not long period storage and inventory aspects are considered. This problem represents many distribution applications where all the operations at the intermediary points are realized in one day, and at the closing times no freight remains there. The main constraint associated to a k level satellites is that freight that arrives to a satellite have to be loaded into k+1 level vehicles.

The most typical objectives which can be considered in a NE-VRP are the following:

- to minimize the global transportation cost, which depends on the global distance traveled and on the fixed costs associated with the used depots, satellites and vehicles. This cost can be also called system transportation costs, noting that the objective is to minimize the cost considering the entire transportation system, and not the sum of each level minimum cost;
- to minimize the number of vehicles (or drivers) required to realize all the distribution operations, which is usually considered secondary respect to other objectives, but in some applications it can be important to consider also this aspect; in some real cases, intermediary level will follow different policies and sometimes the minimization of the number of vehicles will be possible only in some levels, and not for the entire system, due to different reasons;
- to minimize client inconvenience, which are defined by the means of penalties associated by the partial or the no respect of customer service quality standards;
- to maximize the profit, considering costs and revenues associated with freight transportation. Profit is then defined as the difference between revenues and costs.

These objectives can be considered separately (using only one of them as an optimization objective), or as a objective which includes two or more of them (by the means of a weighted combination of them, or, in some cases, as a multi-objective or multilevel optimization).

The next step should be to build a decision model which represent this problem. The first step to build this model is to define the system transportation cost optimization, not dividing the problem into different independent sub-problems but considering the entire system defined above, defining the objective and the transportation costs. Then, in a second time, the advances in multi-echelon inventory should be integrated. Due to the complexity of this problem, and to the fact many real applications are based on two-echelon or three-echelon distribution systems, we will not present a general model for the NE-VRP. We will focus instead on the simplest NE-VRP version, the Two-Echelon Vehicle Routing Problem, which will be defined in next chapter. From this model, an easy extrapolation can be made to three or more echelon problems, in function of the application needs and constraints.

## 4.2 Basic variants of the N-Echelon VRP

In the precedent section we defined a general basic version of the problem, without specifying the different constraints due to capacity and total travel distances or times for a route, which are considered in the basic versions of VRP problems. In this section we will define two groups of basic problems, the basic NE-VRP problems, which are the uncapacitated and the capacitated versions, and the basic NE-VRP variants, following the same classification presented in 3.2.1.

### 4.2.1 Basic multi-echelon route optimization problems

From the definitions presented above, different basic problems can be enounced based on the vehicles and satellites characteristics. The complexity of these different problems is higher than classical VRP complexity, due to the different nature of these two types of capacity. In classical VRP, the simplest variant is the uncapacitated version of the problem, which is equivalent to a multi-TSP, term which is usually adopted to name the problem. However, different problems can be derived from the generalization of multi-TSP to N-echelon distribution systems.

The first of them is the nested TSP. The N-nested TSP (N-NTSP) is the simplest version of the multi-echelon route optimization problem. To the depot and to each k-level satellite is assigned only one vehicle of unlimited capacity. The objective of this variant is to find the best combination of intermediary levels which total transportation cost is the minimum. The main aspects of this problem are no only to find the minimum transportation costs (the combination of arcs which give the lowest possible overall cost) but also to assign each customer to a combination of satellites (one per level), which will represent the subset of active satellites. In this version of the problem, satellites can be uncapacitated (US N-NTSP)or capacitated (CS N-NTSP). Similar to it, the Uncapacitated NE-VRP (NE-UVRP), or NE-multi TSP which is the version of the problem presented. In this version of the problem, only the network constraints described at precedent section are considered, whereas vehicle ans satellite capacity are considered unlimited. Unlikely to classical TSP and VRP problems, N-NTSP and NE-UVRP are less common, but a formulation derived from this problem can be used, for example, in some air or sea transport cases (also for passengers), where the "‘hub and spoke"’ strategy is common, if the capacity constraints are not very restictive.

The capacitated version of the problem NE-CVRP is the variant where some vehicles and/or satellites are capacitated. The most common versions are those where at least vehicles of one level are capacitated. The vehicle capacity, which is defined in the same way as for classical VRP variants, is defined as the maximum freight quantity that can be loaded into the considered vehicle, in terms of volume or weight. The satellite capacity represents the maximum freight quantity which can be temporarily stored in a satellite. Note that, in the basic versions of NE-VRP, inventory and long storage for freight is not allowed, to focus on the transportation system. The satellite capacity represents then the space which is available in a satellite to realize the different loading, unloading and trans-dock operations, and eventually other additional services (as control, package, etc, which will not be considered in a first time, but can be easily taken into account in satellite capacities and operational costs). Satellite capacity can be represented in more different ways. As well as classical representations in terms of maximum freight volume or wight, we can use as a k-level satellite capacity limitation the number of k-level or k+1 level vehicles (the most restrictive of both). In NE-CVRP, it can be possible that not all the levels present capacity constraints. The problems which have at least one of these constraints, even if it's present in one specific level of the whole system, are considered NE-CVRP cases, due to the connexions between levels, which can implicitly limit the usage of uncapacitated vehicles to assure the global function of the system.

The two last families of route optimization problems can be considered as the basic N-echelon VRP (Uncapacitated and Capacitated). Capacity limitations are one of the main aspects in multi-echelon distribution, and, due to the dependence between different levels, can become more restrictive in multi-echelon problems than in single-echelon distribution. Most of the multi-echelon distribution applications presented in precedent chapters are subject to these limitations. For these reasons, the concepts, methods and variants presented in next section, and also in the rest of the work, will mainly refer to capacitated problems, even if the concepts which will be described can also be applied to the other versions presented above.

### 4.2.2 Basic NE-VRP variants

In this section we will define briefly the main NE-VRP variants which can be considered, following the same classification used in chapter 3.2.1. In this classification, we consider three main aspects: network and service features, and route limitations.

Route limitations are applied to one or more routes, on one or more levels. Two types of limitations can be considered, and are distance and time constraints. A N-echelon distance constrained VRP is a variant of the basic problem of the same family where one or more k-levels present maximum distance limits. These limits are expressed in terms of maximum distance that vehicles can travel, and they will be related to the vehicle's characteristics. This distance can be explained in terms or travel distance (in km), or in terms of travel time. In this second group of constraints, different factors like, times related to loading and unloading operations, and slack pauses can be considered, and represent the maximum time a vehicle can be on service, for maintenance, crew working hours and other reasons.

In network features variants, the main important problem should be the N-Echelon multi-depot VRP. This problem, analogously to classical VRP, present more than one depot, so the starting point of each 1st-level route can be different to the others. In these problems, two main policies are considered. The first is that the freight type is the same for all customers and all depots have an enough quantity of freight to serve all customers. Analogously to MDVRP, in ME-MDVRP, freight requested can be assigned to one of the available depots. Additional constraints can be added to the depot availability, as for instance the depot capacity, time period for service (defined by the opening and closing hours which correspond to the limits beyond which it is not possible to arrive to the depot), but in all cases it is supposed that all customers can be assigned to all depots. A possible method to solve this problem can be to transform it into a N+1 echelon VRP where all the depots are transformed into 1st-level satellites and a new depot is added. The new 1st-level costs are considered as equal to 0. In some cases, the complexity added using this approach, will be negligible but for problems with few levels an approach based on multidepot VRP should be more effective. A second possibility is to assign 'a priori' each customer to each depot. In this case, the NE-MDVRP is not equivalent to a number of independent NE-VRP, because the freight at one k-level satellite can be merged into the same k+1 route.

Service features variants refer to some aspects which the distribution service company offers in the transportation service. Two main families of variants are presented, analogously to classical VRP. The first of, and maybe the most important, due to time limitations, is the ME-VRP with Time Constraints. Several types of time constraints, which represent different temporal aspects of multi-echelon transport organization, can be considered. We will describe those which can be observed in most real applications where time constitutes one of the main factors describing the proposed service features.

The most common time limitation, analogously to classical VRP and distribution problems, is the time interval in which the vehicle can visit a facility. This interval is known as *Time Window* (TW), and is defined by an early arrival time of the vehicle (EAT), which corresponds to the moment (in time) from which the facility can receive the service, and the last arrival time of the vehicle, which is the moment after which it's not possible to arrive to the facility. The problem is called NE-CVRP with time windows (NE-VRPTW), and the TW are associated to nodes (usually, time limitations are not directly associated to arcs, but to customers or k-level satellites, even if TW can also be associated to the depot). When the TW are associated only to customers, only the last level follows VRPTW logics whereas for the other levels time constraints will not influence it directly, but indirectly assuring that the freight arrives on time to N-1 level satellites. When the TW are associated to satellites, the complexity of the problem increases, because of restriction due to connexion constraints between levels.

Other time constraints, which are more restrictive, are vehicle synchronization at satellites. In some real applications, satellites are not projected to store freight even for a few time interval, and vehicles cannot wait for a long time at satellites, waiting to be loaded or unloaded. We can formulate a problem which represent these cases, which can be noted as N-Echelon Capacitated VRP with Satellites Synchronization (2E-CVRP-SS). In this problem, time constraints on the arrival and the departure of the vehicles at the satellites are considered. In fact, the k-1 level vehicles arriving in a satellite unload their freight, which must be immediately loaded into a k-level vehicle. These constraints can be of two types: hard and soft. In general, a small time interval, called synchronization margin $T^s$, is defined. In hard SS, every time a k-1 level vehicle unloads its freight, k-level vehicles must be ready to deliver it. This is represented as follows: k-1 level vehicles cannot wait more than $T^s$, and this is expressed by a very restrictive pseudo-TW, which does not have a fixed EAT but, when a k-level vehicle arrives at a k-level satellite at a time $t$, the corresponding complementary k-1 level vehicles must arrive at most at time $t' = t + T^s$, and vice versa. In soft SS, when k-1 level vehicles arrive, if k-level vehicles are not available, the demand is lost and a penalty is paid.

A more complex version which derives from Multi-depot NE-VRP but consider feature services which are different from time constraints is Multi-depot multi-request NE-VRP (MD-MR NE-VRP). This problem is only considered if freight can be merged at satellites. In this case, given a k-level satellite, the freight coming from k-1 level routes assigned to different depots can be merged or reorganized to put on the same k-level vehicle freight with different origin depot and having to be delivered to the same customer. The main difficulty of this variant arises in the fact of selecting the k-level satellites to merge the freight which allow to minimize the overall costs.

Another service feature policy represents services with Pickup and Deliveries (2E-CVRP-PD). Pickup and deliveries, in the three modalities presented in chapter 3 (Backhauls, Mixed PD and Simultaneous PD). In this case we can consider the satellites as intermediate depots to store both the freight that has been picked-up from or must be delivered to the customers.

A particular case of NE-VRP is obtained when considering a transportation system where taxi services are considered (NE-VRP-TS). In this variant, direct shipping from the depot or a k-level satellites to customers is allowed if it helps to decrease the cost, or to satisfy time and/or synchronization constraints, without passing through the rest of the levels.

# Chapter 5

# Models for the 2E-CVRP

The most common version of Multi-Echelon Vehicle Routing Problem arising in practice is the *Two-Echelon Vehicle Routing Problem*, where just two levels are considered. In multi-echelon freight distribution, when a considerable number of levels have to be managed, it is difficult to consider the overall system, and in many cases the problem is divided into a number of smaller problems which are solve independently or approximating some costs for the rest of the system. However, for multi-echelon problems where the number of levels is low (2,3 or 4), it can be possible to consider the entire system in optimization. As we have seen in chapters 2 and 3, many distribution systems are organized using two-echelon distribution strategies. Moreover, the main aspect of multi-echelon distribution is to model and manage the connexion and the eventual synchronization between two levels. This connexion can be studied in the simplest version of the NE-VRP, i.e. when only two echelons are considered, and then the conclusions of this study should be extended to each connexion between two levels. For these reasons, we will focus on the simplest problem of the family, the Two Echelon Vehicle Routing problem (2E-VRP). In this chapter we give a formal definition of the problem, describing the problem, and we present the basic notation in order to build different decision models.

## 5.1 Problem description

In this section give a formal definition of the basic Two-Echelon Vehicle Routing problems, which are obtained from the extension to two-echelon distribution of the classical VRP variants. First of all, we will specify how the transportation system works on a physical point of view, describing the interactions between the two levels and showing the different assumptions and hypothesis. Then, the basic variants of the 2E-VRP will be presented, as a particular case of the NE-VRP variants described in chapter 5.

Consider a two-echelon freight distribution system, which network is composed by one depot, a number of 1st-level intermediary points, which will be noted as satellites for simplicity, and a number of customers, each of them with a demand associated. All these nodes are represented as described in section 4.1. Two levels are then defined: the first level connects the depot with satellites, and the second level satellites with customers. To each level a fleet of vehicles is associated, which will operate only to bring the freight to the considered echelon.

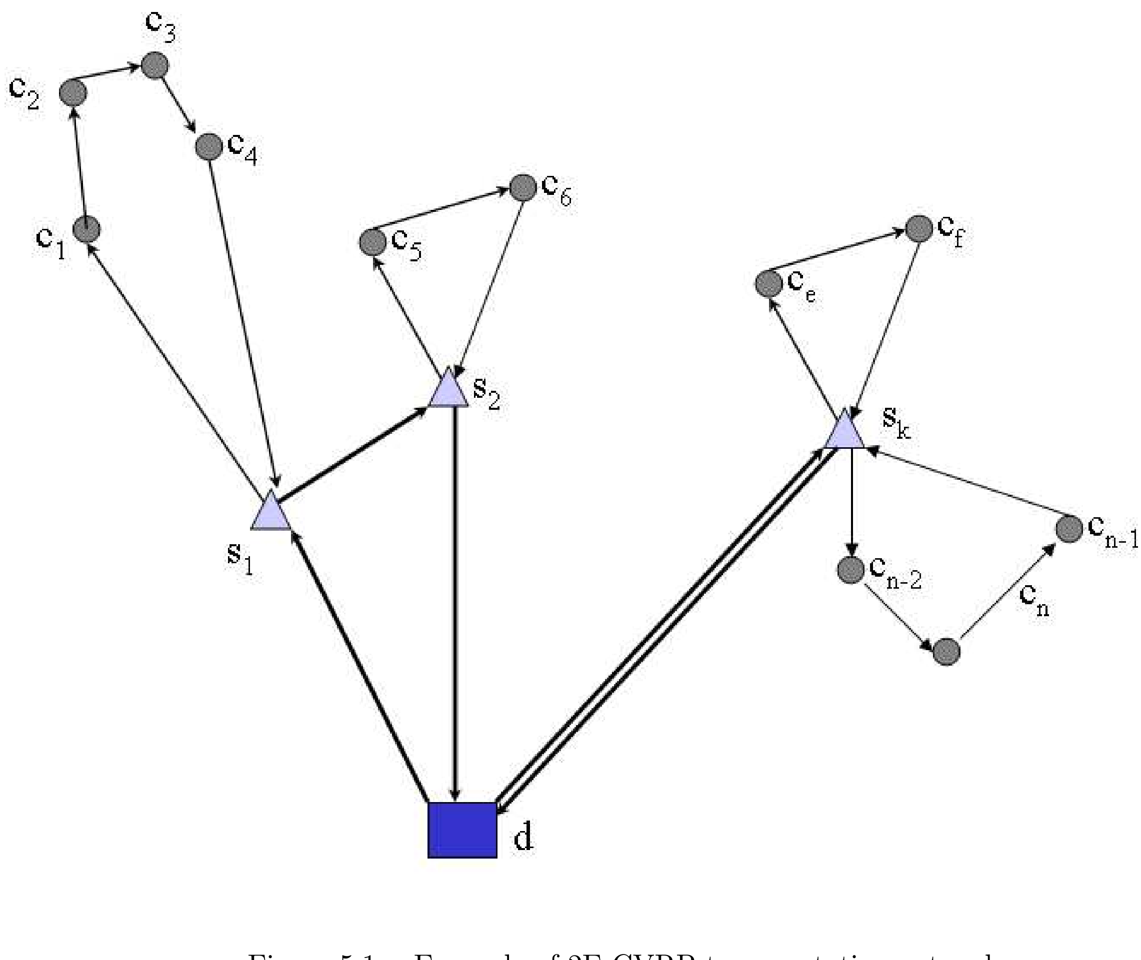


Figure 5.1. Example of 2E-CVRP transportation network

htb

From a physical point of view, a Two-Echelon Capacitated Vehicle Routing system represent a distribution system which operates as follows (see figure 5.1). Freight arrives at the depot, where it is consolidated into the 1st-level vehicles, unless it is already carried into a fully-loaded vehicle. Each 1st-level vehicle travels to a subset of satellites that will be determined by the model and then it will return to the depot. At a satellite, freight is transferred from 1st-level to 2nd-level vehicles. Each 2nd-level vehicle performs a route to serve the designated customers, and then travels to a satellite for its next cycle of operations. The 2nd-level vehicles return to their departure satellite.

Demand at satellites is not known, i.e., it is not a priori determined which satellite will receive the freight having to be delivered to each customer. In most real applications, it is not important to know through which satellite each customer's demand request passes, but the primordial aspect to consider will be to deliver the freight to the correspondent customer.a primordial decision parameter.

The basic version of the problem is the Two-Echelon Capacitated Vehicle Routing Problem (2E-CVRP). In these problems, vehicles and satellites have maximum quantity of freight which can be deposited in. In the case of vehicles, we consider that each level vehicle fleet is homogeneous, so each element of the same vehicle fleet level has the same fixed capacity. The objective is to serve customers by minimizing the total transportation cost, and satisfying the capacity constraints of the vehicles. There is a single depot and a fixed number of capacitated satellites. All customer demands are fixed and known in advance. Moreover, no time window is defined for the deliveries and the satellite operations. All customer demands must be satisfied. For the 2nd level, the demand of each customer is smaller than each vehicle's capacity and the demands cannot be split in multiple routes of the same level. For the 1st level we can consider two complementary distribution strategies. In the first case, each satellite is served by just one 1st-level vehicle and the demand passing through the satellite cannot be split into different 1st-level vehicles. This strategy is similar to the classical CVRP, and the capacity of 1st-level vehicles has to be greater than the demand of each satellite. In the second case, a satellite can be served by more of one 1st-level vehicle, so each satellite demand can be split. This strategy has some analogies to the VRP with split deliveries and allow to have 1st-level vehicles with capacity which is lower than each satellite demand.

Another basic problem is the Two-Echelon Distance Constrainted VRP. In this problem, routes cannot exceed a maximum distance, which can be expressed in terms of geographical distance or in terms of travel time. This constraint is related to vehicle capacities, so, in a homogeneous vehicle fleets problem, all the routes of the same level will be subjected to the same maximum traveled distance limitations.

From these two basic versions, other variants can be obtained. The main variants which have important real applications are presented below. Network features which can be added to the basic problem are related to multi-depot transportations systems. In variant, which will be called Multi-depot 2E-VRP (MD 2E-VRP), we will consider a network with more than one depot, and each customer will receive freight for only one of them. Freight is supposed to be the same, so each depot is equivalent to the others in terms of freight compatibility respect to customer's demand. Another interesting variant is Multi-depot 2E-CVRP with satellite freight reorganization (2E-MDCVRP-SR). This variant represent many real distribution cases in which a customer receives freight from more than one depot. This freight goes, for each customer, from each depot to the same satellite, where it is consolidated into one vehicle, then is distributed to the corresponding customer.

Variants which refer to service features can be divided into two groups. The first group variants have several time-related constraints. The most common problem of this group is Two-Echelon Capacitated VRP with Time Windows (2E-CVRP-TW). This case is the extension of 2E-CVRP where time windows on the arrival or departure time at the satellites and/or at the customers are considered. The time windows can be hard or soft. In the first case the time windows cannot be violated, while in the second one if they are violated a penalty cost is paid. A similar variant, which has more restrictive constraints, is Two-Echelon Capacitated VRP with Satellites Synchronization (2E-CVRP-SS). In this problem, time constraints on the arrival and the departure of the vehicles at the satellites are considered. In fact, the vehicles arriving in a satellite unload their cargo, which must be immediately loaded into a 2nd-level vehicle. Also this kind of constraints can be of two types: hard and soft. In the hard case, every time a 1st-level vehicle unloads its freight, 2nd-level vehicles must be ready to deliver it (this constraint is formulated through a very small hard time window). In the second case, if 2nd-level vehicles are not available, the demand is lost and a penalty is paid.

The other group of variants is related to other service policies. The most common problem of this group is 2E-CVRP with Pickup and Deliveries (2E-CVRP-PD). In this case we can consider the satellites as intermediate depots to store both the freight that has been picked-up from or must be delivered to the customers. Also the possibility to serve the customers directly from the depot can be considered. This is the case of 2E-CVRP with taxi services (2E-CVRP-TS), where direct shipping from the depot to customers is allowed if it helps to decrease the cost, or to satisfy time and/or synchronization constraints.

## 5.2 The Two-Echelon Capacitated Vehicle Routing Problem

In this section we will describe, in terms of combinatorial optimization concepts, the 2E-CVRP, giving the basic notation and definitions in order to build mathematical decision models. In a first time, we will define the combinatorial optimization concepts related to vehicle routing optimization for the two-echelon distribution system described above, extending the VRP basic definitions to two-echelon freight distribution. Then, two first models, which respectively derive from multicommodity network design and from classical VRP vehicle flow formulations, will be presented. The first model will be realized on an oriented graph, so it can be used for symmetric and asymmetric problems, whereas the second one, due to its nature, will be presented only for an undirected graph, for reasons of simplicity, and will be then used for symmetric problems.

### 5.2.1 General definitions and notation

Consider the Two-echelon distribution system described above. This system is operating in a network which can be represented by a graph $G$. This graph, as already presented, is composed by three sets of nodes: depot, satellites and customers. Let us denote the depot by $v_0$, the set of satellites by $V_s$ and the set of customers by $V_c$. Let $n_s$ be the number of satellites and $n_c$ the number of customers. The depot is the starting point of the freight and does not present limitations in terms of capacity or number of vehicles. Each satellite $k$ is supposed to have its own capacity, usually expressed in terms of maximum number of 2nd-level routes starting from the satellite or freight volume. In this case, we will use the fist measure unit for satellite $k$ capacity, which will be $m_{s_k}$. The customers are the destinations of the freight and each customer $i$ has associated a demand $d_i$, i.e. the quantity of freight that has to be delivered to that customer. Define the arc $(i,j)$ as the direct route connecting node $i$ to node $j$. If both nodes are satellites or one is the depot and the other is a satellite, we define the arc as belonging to the 1st-level network, while if both nodes are customers or one is a satellite and the other is a customer, the arc belongs to the 2nd-level network.

Two fleets, each of them composed by identical vehicles, are defined. The number of vehicles corresponding to 1st-level fleet is $m^1$, and each of them has the same capacity $K^1$. In the same way, a number of $m^2$ vehicles with identical capacity $K^2$ constitute the 2nd-level fleet. We define as *1st-level route* a route made by a 1st-level vehicle which starts from the depot, serves one or more satellites and ends at the depot. A *2nd-level route* is a route made by a 2nd-level vehicle which starts from a satellite, serves one or more customers and ends at the same satellite.

The distribution of the freight can not be managed by direct shipping from the depot to the customers, but the freight must be consolidated from the depot to a satellite and then delivered from the satellite to the desired customer. This implicitly defines a two-echelon transportation system: the 1st level interconnecting the depot to the satellites and the 2nd one the satellites to the customers (see Figure 5.1). We consider only one type of freight, i.e. the volumes of freight belonging to different customers can be stored together and loaded in the same vehicle for both the 1st and the 2nd-level vehicles. Moreover, each customer's demand cannot be split among different vehicles *at the 2nd level*. In this way, each customer must be visited by one and only one 2nd-level vehicle. Only if customer's demand is greater than each 2nd-level vehicle capacity, more than one vehicle is allowed to visit it, but with the following distribution strategy: a number of full vehicles will realize a direct shipping TL operation from the corresponding satellite, whereas the remaining freight having to be delivered to this customer will be loaded into a vehicle which will follow a LTL distribution strategy. The same policy is applied to first level, but, whereas for 2nd-level vehicles this situation will be not usual, due to the two-echelon distribution organization, it will be common to observe it for the 1st-level vehicle trips. For this reason, we consider that each satellite can be served by more than one 1st-level vehicle, so the aggregated freight assigned to each satellite can be split into two or more vehicles. Each 1st level vehicle can deliver the freight of one or more customers, as well as serve more than one satellite in the same route.

The freight must be delivered from the depot $v_0$ to the customers set $V_c = \{v_{c_1},v_{c_2},...,v_{c_{n_c}}\}$. Let $d_i$ the demand of the customer $c_i$. The number of 1st-level vehicles available at the depot is $m_1$. These vehicles have the same given capacity $K^1$. The total number of 2nd-level vehicles available for the second level is equal to $m_2$. The total number of active vehicles can not exceed $m_2$ and each satellite $k$ have a maximum capacity $m_{s_k}$. The 2nd-level vehicles have the same given capacity $K^2$.

The problem is easily seen to be NP-Hard via a reduction to VRP, which is a special case of 2E-CVRP arising when just one satellite is considered. If we consider a two-echelon VRP with only one satellite of unlimited capacity, where 1st-level costs are negligible, and with one 1st-level vehicle of unlimited capacity, we obtain an equivalent problem to a classical CVRP. Since CVRP is known to be NP-hard, and an equivalent 2E-CVRP can be defined, we can affirm that the problem we are defining is also NP-hard.

In the following we will present two vehicle flow formulations for the 2E-CVRP. The first of them is a two-index vehicle flow formulation which derives from multicommodity network design, based on the concept of freight flow to represent vehicle load. Then, a second model, using the three-index vehicle flow formulation obtained by extension two two-echelon distribution of the three-index VRP formulation presented in (103)

According to the definition of 2E-CVRP, if the assignments between customers and satellites are determined, the problem reduces to $1 + n_s$ CVRP (1 for the 1st-level and $n_s$ for the 2nd-level). In the same way, if the first level operates a TL policy (with a number of direct routes depot-satellite-depot), the problem can easily be converted into a Multidepot CVRP. In the other cases, the main question when modeling 2E-CVRP is how to connect the two levels and manage the dependence of the 2nd-level from the 1st one. The formulation we present derives from the multi-commodity network design and uses the flow of the freight on each arc as main decision variables. In this model we will not consider the fixed costs of the vehicles, since we suppose they are available in fixed number.

We consider the following costs. First of all, we define the travel costs $c_{ij}$, which are of two types: costs of the arcs traveled by 1st-level vehicles, i.e. arcs connecting the depot to the satellites and the satellites between them, and costs of the arcs traveled by 2nd-level vehicles, i.e. arcs connecting the satellites to the customers and the customers between them. Another type of costs that can be considered part of the global transport operation, even if they are not classical transport costs, are those related to the loading and unloading operations at the satellites. Supposing that the number of workers in each satellite $v_{s_k}$ is fixed, we consider only the cost incurred by the management of the freight and we define $S_k$ as the unit cost of freight handling at the satellite $v_{s_k}$.

Summarizing, all these concepts can be resumed in the following table.

| | |
|---|---|
| $V_0 = \{v_0\}$ | Depot |
| $V_s = \{v_{s_1},v_{s_2},\ldots,v_{s_{n_s}}\}$ | Set of satellites |
| $V_c = \{v_{c_1},v_{c_2},\ldots,v_{c_{n_c}}\}$ | Set of customers |
| $n_s$ | number of satellites |
| $n_c$ | number of customers |
| $m_1$ | number of the 1st-level vehicles |
| $m_2$ | number of the 2nd-level vehicles |
| $m_{s_k}$ | maximum number of 2nd-level routes starting from satellite $k$ |
| $K^1$ | capacity of the vehicles for the 1st level |
| $K^2$ | capacity of the vehicles for the 2nd level |
| $d_i$ | demand required by customer i |
| $c_{ij}$ | cost of the arc $(i,j)$ |
| $S_k$ | cost for loading/unloading operations of a unit of freight in satellite $k$ |

Table 5.1. Definitions and notation

### 5.2.2 A Flow-based Model for 2E-CVRP

The concepts presented below are common to all vehicle flow formulations. These models define a group of decision variables associated to arcs, and also another group of variables to represent the satellite-customer assignments. In this subsection we will define a first formulation, which derives from multicommodity network design, which will be noted as *Flow Model*. In order to define a route and avoid internal cycles, we will use the concept of freight flow, which is associated to vehicle load, in the following way. Given an arc $ij$, the load flow is defined as the freight quantity traveling from node i to node j, without specifying how many routes will be used to do it (for the second level, each customer is visited by one vehicle, but for the 1st level routes, satellites do not have constraints for the minimum and maximum number of 1st-level vehicles that can receive, only a capacity limitation). The load flow will be used as a variable and defined above. The decision variables which will be considered in this model are grouped in five sets, which can be divided related to its nature in following three groups:

- The first group represents the arc usage variables. We define two sets of such variables, one for each level. The variable $x_{ij}$ is an integer variable of the 1st-level routing and is equal to the number of 1st-level vehicles using arc ($i$,$j$). The variable $y_{ij}^k$ is a binary variable representing the 2nd-level routing. It is equal to 1 if a 2nd-level vehicle makes a route starting from satellite $k$ and goes from node $i$ to node $j$, 0 otherwise.

- The second group of variables represents the assignment of each customer to one satellite and are used to link the two transportation levels. More precisely, we define $z_{kj}$ as a binary variable that is equal to 1 if the freight to be delivered to customer $j$ is consolidated in satellite $k$ and 0 otherwise.

- The third group of variables, split into two subsets, one for each level, represents the freight flow passing through each arc. We define the freight flow as a variable $Q_{ij}^1$ for the 1st-level and $Q_{ijk}^2$ for the 2nd level, where $k$ represents the satellite where the freight is passing through. Both variables are continuous.

In order to lighten the model formulation, we define the auxiliary quantity

$$D_k = \sum_{j \in V_c} d_j z_{kj}, \forall k \in V_s, \tag{5.1}$$

which represents the freight passing through each satellite $k$.

Summarizing, all these variables, as long as the auxiliary quantity $D_k$ can be resumed in the following table.

| | |
|---|---|
| $D_k$ | demand changing vehicle at satellite $k$ |
| $Q^1_{ij}$ | flow passing through the 1st-level arc $(i,j)$ |
| $Q^2_{ijk}$ | flow passing through the 2st-level arc $(i,j)$ and coming from satellite $k$ |
| $x_{ij}$ | number of 1st-level vehicles using the 1st-level arc $(i,j)$ |
| $y^k_{ij}$ | boolean variable equal to 1 if the 2st-level arc $(i,j)$ is used by the 2nd-level routing starting from satellite $k$ |
| $z_{kj}$ | variable set to 1 if the customer $c_i$ is served by the satellite $k$ |

Table 5.2. Definitions and notation

We can then build a mathematical formulation for the travel costs optimization problem we have defined above. The model to minimize the total cost of the system due to freight distribution may be formulated as follows:

$$\min \sum_{i,j \in V_0 \cup V_s, i \neq j} c_{ij} x_{ij} + \sum_{k \in V_s} \sum_{i,j \in V_s \cup V_c, i \neq j} c_{ij} y^k_{ij} + \sum_{k \in V_s} S_k D_k$$

Subject to

$$\sum_{i \in V_s} x_{0i} \leq m_1 \tag{5.2}$$

$$\sum_{j \in V_s \cup V_0, j \neq k} x_{jk} = \sum_{i \in V_s \cup V_0, i \neq k} x_{ki} \quad \forall k \in V_s \cup V_0 \tag{5.3}$$

$$\sum_{k \in V_s} \sum_{j \in V_c} y^k_{kj} \leq m_2 \tag{5.4}$$

$$\sum_{j \in V_c} y^k_{kj} \leq m_{s_k} \quad \forall k \in V_s \tag{5.5}$$

$$\sum_{j \in V_c} y^k_{kj} = \sum_{j \in V_c} y^k_{jk} \quad \forall k \in V_s \tag{5.6}$$

$$\sum_{i \in V_s \cup v_0, i \neq j} Q^1_{ij} - \sum_{i \in V_s \cup v_0, i \neq j} Q^1_{ji} = \begin{cases} D_j & j \text{ is not the depot} \\ \sum_{i \in V_c} -d_i & \text{otherwise} \end{cases} \quad \forall j \in V_s \cup V_0 \tag{5.7}$$

$$Q^1_{ij} \leq K^1 x_{ij} \quad \forall i,j \in V_s \cup V_0, i \neq j \tag{5.8}$$

$$\sum_{i \in V_c \cup k, i \neq j} Q^2_{ijk} - \sum_{i \in V_c \cup k, i \neq j} Q^2_{jik} = \begin{cases} z_{kj} d_j & j \text{ is not a satellite} \\ -D_j & \text{otherwise} \end{cases} \quad \forall j \in V_c \cup V_s, \forall k \in V_s \tag{5.9}$$

$$Q^2_{ijk} \leq K^2 y^k_{ij} \quad \forall i,j \in V_s \cup V_c, i \neq j, \forall k \in V_s \tag{5.10}$$

$$\sum_{i \in V_s} Q^1_{iv_0} = 0 \tag{5.11}$$

$$\sum_{j \in V_c} Q^2_{jkk} = 0 \quad \forall k \in V_s \tag{5.12}$$

$$y^k_{ij} \leq z_{kj} \quad \forall i \in V_s \cup V_c, \forall j \in V_c, \forall k \in V_s \tag{5.13}$$

$$y^k_{ji} \leq z_{kj} \quad \forall i \in V_s, \forall j \in V_c, \forall k \in V_s \tag{5.14}$$

$$\sum_{i \in V_s \cup V_c} y^k_{ij} = z_{kj} \quad \forall k \in V_s, \forall j \in V_c \tag{5.15}$$

$$\sum_{i \in V_s} y^k_{ji} = z_{kj} \quad \forall k \in V_s, \forall j \in V_c \tag{5.16}$$

$$\sum_{i \in V_s} z_{ij} = 1 \quad \forall j \in V_c \tag{5.17}$$

$$y^k_{kj} \leq \sum_{l \in V_s \cup V_0} x_{kl} \quad \forall k \in V_s, \forall j \in V_c \tag{5.18}$$

$$y^k_{ij} \in \{0,1\}, z_{kj} \in \{0,1\}, \quad \forall k \in V_s \cup V_0, \forall i,j \in V_c \tag{5.19}$$

$$x_{kj} \in \mathbb{Z}^+, \quad \forall k,j \in V_s \cup V_0 \tag{5.20}$$

$$Q^1_{ij} \geq 0, \forall i,j, \in V_s \cup V_0,\ Q^2_{ijk} \geq 0,\ \forall i,j \in V_s \cup V_c, \forall k \in V_s. \tag{5.21}$$

The objective function minimizes the sum of the traveling and handling operations costs. Constraints (5.3) show, for $k = v_0$, that each 1st-level route begins and ends at the depot, while when $k$ is a satellite, impose the balance of vehicles entering and leaving that satellite. The limit on the satellite capacity is satisfied by constraints (5.5). They limit the maximum number of 2nd-level routes starting from every satellite (notice that the constraints also limit at the same time the freight capacity of the satellites). Constraints (5.6) force each 2nd-level route to begin and end to one satellite and the balance of vehicles entering and leaving each customer. The number of the routes in each level must not exceed the number of vehicles for that level, as imposed by constraints (5.2) and (5.4).

Constraints (5.7) and (5.9) indicate that the flows balance on each node is equal to the demand of this node, except for the depot, where the exit flow is equal to the total demand of the customers, and for the satellites at the 2nd-level, where the flow is equal to the demand (unknown) assigned to the satellites. Moreover, constraints (5.7) and (5.9) forbid the presence of subtours not containing the depot or a satellite, respectively. In fact, each node receives an amount of flow equal to its demand, preventing the presence of subtours. Consider, for example, that a subtour is present between the nodes $i$, $j$ and $k$ at the 1st level. It is easy to check that, in such a case, does not exist any value for the variables $Q^1_{ij}$, $Q^1_{jk}$ and $Q^1_{ki}$ satisfying the constraints (5.7) and (5.9). The capacity constraints are formulated in (5.8) and (5.10), for the 1st-level and the 2nd-level, respectively. Constraints (5.11) and (5.12) do not allow residual flows in the routes, making the returning flow of each route to the depot (1st-level) and to each satellite (2nd-level) equal to 0.

Constraints (5.13) and (5.14) indicate that a customer $j$ is served by a satellite $k$ ($z_{kj} = 1$) only if it receives freight from that satellite ($y^k_{ij} = 1$). Constraint (5.17) assigns each customer to one and only one satellite, while constraints (5.15) and (5.16) indicate that there is only one 2nd-level route passing through each customer. At the same time, they impose the condition that a 2nd-level route departs from a satellite $k$ to deliver freight to a customer if and only the customer's freight is assigned to the satellite itself. Constraints (5.18) allow a 2nd-level route to start from a satellite $k$ only if a 1st-level route has served it.

### 5.2.3 TSP-based Model

The model presented above is not derived from the classical VRP formulation, which uses subtour elimination instead of flow constraints to avoid internal cycles and define the capacity limitations. Other way to produce a formulation for the 2E-CVRP can be to extend to multi-echelon distribution systems the classical VRP models, based on TSP formulation. For simplicity, and because the first test cases will be realized on instances with symmetric euclidean costs, we will consider a network with a symmetric cost matrix. We consider then an unoriented graph $G$ in which we define three sets of nodes: one set containing the depot, one containing $n_s$ satellites and one containing $n_c$ customers, as defined in section 5.2.1. We consider all the definitions described in that section, with the only specification in the arc costs that, given a node $i$ and a node $j$, costs $c_(ij)$ and $c_(ji)$ will be the same. In this case, each node will be defined by a numeric index. Given two nodes of the same echelon $i$ and $j$, index $i$ being lower than index $j$, the edge will be noted as $ij$ and to each variable, the double index referring to the considered edge will be written in the same way $(ij)$.

As seen while defining the flow model, the first level has the particularity that each satellite can be visited by more than one 1st-level vehicle, or, in some cases, they can remain unused so unvisited. The satellites demand is also to be determined. For these two reasons, we have considered that a simpler way to define the TSP-like model is to use the three index formulation, associating each variable to a route. Note that in the flow model the second level is defined by three-index variables, where two indexes refer to the arc and the third to the satellite. In the three-index formulation we will present each x and y variable has an index which refers to the vehicle that performs the route. The routes are then explicited, and in general the number of constraints does not change in a considerable way respect to the two-index TSP-like model (the number of available vehicles is in general small). This can be also useful to define problems where the distribution service presents heterogeneous vehicle fleets.

The decision variables which will be considered in the TSP-like model are grouped in four sets, which are the following:

- The first and the second sets of variables represents the arc usage variables, each of them representing one level. The variable $x_{ij}^k$ is a binary variable of the 1st-level routing and is equal 1 if the 1st-level vehicle $k$ travels the arc connecting nodes $i$ and $j$, otherwise it is equal to 0.
- The variable $y_{ij}^k$ is a binary variable representing the 2nd-level routing, and is equal to 1 if the 2nd-level vehicle $k$ goes from node $i$ to node $j$, 0 otherwise.
- The third set of variables represents the assignment of each satellite to one 1st-level route. We define $w_i^{1k}$ as a binary variable that is equal to 1 if satellite i is visited by 1st-level route $k$ and 0 otherwise.
- The fourth set of variables represents the assignment of each customer to one 2nd-level route and are also used to link the two transportation levels. More precisely, we define $w_{is}^{2k}$ as a binary variable that is equal to 1 if customer $j$ is served by route $k$ starting and ending at satellite $s$ and 0 otherwise.

In order to lighten the model formulation, we define the auxiliary quantity

$$D_s^k = \sum_{l \in V_c} d_l w_{li}^{2k}; \forall s \in V_s; \forall k \in 1..m_1 \tag{5.22}$$

which represents the freight traveling to satellite $s$ into vehicle $k$.

Also another auxiliary quantity is defined:

$$D_s = \sum_{k=1}^{k=m_1} D_s^k; \forall s \in V_s \tag{5.23}$$

which represents the freight passing through each satellite $s$.

Summarizing, all these variables, as long as the auxiliary quantity $D_k$ can be resumed in the following table.

We can then build a mathematical model for the route optimization problem we have defined above. This formulation has as objective to minimize the total cost of the system due to freight distribution, and may be formulated as follows:

| | |
|---|---|
| $D_s$ | demand changing vehicle at satellite $s$ |
| $x_{ij}$ | boolean variable equal to 1 if the 1st-level edge $(i,j)$ is used by the 1st-level route $k$ |
| $y_{ij}^k$ | boolean variable equal to 1 if the 2nd-level edge $(i,j)$ is used by the 2nd-level route $k$ which starts from satellite $k$ |
| $w_i^{2k}$ | variable set to 1 if the satellite $s_i$ is served by the 1st-level |
| $w_{is}^{2k}$ | variable set to 1 if the customer $c_i$ is served by the 2nd-level route $k$, which starts from satellite $s$ |

Table 5.3. Definitions and notation

$$\min \sum_{k=1}^{k=m_1} \sum_{i,j \in V_0 \cup V_s} c_{ij} x_{ij}^k + \sum_{k=1}^{k=m_2} \sum_{i,j \in V_s \cup V_c} c_{ij} y_{ij}^k + \sum_{k \in V_s} S_k D_k \tag{5.24}$$

Subject to

$$\sum_{i \in V_s} x_{0i}^k \leq 1; \forall k \in 1..m_1 \tag{5.25}$$

$$\sum_{j \in V_s} x_{j0}^k = \sum_{i \in V_s} x_{0i}^k; \forall k \in 1..m_1 \tag{5.26}$$

$$\sum_{i \in V_s, j \in V_c} y_{ij}^k \leq 1; \forall k \in 1...m_2 \tag{5.27}$$

$$\sum_{j \in V_c} y_{ij}^k = \sum_{j \in V_c} y_{ji}^k; \forall i \in V_s, k \in 1...m_2 \tag{5.28}$$

$$\sum_{i \in V_s \cup \{0\}} x_{ij}^k = w_j^{1k}, \forall j \in V_s, k \in 1...m_1 \tag{5.29}$$

$$\sum_{i \in V_s \cup \{0\}} x_{ji}^k = w_j^{1k}, \forall j \in V_s, k \in 1...m_1 \tag{5.30}$$

$$\sum_{i \in V_s \cup V_c} y_{ij}^k = w_{js}^{2k}, \forall j \in V_c, s \in V_s, k \in 1...m_2 \tag{5.31}$$

$$\sum_{i \in V_s \cup V_c} y_{ji}^k = w_{js}^{2k}, \forall j \in V_c, s \in V_s, k \in 1...m_2 \tag{5.32}$$

$$\sum_{i,j \in S_s} x_{ij}^k \leq |S_s| - 1; \forall S_s \subseteq V_s, k \in 1...m_1 \tag{5.33}$$

$$\sum_{i,j \in S_c} y_{ij}^k \leq |S_c| - 1; \forall S_c \subseteq V_c, k \in 1...m_2 \tag{5.34}$$

$$\sum_{i \in V_s \cup \{0\}} x_{ij}^k \geq D_j^k / K^1; \forall j \in V_s; \forall k \in 1...m_1 \tag{5.35}$$

$$\sum_{i \in V_c} d_i * w_{is}^{2k} \leq K^2; \forall s \in V_s; \forall k \in 1...m_2 \tag{5.36}$$

$$\sum_{k=1}^{k=m_1} \sum_{i \in V_c} y_{si}^k \leq m_{s_s}; \forall s \in V_s \tag{5.37}$$

$$x_{ij}^k \leq w_i^{1k}; \forall i \in V_s \cup \{0\}, j \in V_s; \forall k \in 1...m_1 \tag{5.38}$$

$$x_{j0}^k \leq w_j^{1k}; \forall j \in V_s; \forall k \in 1...m_1 \tag{5.39}$$

$$y_{ij}^k \leq w_{is}^{2k}; \forall i \in V_s \cup V_c, j \in V_c; \forall s \in V_s, k \in 1...m_2 \tag{5.40}$$

$$y_{ji}^k \leq w_{js}^{2k}; \forall i \in V_s, j \in V_c; \forall s \in V_s, k \in 1...m_2 \tag{5.41}$$

$$\sum_{k \in 1...m_2} w_{is}^{2k} = 1; \forall i \in V_c, \forall s \in V_s \tag{5.42}$$

$$\sum_{s \in V_s} w_{is}^{2k} = 1; \forall i \in V_c, k \in 1...m_2 \tag{5.43}$$

$$\sum_{k=1}^{k=m_2} w_{is}^{2k} \leq \sum_{k=1}^{k=m_1} w_s^{1k}; \forall s \in V_s; \forall i \in V_c \tag{5.44}$$

$$x_{ij}^k \in \{0,1\}, y_{ij}^k \in \{0,1\}, w_i^{1k} \in \{0,1\}, w_{is}^{2k} \in \{0,1\}; D_i, D_i^k \in \mathbb{N} \tag{5.45}$$

Similarly to the flow model, the objective function minimizes the sum of the traveling and handling operations costs. Constraints (5.25) show that each 1st-level route starts and end at the depot, and constraints (5.27) show that each 2nd-level route starts and ends at the same satellite. Constraint (5.26) show that the total number of 1st-level routes cannot exceed the number of available 1st-level vehicles $m_1$. In the same way, constraint (5.28) show that the total number of 2nd-level routes cannot exceed the number of available 2nd-level vehicles $m_2$. Each 1st-level route visits each satellite at most one, and will visit only the satellites which are assigned to the considered route, as shown in constraints 5.29 and 5.30. The same behavior is represented for the second level in 5.31 and 5.32. To avoid internal cycles in a route, subtour elimination constraints are defined respectively for each level at 5.33 and 5.34. This form is enough to avoid subtours, even if it's not the most restrictive, but due to the connexion between the two levels and the variability of the satellites demands (which is not fixed or a priori easily determinable) these constraints are valid and easy t insert in our model.

Each vehicle cannot be loaded with a higher freight quantity than its capacity, respectively for each level, as show in 5.35 and 5.36. The limit on the satellite capacity is satisfied by constraints (5.37). They limit the maximum number of 2nd-level routes starting from every satellite (notice that the constraints also limit at the same time the freight capacity of the satellites). Constraints 5.38 and 5.39 show that each satellite will be visited by one 1st-level vehicle only if it's assigned to it, and analogously constraints 5.40 and 5.41 show that each customer is visited by one 2nd-level vehicle only if it's assigned to it. Each customer is visited by only one 2nd-level vehicle, and it's assigned to only one satellite, as seen at 5.42 and 5.43. The freight which arrives at one satellite with 1st-level routes has to start with 2nd-level routes from the same satellite, as represent the connexion constraints 5.44

## 5.3 Valid inequalities for 2E-CVRP

In order to strengthen the continuous relaxation of the flow model, we introduce cuts derived from VRP formulations. In particular, we use two families of cuts, one applied to the assignment variables derived from the subtour elimination constraints (edge cuts) and the other based on the flows.

The *edge cuts* explicitly introduce the well-known subtours elimination constraints derived from the TSP. They can be expressed as follows:

$$\sum_{i,j \in S_c} y_{ij}^k \leq |S_c| - 1, \forall S_c \subset V_c,\ 2 \leq |S_c| \leq |V_c| - 2 \tag{5.46}$$

These inequalities explicitly forbid the presence in the solution of subtours not containing the depot, already forbidden by Constraints (5.9). The number of potential valid inequalities is exponential, so we should need a separation algorithm to add them. As we will show in chapter 6, in practice the inequalities involving sets $S_c$ with cardinality more than 3 are unuseful and the separation algorithm can be substituted by a direct inspection of the constraints up to cardinality equal to 3.

The aim of flow cuts is to reduce the splitting of the values of the binary variables when the continuous relaxation is performed, strengthening the BigM constraints (5.10). The idea is to reduce the constant $K^2$ by considering that each customer reduces the flow by an amount equal to its demand $d_i$. Thus the following inequalities are valid:

$$\begin{cases} Q_{ijk}^2 \leq (K^2 - d_i) y_{ij}^k, \forall i,j \in V_c \ \ \forall k \in V_s \\ Q_{ijk}^2 - \sum_{l \in V_s} Q_{jlk}^2 \leq (K^2 - d_i) y_{ij}^k \ \ \forall i,j \in V_c, \forall k \in V_s. \end{cases} \tag{5.47}$$

Constraints (5.47) are of the same order of magnitude of (5.10), so they can be directly introduced into the model.

# Chapter 6

# Computational tests

In this chapter, we will present and analyze the behavior of the models presented below using a commercial solver. To study these models, we needed to build new sets of instance, due to the fact this problem has not been studied in a general way in literature and no benchmarks were proposed which should be used to test our formulations. In Section 6.1 we present four sets of benchmark instances, and we describe how these instances were built. The first of these sets is composed by instances with a small number of customers. In Section 6.2 we present the results of the vehicle flow models on a set of small-sized instances highlighting the properties of 2E-CVRP, and the limits of each formulation. Moreover, we present, from the same instances, a study which results show the cost distribution according to the geographic distribution of the satellites. Finally, Section 6.3 is devoted to present the computational results on all the benchmark instances and the impact on the computational results of the valid inequalities of Section 5.3, testing the flow model on instances with a number of nodes up to 5 satellites and 50 customers.

## 6.1 Construction of the instance sets

In this section we introduce different instance sets for 2E-CVRP. The instances cover up to 51 nodes (1 depot and 50 customers) and are grouped in four sets. The first three sets have been built from the existing instances for VRP by Christofides and Eilon denoted as E-n13-k4, E-n22-k4, E-n33-k4 and E-n51-k5 (21) , while the third set is constituted by randomly generated instances simulating different geographical distributions, including customers distribution in urban and regional areas. All the instance sets can be downloaded from the web site of OR-Library (10) .

The first instance set is made by 66 small-sized instances with 1 depot, 12 customers and 2 satellites. All the instances have the cost matrix of the instance E-n13-k4 (the costs of the matrix of the original instance is read as an upper triangular matrix and the corresponding optimal cost of the VRP instance is 290). The two satellites are placed over two customers in all the $\begin{pmatrix} 12 \\ 2 \end{pmatrix} = 66$ possible ways (the case where some customers are used as satellites is quite common for different kinds of distribution, e.g. grocery distribution). When a node is both a customer and a satellite, the arc cost $c_{ki}$ is set equal to 0. The number of vehicles for the 1st-level is set to 2, while the 2nd-level vehicles are 4, as in the original VRP instance. The capacity of the 1st-level vehicles is 2.5 times the capacity of the 2nd-level vehicles, to represent cases in which the 1st-level is made by trucks and the 2nd-level is made by smaller vehicles (e.g., vehicles with a maximum weight smaller than 3.5 t). The capacity of the 2nd-level vehicles is equal to the capacity of the vehicles of the VRP instance. The cost due to loading/unloading operations is set equal to 0, while the arc costs are the same of the VRP instances.

The second set of instances is obtained in a similar way from the instances E-n13-k4, E-n22-k4, E-n33-k4 and E-n51-k5. The instances are obtained by considering 6 pairs randomly generated satellites. For the instance E-n51-k5, which has 50 customers, we build an additional group of 3 instances obtained randomly placing 4 satellites instead of 2. The cost due to loading/unloading operations is set equal to 0, while the arc costs are the same of the VRP instances.

The main issue in the original instances by Christofides and Eilon is that the depot is in an almost central position in respect to the area covered by the customers. The third set of instances also considers the instances E-n13-k4, E-n22-k4, E-n33-k4 and E-n51-k5 by considering six pairs of satellites randomly chosen between the customers on the external border of the area determined by the customers distribution. Moreover, the depot is external to the customers areas, being placed at the coordinate (0,0) (the southeast corner of the customers area).

Finally, the fourth set includes instances generated in order to simulate different geographical distributions of the customers as well as of the satellites arising in urban and regional applications. The instances are generated according to the following parameters:

- Depot. The depot is external to the customers' areas and is located at the South-East corner of the square of side 100.
- Customers. They are generated according to three rules:
  - Random. In this case the customers' positions are randomly generated. This rule is the same used in the Christofides and Eilon's instances.
  - Centroids. The customers area is a circle of radius 100. In the customers' area are located 8 centroids, 4 randomly generated in an inner circle of radius 33 and 4 outside the circle. For each centroid, a cluster of customers is randomly created. This distribution simulates the situation in a urban area, where the centroids represents the neighborhoods of the city. In fact, usually the clustering is easy to determine in the peripheral neighborhoods, while the clusters in the areas near to the city center intersect each other.
  - Quadrants. The customers area is a circle of radius 100. The circle is split into 4 equal quadrants and one centroid is randomly located for each quadrant. For each centroid, a cluster of customers is randomly created. This distribution simulates the regional distribution, where the centroids represents villages and small cities.
- Satellites. The number of the satellites is 2, 3 and 5. The satellites are capacitated in terms of maximum number of 2nd-level routes starting from every satellite, while their cost for loading/unloading operations is set equal to 0. They are located according to the following rules:
  - Border Random. The satellites are randomly located on the external border of the area determined by the customers distribution. This distribution occurs when the satellites must be compulsory placed in existing areas, such as disused industrial areas or railway stations.
  - Sliced. The external border of the area determined by the customers distribution is split into a number of slices equal to the number of satellites. For each slice a satellite in randomly located. This occurs when the municipalities have freedom in choosing the satellite location.

  - Forbidden. Given the external border of the area determined by the customers distribution, 1/3 of this area is considered forbidden, while the remaining is split into a number of slices equal to the number of satellites. For each slice a satellite in randomly located. This distribution simulates the case where the satellites can not be located in a portion of the area as in the case of cities close to mountains or the sea.

- Customer demand. The demand is randomly selected in the range [0,1000].

- Vehicles. The capacity is set to 5000 for the 2nd-level vehicles and to 12500 for the 1st-level vehicles. The number of the 2nd-level vehicles ensures to have a ratio between the total demand of the customers and the loading capacity of the 2nd-level fleet from 0.7 to 0.92.

- Arc costs. The arc costs are integer and are computed as the Euclidean distances between the different coordinates (depot, satellites, customers).

Far each combination of number of satellites, customer and satellite distribution, two instances are created, for a total of 54 instances.

A summary of the main features of the different sets are reported in Table 6.1. The first column reports the set of instances, while the number of instances in shown in Column 2. Columns 3 and 4 contain the number of satellites and customers, respectively. The number of vehicles for the 1st and the 2nd level can be read in Columns 5 and 6, while Columns 7 and 8 give the capacity of the vehicles of the two levels. Finally, Column 9 shows the capacity of the satellites in terms of maximum number of routes starting from each satellite. In the remaining columns the rule used to localize the satellites and the customers are specified. More in detail, for the satellites the value *All pairs* indicates that all the possible pairs have been computed, *Random* that the satellites are randomly selected, while *Border Random*, *Sliced* and *Forbidden* have the meanings specified in the description of the set 4. About the customers, the value *From 'Instance-name' instance* indicates that we used the same customer distribution of the instance named *Instance-name*, while *Random*, *Centroids* and *Quadrants* have the meanings specified in the description of the set 4.

| Set | N° of Instances | $n_s$ | $n_c$ | $m_1$ | $m_2$ | $K^1$ | $K^2$ | Sat. capacity | Sat. distribution | Customer distribution |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 66 | 2 | 12 | 3 | 4 | 15000 | 6000 | ∞ | All pairs | From E-n13-k4 instance |
| 2 | 6 | 2 | 21 | 3 | 4 | 15000 | 6000 | ∞ | Random | From E-n22-k4 instance |
| 2 | 6 | 2 | 32 | 3 | 4 | 20000 | 8000 | ∞ | Random | From E-n33-k4 instance |
| 2 | 6 | 2 | 50 | 3 | 5 | 400 | 160 | ∞ | Random | From E-n51-k5 instance |
| 2 | 3 | 4 | 50 | 3 | 5 | 400 | 160 | ∞ | Random | From E-n51-k5 instance |
| 3 | 6 | 2 | 21 | 3 | 4 | 15000 | 6000 | ∞ | Border Random | From E-n22-k4 instance |
| 3 | 6 | 2 | 32 | 3 | 4 | 20000 | 8000 | ∞ | Border Random | From E-n33-k4 instance |
| 3 | 6 | 2 | 50 | 3 | 5 | 400 | 160 | ∞ | Border Random | From E-n51-k5 instance |
| 4 | 2 | 2 | 50 | 3 | 6 | 12500 | 5000 | 4 | Border Random | Random |
| 4 | 2 | 2 | 50 | 3 | 6 | 12500 | 5000 | 4 | Sliced | Random |
| 4 | 2 | 2 | 50 | 3 | 6 | 12500 | 5000 | 4 | Forbidden | Random |
| 4 | 2 | 2 | 50 | 3 | 6 | 12500 | 5000 | 4 | Border Random | Centroids |
| 4 | 2 | 2 | 50 | 3 | 6 | 12500 | 5000 | 4 | Sliced | Centroids |
| 4 | 2 | 2 | 50 | 3 | 6 | 12500 | 5000 | 4 | Forbidden | Centroids |
| 4 | 2 | 2 | 50 | 3 | 6 | 12500 | 5000 | 4 | Border Random | Quadrants |
| 4 | 2 | 2 | 50 | 3 | 6 | 12500 | 5000 | 4 | Sliced | Quadrants |
| 4 | 2 | 2 | 50 | 3 | 6 | 12500 | 5000 | 4 | Forbidden | Quadrants |
| 4 | 2 | 3 | 50 | 3 | 6 | 12500 | 5000 | 3 | Border Random | Random |
| 4 | 2 | 3 | 50 | 3 | 6 | 12500 | 5000 | 3 | Sliced | Random |
| 4 | 2 | 3 | 50 | 3 | 6 | 12500 | 5000 | 3 | Forbidden | Random |
| 4 | 2 | 3 | 50 | 3 | 6 | 12500 | 5000 | 3 | Border Random | Centroids |
| 4 | 2 | 3 | 50 | 3 | 6 | 12500 | 5000 | 3 | Sliced | Centroids |
| 4 | 2 | 3 | 50 | 3 | 6 | 12500 | 5000 | 3 | Forbidden | Centroids |
| 4 | 2 | 3 | 50 | 3 | 6 | 12500 | 5000 | 3 | Border Random | Quadrants |
| 4 | 2 | 3 | 50 | 3 | 6 | 12500 | 5000 | 3 | Sliced | Quadrants |
| 4 | 2 | 3 | 50 | 3 | 6 | 12500 | 5000 | 3 | Forbidden | Quadrants |
| 4 | 2 | 5 | 50 | 3 | 6 | 12500 | 5000 | 2 | Border Random | Random |
| 4 | 2 | 5 | 50 | 3 | 6 | 12500 | 5000 | 2 | Sliced | Random |
| 4 | 2 | 5 | 50 | 3 | 6 | 12500 | 5000 | 2 | Forbidden | Random |
| 4 | 2 | 5 | 50 | 3 | 6 | 12500 | 5000 | 2 | Border Random | Centroids |
| 4 | 2 | 5 | 50 | 3 | 6 | 12500 | 5000 | 2 | Sliced | Centroids |
| 4 | 2 | 5 | 50 | 3 | 6 | 12500 | 5000 | 2 | Forbidden | Centroids |
| 4 | 2 | 5 | 50 | 3 | 6 | 12500 | 5000 | 2 | Border Random | Quadrants |
| 4 | 2 | 5 | 50 | 3 | 6 | 12500 | 5000 | 2 | Sliced | Quadrants |
| 4 | 2 | 5 | 50 | 3 | 6 | 12500 | 5000 | 2 | Forbidden | Quadrants |

Table 6.1. Summary of the benchmark tests

## 6.2 Small-sized instances results

In this section, we report the results obtained by solving to optimality all the 66 instances of the first set (12 customers and 2 satellites). The objective function values are reported in Table 6.2. The table contains, in the first column, the customer's number of the VRP instance E-n13-k4 where the satellites are placed. Column 2 reports the value of the optimum. Column 3 contains the percentage variation of the optimum of the 2E-CVRP compared to the optimum of the VRP instance. Column 4 shows the mean value of the accessibility index (58) computed on each satellite as

$$A_k = \frac{1}{|V_c|} \sum_{i \in V_c} \frac{d_i}{d_{max}} e^{-\beta \frac{c_{ki} - c^2_{min}}{c^2_{max} - c^2_{min}}}, \tag{6.1}$$

where $d_i$ is the demand of the customer $i$, $d_{max}$ the maximum demand overall the customers, $c_{ki}$ the transportation cost between the satellite $k$ and the customer $i$, $c^2_{min}$ and $c^2_{max}$ the minimum and maximum values of the transportation costs at the 2nd-level, respectively, and $\beta > 0$ is a given parameter (we have assumed $\beta = 0.1$). Finally, Column 5 reports the mean normalized transportation cost of the satellites with respect to the depot, where the normalized transportation cost of each satellite $k$ is given by:

$$\overline{c}_k = 100 \frac{c_{0k} - c^1_{min}}{c^1_{max} - c^1_{min}}, \tag{6.2}$$

where $c_{0k}$ is the transportation cost between the depot and the satellite $k$ and $c^1_{min}$ and $c^1_{max}$ are the minimum and maximum values of the transportation costs of the 1st-level. In the following we discuss advantages and disadvantages of the proposed two-level distribution system, by considering all the pairs of customers as possible satellite location and comparing the results with the optimal solution of the original VRP instance with optimum 290.

From the results, it is clear the benefit of using the 2E-CVRP distribution model instead of the VRP one. Indeed, the former is able to achieve a smaller cost in 45 instances, while the decreasing/increasing of the costs is, except for satellites 11,12 with +38%, in the range $[-25\%, +25\%]$ of the corresponding VRP instance. The mean decrease in the 45 instances with a reduced transportation cost is 11.33%, which can be used to balance the costs due to the loading/unloading operations at the satellites. In the city logistics field, this means that the 2E-CVRP distribution model can be introduced without increasing the transportation cost, and obtaining indirect advantages, such as the reduction of the traffic flows and pollution level.

| Satellites | OPT | Variation (%) | Mean acc | Mean transp. Cost |
|---|---|---|---|---|
| 1,2 | 262 | 6.07 | 62 | 6 |
| 1,3 | 274 | 10.93 | 66 | 14 |
| 1,4 | 274 | 10.93 | 50 | 16 |
| 1,5 | 208 | -15.79 | 57 | 15 |
| 1,6 | 212 | -14.17 | 76 | 19 |
| 1,7 | 220 | -10.93 | 88 | 27 |
| 1,8 | 214 | -13.36 | 102 | 31 |
| 1,9 | 226 | -8.50 | 89 | 34 |
| 1,10 | 234 | -5.26 | 101 | 38 |
| 1,11 | 258 | 4.45 | 67 | 48 |
| 1,12 | 266 | 7.69 | 64 | 50 |
| 2,3 | 262 | 6.07 | 52 | 20 |
| 2,4 | 262 | 6.07 | 36 | 22 |
| 2,5 | 208 | -15.79 | 43 | 21 |
| 2,6 | 212 | -14.17 | 62 | 24 |
| 2,7 | 218 | -11.74 | 74 | 33 |
| 2,8 | 214 | -13.36 | 88 | 37 |
| 2,9 | 230 | -6.88 | 75 | 40 |
| 2,10 | 234 | -5.26 | 87 | 44 |
| 2,11 | 256 | 3.64 | 53 | 53 |
| 2,12 | 262 | 6.07 | 50 | 56 |
| 3,4 | 278 | 12.55 | 40 | 30 |
| 3,5 | 218 | -11.74 | 48 | 29 |
| 3,6 | 226 | -8.50 | 67 | 33 |
| 3,7 | 226 | -8.50 | 79 | 41 |
| 3,8 | 228 | -7.69 | 93 | 45 |
| 3,9 | 244 | -1.21 | 79 | 48 |
| 3,10 | 236 | -4.45 | 91 | 52 |
| 3,11 | 256 | 3.64 | 57 | 62 |
| 3,12 | 266 | 7.69 | 54 | 64 |
| 4,5 | 218 | -11.74 | 31 | 31 |
| 4,6 | 226 | -8.50 | 50 | 35 |
| 4,7 | 228 | -7.69 | 63 | 43 |

| Satellites | OPT | Variation (%) | Mean acc | Mean transp. Cost |
|---|---|---|---|---|
| 4,8 | 228 | -7.69 | 76 | 48 |
| 4,9 | 244 | -1.21 | 63 | 50 |
| 4,10 | 236 | -4.45 | 75 | 55 |
| 4,11 | 254 | 2.83 | 41 | 64 |
| 4,12 | 258 | 4.45 | 38 | 66 |
| 5,6 | 210 | -14.98 | 58 | 34 |
| 5,7 | 210 | -14.98 | 70 | 42 |
| 5,8 | 214 | -13.36 | 84 | 47 |
| 5,9 | 218 | -11.74 | 70 | 49 |
| 5,10 | 218 | -11.74 | 82 | 53 |
| 5,11 | 218 | -11.74 | 48 | 63 |
| 5,12 | 218 | -11.74 | 45 | 65 |
| 6,7 | 234 | -5.26 | 89 | 45 |
| 6,8 | 230 | -6.88 | 103 | 50 |
| 6,9 | 230 | -6.88 | 89 | 52 |
| 6,10 | 230 | -6.88 | 101 | 57 |
| 6,11 | 230 | -6.88 | 67 | 66 |
| 6,12 | 230 | -6.88 | 64 | 69 |
| 7,8 | 234 | -5.26 | 115 | 58 |
| 7,9 | 246 | -0.40 | 101 | 60 |
| 7,10 | 240 | -2.83 | 114 | 65 |
| 7,11 | 246 | -0.40 | 80 | 74 |
| 7,12 | 246 | -0.40 | 77 | 77 |
| 8,9 | 254 | 2.83 | 115 | 65 |
| 8,10 | 254 | 2.83 | 127 | 70 |
| 8,11 | 254 | 2.83 | 93 | 79 |
| 8,12 | 254 | 2.83 | 90 | 81 |
| 9,10 | 270 | 9.31 | 114 | 72 |
| 9,11 | 274 | 10.93 | 80 | 81 |
| 9,12 | 274 | 10.93 | 77 | 84 |
| 10,11 | 274 | 10.93 | 92 | 86 |
| 10,12 | 274 | 10.93 | 89 | 88 |
| 11,12 | 308 | 24.70 | 55 | 98 |

Table 6.2. 12 customers and 2 satellites instances: detailed results

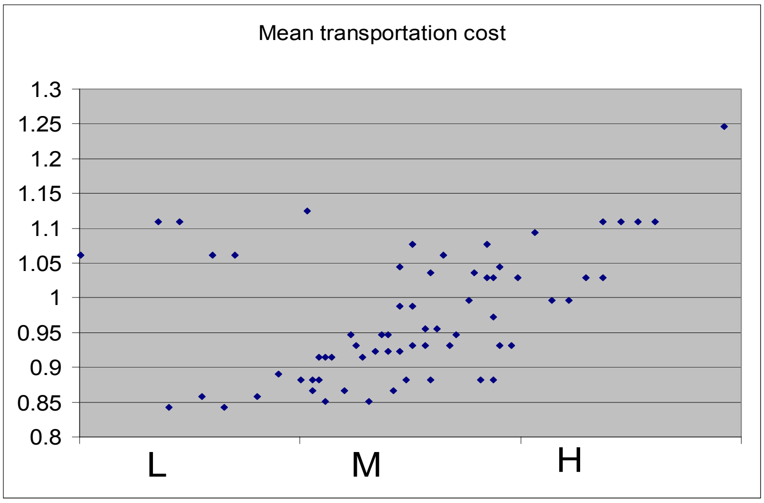


Figure 6.1. 12 customers instances: dispersion of the travelling costs

In Figure 6.1 we report the dispersion of the optima of the 66 instances with respect to the mean transportation cost from the depot to the satellites. These costs have been categorized in three sets: low (L), medium (M) and high (H) as:

- Low: mean transportation cost of the satellites in the interval [0,50];
- Medium: mean transportation cost of the satellites in the interval [50,67];
- High: mean transportation cost of the satellites in the interval [67,100].

On the $X$ axis the mean transportation cost is reported, while on $Y$ we report the ratio between the optimum of the 2E-CVRP instance and the optimum of the VRP instance. Thus, a ratio greater than 1 means that the optimum of the 2E-CVRP instance is worse than the VRP one.

According to the figure, it is clear that the instances with an optimum better than the VRP are characterized by a low mean transportation cost from the depot to the satellites. The greater the mean transportation cost the less likely to obtain an improved optimum. On the other hand, it is possible to obtain a gain even with a high mean transportation cost, which means that the mean transportation cost from the depot to the satellites is not the only parameter to be taken into account.

In Figure 6.2 we show the dispersion of the optima of the 66 instances with respect to the mean accessibility index of the satellites. The mean transportation cost is split into three sets: low (L), medium (M) and high (H) accessibility as follows:

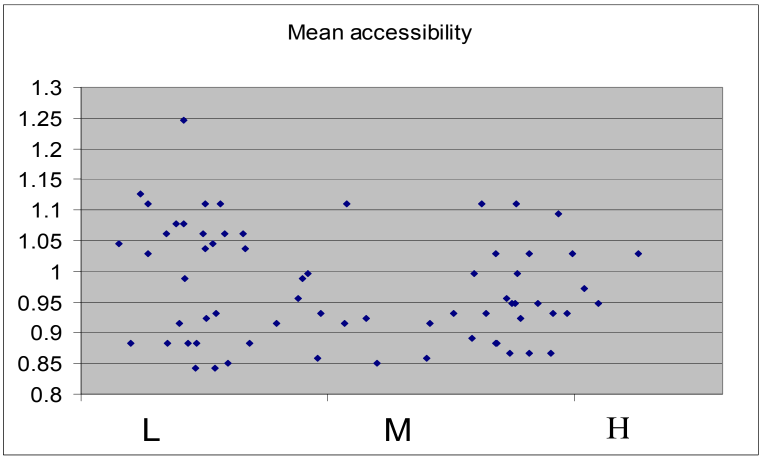


Figure 6.2. 12 customers instances: dispersion of the accessibility index

Mean accessibility

| | | | |
|---|---|---|---|
| H | 9/0 | 5/2 | 0/6 |
| M | 10/3 | 3/0 | 1/4 |
| L | 9/5 | 2/6 | 0/1 |
| | L | M | H |

Mean transportation cost

Table 6.3. 12 customers instances: resume of the results

- Low: mean accessibility in the interval [$A_{min}$,33% of [$A_{min}$,$A_{max}$]];
- Medium: mean accessibility in the interval [33% of [$A_{min}$,$A_{max}$],66% of [$A_{min}$,$A_{max}$]];
- High: mean accessibility in the interval [66% of [$A_{min}$,$A_{max}$],$A_{max}$];

where $A_{min} = \min_k\{A_k\}$ and $A_{max} = \max_k\{A_k\}$.

The $X$ axis shows the mean accessibility index, while $Y$ indicates the ratio between the optimum of the 2E-CVRP instance and the optimum of the VRP.

According to the figure, when accessibility increases the number of the instances of the 2E-CVRP with a gain does increase. However, even in the instances with a high accessibility, it is possible to have a deterioration of the optimum.

Table 6.3 presents a summary of the instances: the accessibility values are given in the rows and the transportation cost in the columns. Each cell contains the number of instances with an objective function better than the VRP and the number of instances with an objective function which is worse. The table shows that 2E-CVRP gives its best results when the mean transportation cost of the satellites is less than 50% of the maximum transportation cost (low transportation cost), with the ratio between gain and loss decreasing while the accessibility index decreases. When the mean transportation cost is medium, the costs of using the satellites are lower than VRP with a medium accessibility, which means that the quality of the result is mainly related to the dispersion of the customers themselves, while with a low accessibility it is difficult to obtain a gain on the total costs. With a high mean transportation cost, it becomes hard to obtain a lower transportation cost, even in presence of a high accessibility index. This is mainly due to the fact that even if the satellites are placed in the neighborhood of the customers, they are usually near the border of the customers' area, so the trasportation cost paid in the 1st level to reach the satellites is not compensated by the gain due to the proximity of the satellites to the customers and the consequent reduction of the 2nd-level fleet routes.

## 6.3 Valid inequalities: computational results

In this section we present the computational results of the first and the second set of instances for 2E-CVRP using the valid inequalities introduced in Setion 5.3 within a computation time limit of 10000 seconds.

With respect to the edge cuts, a series of tests was carried out using a simple procedure testing all the subtours up to cardinality 5. The procedure, coded in Mosel, has been tested on the instances of the sets 1, 2, and 3. According to the results, the subtours of cardinality greater than 3 are ineffective for the quality of both lower bounds and final solution. As the edge cuts of cardinality up to 3 are $O(n^3)$, we tested the model directly, adding them to the model at the root node, using a procedure to remove those cuts which are ineffective after five levels of the search tree.

In table 6.4 the results of the 66 instances corresponding to the problem with 12 customers and 2 satellites are given. The optimum is reported in the second column, while columns 3 and 4 contain the time in seconds needed to solve the instances without and with the valid inequalities introduced in Section 5.3. Finally, the last column presents the percentage of decreasing/increasing of computational time due to the usage of the valid inequalities.

According to the results most instances are solved in less than one minute, and only 10 of them need more than 2 minutes to be solved. There are however seven instances for which the computational times are greater than 10 minutes. This gap is mostly related to the satellite location. In fact, the greatest computational times are related to the situation where choosing which satellite to use has little or no effect on the final solution. In this situation, the model finds an optimal solution quickly, but spends a lot of time closing the nodes of the decision tree due to the poor quality of the lower bound obtained by the continuous relaxation of the model. Better behavior is obtained with the valid inequalities. As a counter effect, on some instances, the computational time still increases, but this is mainly due to the fact that the management of the additional inequalities can affect the computational times on small-sized instances, which show a rather small computational time without the cuts.

The results on the second set of instances are presented in Tables 6.5 and 6.6.

Table 6.5 presents the behavior of the lower bound computed with a continuous relaxation of the model found without and with the valid inequalities. More precisely, columns 1 to 4 contain, respectively, the number of customers in the original Christofides and Eilon's instances, the position of the satellites given as customer number, the mean accessibility as defined by (6.1) and the mean transportation cost of the satellites computed according to (6.2). The values and the gap with the best lower bound of the first lower bound (calculated at the root node) without and with the valid inequalities are reported in columns 5-8, while the final lower bound (calculated at the end of the optimization process), increased by letting the solver apply lift-and-project cuts during the optimization, and its gap are presented in columns 9-13. The last column summarizes the best lower bound obtained for each instance (bold values mean optimal values).

From these results it can be seen that the use of the cuts helps the model to reduce the gap by up to 26%. The behavior is confirmed by considering, in Table 6.6, the values of the feasible solutions found by the model without and with the valid inequalities. More precisely, columns 1 to 6 contain, respectively, the name of the original Christofides and Eilon's instance, the position of the satellites given as customer number, the mean accessibility as defined by (6.1), the mean transportation cost of the satellites computed according to (6.2), the best solution found and the best lower bound. The other columns contain respectively the values and the gap with the best lower bound of the first feasible solution, the best solutions after 100, 1000 and 5000 seconds, and the best solution. For each column, the results without and with the cuts are given.

According to these results, for up to 32 customers the model is able to find good quality solutions in 5000 seconds at most. When the number of customers increases to 50, more than 5000 seconds are required to find a good solution. Moreover, the use of the cuts increases the average model quality in terms of the initial solutions and the lower bounds. The gaps between the best solutions and the best bounds are quite small for instances involving up to 32 customers, but increase for 50-customer instances, with a gap up to 54% for the 4 satellite case.

| Satellites | Cost | Time Without cuts | Time With cuts | % Time |
|---|---|---|---|---|
| 1,2 | 262 | 8507 | 3590.9 | -57.79% |
| 1,3 | 274 | 3337.2 | 915.9 | -72.55% |
| 1,4 | 274 | 3274.3 | 3922.1 | 19.78% |
| 1,5 | 208 | 16.4 | 6.2 | -62.20% |
| 1,6 | 212 | 9.6 | 10.3 | 7.29% |
| 1,7 | 220 | 24.7 | 15.4 | -37.65% |
| 1,8 | 214 | 8.5 | 4.4 | -48.24% |
| 1,9 | 226 | 13.8 | 9.2 | -33.33% |
| 1,10 | 234 | 24.1 | 20.2 | -16.18% |
| 1,11 | 258 | 83.3 | 178.1 | 113.81% |
| 1,12 | 266 | 348.2 | 375.7 | 7.90% |
| 2,3 | 262 | 1526.4 | 749.5 | -50.90% |
| 2,4 | 262 | 693.6 | 940.2 | 35.55% |
| 2,5 | 208 | 8.4 | 5.4 | -35.71% |
| 2,6 | 212 | 11.1 | 7.1 | -36.04% |
| 2,7 | 218 | 10.1 | 10.4 | 2.97% |
| 2,8 | 214 | 10.2 | 6 | -41.18% |
| 2,9 | 230 | 13.1 | 14.9 | 13.74% |
| 2,10 | 234 | 25.7 | 21.5 | -16.34% |
| 2,11 | 256 | 207.5 | 109.8 | -47.08% |
| 2,12 | 262 | 672.6 | 197.5 | -70.64% |
| 3,4 | 278 | 5393.8 | 3307 | -38.69% |
| 3,5 | 218 | 5.2 | 6.3 | 21.15% |
| 3,6 | 226 | 23.4 | 11.3 | -51.71% |
| 3,7 | 226 | 8.7 | 9 | 3.45% |
| 3,8 | 228 | 7.4 | 7.8 | 5.41% |
| 3,9 | 244 | 13.2 | 15 | 13.64% |
| 3,10 | 236 | 28.5 | 23.1 | -18.95% |
| 3,11 | 256 | 12.1 | 11.6 | -4.13% |
| 3,12 | 266 | 40.3 | 15.6 | -61.29% |
| 4,5 | 218 | 8.6 | 6 | -30.23% |
| 4,6 | 226 | 30.4 | 28.5 | -6.25% |
| 4,7 | 228 | 16.3 | 14 | -14.11% |

| Satellites | Cost | Time Without cuts | Time With cuts | % Time |
|---|---|---|---|---|
| 4,8 | 228 | 7 | 4.2 | -40.00% |
| 4,9 | 244 | 14.6 | 14.2 | -2.74% |
| 4,10 | 236 | 18.3 | 23.2 | 26.78% |
| 4,11 | 254 | 18.5 | 19.2 | 3.78% |
| 4,12 | 258 | 15.1 | 15.4 | 1.99% |
| 5,6 | 210 | 10.5 | 6.3 | -40.00% |
| 5,7 | 210 | 4.6 | 5.2 | 13.04% |
| 5,8 | 214 | 12.7 | 10.9 | -14.17% |
| 5,9 | 218 | 9.2 | 5.1 | -44.57% |
| 5,10 | 218 | 14.4 | 5.6 | -61.11% |
| 5,11 | 218 | 5.7 | 2.6 | -54.39% |
| 5,12 | 218 | 2.6 | 3 | 15.38% |
| 6,7 | 234 | 4.8 | 4.6 | -4.17% |
| 6,8 | 230 | 23.4 | 13.3 | -43.16% |
| 6,9 | 230 | 14.3 | 16.4 | 14.69% |
| 6,10 | 230 | 32.9 | 11 | -66.57% |
| 6,11 | 230 | 11.8 | 13.5 | 14.41% |
| 6,12 | 230 | 11 | 7.2 | -34.55% |
| 7,8 | 234 | 6.1 | 9.6 | 57.38% |
| 7,9 | 246 | 95.4 | 56.9 | -40.36% |
| 7,10 | 240 | 12.4 | 14.1 | 13.71% |
| 7,11 | 246 | 42.4 | 33.4 | -21.23% |
| 7,12 | 246 | 19.9 | 19.2 | -3.52% |
| 8,9 | 254 | 64.1 | 37.5 | -41.50% |
| 8,10 | 254 | 43.5 | 37.2 | -14.48% |
| 8,11 | 254 | 8.4 | 6 | -28.57% |
| 8,12 | 254 | 12 | 12.1 | 0.83% |
| 9,10 | 270 | 286.2 | 148.3 | -48.18% |
| 9,11 | 274 | 31.5 | 23.3 | -26.03% |
| 9,12 | 274 | 23.5 | 12.2 | -48.09% |
| 10,11 | 274 | 84.1 | 75.9 | -9.75% |
| 10,12 | 274 | 64.3 | 73.5 | 14.31% |
| 11,12 | 308 | 234.8 | 78.5 | -66.57% |

Table 6.4. 12 customers and 2 satellites instances: valid inequalities improvements

| CVRP Instance | Satellites | Accessibility | Distance | First Bound: Without cuts Bound | Gap | First Bound: With cuts Bound | Gap | Final Bound: Without cuts Best Bound | Gap | Final Bound: With cuts Best Bound | Gap | Best Bound Best Sol. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| E-n13-k4 | 2,3 | 52 | 20 | 184.13 | -29.72% | 228.69 | -12.72% | 262 | 0.00% | 262 | 0.00% | **262.00** |
| | 2,5 | 43 | 21 | 151.84 | -27.00% | 158.23 | -23.93% | 208 | 0.00% | 208 | 0.00% | **208.00** |
| | 3,6 | 67 | 33 | 155.12 | -31.36% | 182.73 | -19.15% | 226 | 0.00% | 226 | 0.00% | **226.00** |
| | 4,8 | 76 | 48 | 158.37 | -30.54% | 212.69 | -6.72% | 228 | 0.00% | 228 | 0.00% | **228.00** |
| | 5,10 | 82 | 53 | 201.42 | -7.60% | 201.76 | -7.45% | 218 | 0.00% | 218 | 0.00% | **218.00** |
| | 7,9 | 101 | 60 | 166.92 | -32.15% | 171.37 | -30.34% | 246 | 0.00% | 246 | 0.00% | **246.00** |
| E-n22-k4 | 7,18 | 69 | 41 | 374.27 | -10.26% | 403.86 | -3.17% | 417.07 | 0.00% | 417.07 | 0.00% | **417.07** |
| | 9,15 | 30 | 31 | 338.12 | -12.17% | 360.16 | -6.44% | 384.96 | 0.00% | 384.96 | 0.00% | **384.96** |
| | 10,20 | 31 | 30 | 403.89 | -14.14% | 425.95 | -9.45% | 470.42 | 0.00% | 470.42 | 0.00% | **470.42** |
| | 11,15 | 38 | 25 | 332.02 | -10.63% | 349.89 | -5.82% | 371.50 | 0.00% | 371.50 | 0.00% | **371.50** |
| | 12,13 | 45 | 53 | 354.05 | -17.13% | 387.03 | -9.41% | 427.22 | 0.00% | 427.22 | 0.00% | **427.22** |
| | 13,17 | 33 | 64 | 334.34 | -14.88% | 355.73 | -9.43% | 392.78 | 0.00% | 392.78 | 0.00% | **392.78** |
| E-n33-k4 | 2,10 | 28 | 41 | 567.30 | -19.35% | 613.01 | -12.85% | 668.63 | -4.94% | 703.39 | 0.00% | 703.39 |
| | 3,14 | 19 | 31 | 593.12 | -12.66% | 609.36 | -10.27% | 665.35 | -2.02% | 679.08 | 0.00% | 679.08 |
| | 4,18 | 31 | 30 | 580.23 | -11.65% | 605.81 | -7.75% | 637.09 | -2.99% | 656.74 | 0.00% | 656.74 |
| | 5,6 | 28 | 25 | 61.18 | -91.59% | 633.37 | -12.96% | 710.06 | -2.43% | 727.71 | 0.00% | 727.71 |
| | 8,26 | 30 | 53 | 615.51 | -16.20% | 632.68 | -13.87% | 699.74 | -4.74% | 734.54 | 0.00% | 734.54 |
| | 15,23 | 18 | 64 | 626.80 | -16.60% | 663.38 | -11.74% | 717.27 | -4.57% | 751.59 | 0.00% | 751.59 |
| E-n51-k5 | 3.18 | 38 | 41 | 530.43 | -2.57% | 524.29 | -3.70% | 527.36 | -3.13% | 544.41 | 0.00% | 544.41 |
| | 5.47 | 48 | 18 | 489.86 | -2.89% | 499.81 | -0.92% | 495.14 | -1.85% | 504.44 | 0.00% | 504.44 |
| | 7,13 | 36 | 18 | 487.18 | -4.64% | 502.69 | -1.61% | 494.88 | -3.14% | 510.91 | 0.00% | 510.91 |
| | 12,20 | 32 | 47 | 534.95 | -2.27% | 539.46 | -1.44% | 539.07 | -1.52% | 547.37 | 0.00% | 547.37 |
| | 28,48 | 37 | 16 | 490.65 | -2.52% | 495.33 | -1.59% | 490.66 | -2.51% | 503.32 | 0.00% | 503.32 |
| | 33,38 | 29 | 28 | 488.94 | -5.61% | 507.99 | -1.93% | 502.80 | -2.93% | 517.99 | 0.00% | 517.99 |
| E-n51-k5 | 3,5,18,47 | 43 | 29 | 422.45 | -11.60% | 468.84 | -1.89% | 459.01 | -3.95% | 477.87 | 0.00% | 477.87 |
| | 7,13,33,38 | 32 | 23 | 432.68 | -10.28% | 471.37 | -2.25% | 445.88 | -7.54% | 482.24 | 0.00% | 482.24 |
| | 12,20,28,48 | 34 | 31 | 446.96 | -8.18% | 466.89 | -4.09% | 462.27 | -5.03% | 486.78 | 0.00% | 486.78 |

Table 6.5. Lower bounds for the instances E-n13-k4, E-n22-k4, E-n33-k4 and E-n51-k5

| CVRP Instance | Satellites | Accessibility | Distance | Final solution Best Sol. | Best Bound | First Solution: Without cuts Solution | Gap | First Solution: With cuts Solution | Gap | Solution after 100 s: Without cuts Solution | Gap | Solution after 100 s: With cuts Solution | Gap |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| E-n13-k4 | 2,3 | 52 | 20 | **262** | **262.00** | 328.00 | 25.19% | 310.00 | 18.32% | 262.00 | 0.00% | 262.00 | 0.00% |
| | 2,5 | 43 | 21 | **208** | **208.00** | 292.00 | 40.38% | 228.00 | 9.62% | 208.00 | 0.00% | 208.00 | 0.00% |
| | 3,6 | 67 | 33 | **226** | **226.00** | 242.00 | 7.08% | 230.00 | 1.77% | 226.00 | 0.00% | 226.00 | 0.00% |
| | 4,8 | 76 | 48 | **228** | **228.00** | 276.00 | 21.05% | 254.00 | 11.40% | 228.00 | 0.00% | 228.00 | 0.00% |
| | 5,10 | 82 | 53 | **218** | **218.00** | 238.00 | 9.17% | 338.00 | 55.05% | 218.00 | 0.00% | 218.00 | 0.00% |
| | 7,9 | 101 | 60 | **246** | **246.00** | 392.00 | 59.35% | 386.00 | 56.91% | 246.00 | 0.00% | 246.00 | 0.00% |
| E-n22-k4 | 7,18 | 69 | 41 | **417.07** | **417.07** | 426.75 | 2.32% | 435.54 | 4.43% | 420.14 | 0.74% | 417.07 | 0.00% |
| | 9,15 | 30 | 31 | **384.96** | **384.96** | 521.77 | 35.54% | 691.53 | 79.64% | 446.57 | 16.01% | 392.42 | 1.94% |
| | 10,20 | 31 | 30 | **470.42** | **470.42** | 841.17 | 78.81% | 1027.18 | 118.35% | 505.33 | 7.42% | 507.86 | 7.96% |
| | 11,15 | 38 | 25 | **371.50** | **371.50** | 754.83 | 103.19% | 631.89 | 70.09% | 378.42 | 1.86% | 403.17 | 8.53% |
| | 12,13 | 45 | 53 | **427.22** | **427.22** | 909.98 | 113.00% | 782.06 | 83.06% | 451.93 | 5.78% | 482.46 | 12.93% |
| | 13,17 | 33 | 64 | **392.78** | **392.78** | 552.39 | 40.63% | 603.67 | 53.69% | 408.66 | 4.04% | 431.65 | 9.89% |
| E-n33-k4 | 2,10 | 28 | 41 | 749.36 | 703.39 | 1078.47 | 53.32% | 1148.15 | 63.23% | 1078.47 | 53.32% | 844.21 | 20.02% |
| | 3,14 | 19 | 31 | 751.74 | 679.08 | 954.72 | 40.59% | 1060.75 | 56.20% | 954.72 | 40.59% | 818.52 | 20.53% |
| | 4,18 | 31 | 30 | 729.91 | 656.74 | 903.22 | 37.53% | 1287.94 | 96.11% | 903.22 | 37.53% | 937.39 | 42.74% |
| | 5,6 | 28 | 25 | 851.78 | 727.71 | 1065.18 | 46.37% | 951.83 | 30.80% | 854.36 | 17.40% | 873.40 | 20.02% |
| | 8,26 | 30 | 53 | 766.94 | 734.54 | 956.46 | 30.21% | 1359.81 | 85.12% | 932.53 | 26.95% | 1359.81 | 85.12% |
| | 15,23 | 18 | 64 | 787.31 | 751.59 | 1048.22 | 39.47% | 1198.34 | 59.44% | 1048.22 | 39.47% | 1017.29 | 35.35% |
| E-n51-k5 | 3.18 | 38 | 41 | 644.96 | 544.41 | 1103.07 | 102.62% | 1166.36 | 114.24% | No solution found | | No solution found | |
| | 5.47 | 48 | 18 | 621.00 | 504.44 | 998.23 | 97.89% | No solution found | | No solution found | | No solution found | |
| | 7,13 | 36 | 18 | 672.84 | 510.91 | 1018.26 | 99.30% | 1134.54 | 122.06% | No solution found | | No solution found | |
| | 12,20 | 32 | 47 | 643.25 | 547.37 | 835.94 | 52.72% | 814.91 | 48.88% | No solution found | | No solution found | |
| | 28,48 | 37 | 16 | 619.19 | 503.32 | 876.10 | 74.06% | 1166.87 | 131.83% | No solution found | | 1166.87 | 131.83% |
| | 33,38 | 29 | 28 | 606.84 | 517.99 | 984.21 | 90.01% | 606.84 | 17.15% | 984.21 | 90.01% | No solution found | |
| E-n51-k5 | 3,5,18,47 | 43 | 29 | 700.63 | 477.87 | 778.37 | 62.88% | 700.63 | 46.61% | No solution found | | No solution found | |
| | 7,13,33,38 | 32 | 23 | 744.73 | 482.24 | 1110.92 | 130.37% | 1204.20 | 149.71% | No solution found | | No solution found | |
| | 12,20,28,48 | 34 | 31 | 775.70 | 486.78 | 1163.71 | 139.06% | 1240.07 | 154.75% | No solution found | | No solution found | |

| CVRP Instance | Satellites | Accessibility | Distance | Final solution Best Sol. | Best Bound | Solution after 1000 s: Without cuts Solution | Gap | Solution after 1000 s: With cuts Solution | Gap | Solution after 5000 s: Without cuts Solution | Gap | Solution after 5000 s: With cuts Solution | Gap | Best solution: Without cuts Solution | Gap | Best solution: With cuts Solution | Gap |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| E-n13-k4 | 2,3 | 52 | 20 | **262** | 262.00 | 262.00 | 0.00% | 262.00 | 0.00% | 262.00 | 0.00% | 262.00 | 0.00% | 262.00 | 0.00% | 262.00 | 0.00% |
| | 2,5 | 43 | 21 | **208** | 208.00 | 208.00 | 0.00% | 208.00 | 0.00% | 208.00 | 0.00% | 208.00 | 0.00% | 208.00 | 0.00% | 208.00 | 0.00% |
| | 3,6 | 67 | 33 | **226** | 226.00 | 226.00 | 0.00% | 226.00 | 0.00% | 226.00 | 0.00% | 226.00 | 0.00% | 226.00 | 0.00% | 226.00 | 0.00% |
| | 4,8 | 76 | 48 | **228** | 228.00 | 228.00 | 0.00% | 228.00 | 0.00% | 228.00 | 0.00% | 228.00 | 0.00% | 228.00 | 0.00% | 228.00 | 0.00% |
| | 5,10 | 82 | 53 | **218** | 218.00 | 218.00 | 0.00% | 218.00 | 0.00% | 218.00 | 0.00% | 218.00 | 0.00% | 218.00 | 0.00% | 218.00 | 0.00% |
| | 7,9 | 101 | 60 | **246** | 246.00 | 246.00 | 0.00% | 246.00 | 0.00% | 246.00 | 0.00% | 246.00 | 0.00% | 246.00 | 0.00% | 246.00 | 0.00% |
| E-n22-k4 | 7,18 | 69 | 41 | **417.07** | 417.07 | 417.07 | 0.00% | 417.07 | 0.00% | 417.07 | 0.00% | 417.07 | 0.00% | 417.07 | 0.00% | 417.07 | 0.00% |
| | 9,15 | 30 | 31 | **384.96** | 384.96 | 386.94 | 0.51% | 384.96 | 0.00% | 384.96 | 0.00% | 384.96 | 0.00% | 384.96 | 0.00% | 384.96 | 0.00% |
| | 10,20 | 31 | 30 | **470.42** | 470.42 | 472.71 | 0.49% | 476.81 | 1.36% | 471.18 | 0.16% | 470.42 | 0.00% | 470.42 | 0.00% | 470.42 | 0.00% |
| | 11,15 | 38 | 25 | **371.50** | 371.50 | 371.50 | 0.00% | 371.50 | 0.00% | 371.50 | 0.00% | 371.50 | 0.00% | 371.50 | 0.00% | 371.50 | 0.00% |
| | 12,13 | 45 | 53 | **427.22** | 427.22 | 432.37 | 1.21% | 433.46 | 1.46% | 427.22 | 0.00% | 427.22 | 0.00% | 427.22 | 0.00% | 427.22 | 0.00% |
| | 13,17 | 33 | 64 | **392.78** | 392.78 | 408.66 | 4.04% | 392.78 | 0.00% | 392.78 | 0.00% | 392.78 | 0.00% | 392.78 | 0.00% | 392.78 | 0.00% |
| E-n33-k4 | 2,10 | 28 | 41 | 749.36 | 703.39 | 798.85 | 13.57% | 783.64 | 11.41% | 774.73 | 10.14% | 752.80 | 7.02% | 761.39 | 8.25% | 749.36 | 6.54% |
| | 3,14 | 19 | 31 | 751.74 | 679.08 | 888.47 | 30.83% | 803.97 | 18.39% | 758.80 | 11.74% | 803.97 | 18.39% | 751.74 | 10.70% | 769.96 | 13.38% |
| | 4,18 | 31 | 30 | 729.91 | 656.74 | 808.39 | 23.09% | 852.36 | 29.79% | 808.39 | 23.09% | 729.91 | 11.14% | 780.17 | 18.79% | 729.91 | 11.14% |
| | 5,6 | 28 | 25 | 851.78 | 727.71 | 854.36 | 17.40% | 873.40 | 20.02% | 851.78 | 17.05% | 866.30 | 19.04% | 851.78 | 17.05% | 866.30 | 19.04% |
| | 8,26 | 30 | 53 | 766.94 | 734.54 | 932.53 | 26.95% | 1061.66 | 44.53% | 819.16 | 11.52% | 779.99 | 6.19% | 804.14 | 9.47% | 766.94 | 4.41% |
| | 15,23 | 18 | 64 | 787.31 | 751.59 | 856.04 | 13.90% | 792.02 | 5.38% | 816.71 | 8.66% | 788.88 | 4.96% | 816.71 | 8.66% | 787.31 | 4.75% |
| E-n51-k5 | 3.18 | 38 | 41 | 644.96 | 544.41 | No solution found | | 742.07 | 36.31% | 1103.07 | 102.62% | 644.96 | 18.47% | 1021.88 | 87.70% | 644.96 | 18.47% |
| | 5.47 | 48 | 18 | 621.00 | 504.44 | No solution found | | No solution found | | 930.73 | 84.51% | 621.27 | 23.16% | 930.73 | 84.51% | 621.27 | 23.16% |
| | 7,13 | 36 | 18 | 672.84 | 510.91 | 861.92 | 68.70% | No solution found | | 861.92 | 68.70% | 812.28 | 58.99% | 816.19 | 59.75% | 672.84 | 31.70% |
| | 12,20 | 32 | 47 | 643.25 | 547.37 | No solution found | | 814.91 | 48.88% | 835.94 | 52.72% | 797.93 | 45.78% | 835.94 | 52.72% | 643.25 | 17.52% |
| | 28,48 | 37 | 16 | 619.19 | 503.32 | No solution found | | 1166.87 | 131.83% | No solution found | | 619.19 | 23.02% | 876.10 | 74.06% | 619.19 | 23.02% |
| | 33,38 | 29 | 28 | 606.84 | 517.99 | 861.50 | 66.32% | 606.84 | 17.15% | 861.50 | 66.32% | 606.84 | 17.15% | 861.50 | 66.32% | 606.84 | 17.15% |
| E-n51-k5 | 3,5,18,47 | 43 | 29 | 700.63 | 477.87 | No solution found | | 700.63 | 46.61% | 778.37 | 62.88% | 700.63 | 46.61% | 727.20 | 52.17% | 700.63 | 46.61% |
| | 7,13,33,38 | 32 | 23 | 744.73 | 482.24 | 1107.04 | 129.56% | 744.73 | 54.43% | 767.34 | 59.12% | 744.73 | 54.43% | 767.34 | 59.12% | 744.73 | 54.43% |
| | 12,20,28,48 | 34 | 31 | 775.70 | 486.78 | 1106.09 | 127.23% | No solution found | | 808.42 | 66.08% | 955.71 | 96.33% | 775.70 | 59.36% | 713.26 | 46.53% |

Table 6.6. Solutions for the instances E-n13-k4, E-n22-k4, E-n33-k4 and E-n51-k5

## 6.4 Overall computational results on medium instances

In this section we present the results of the tests in the sets 2, 3 and 4. All the results have been obtained using the model with all the valid inequalities activated. The results of each set are summarized in Tables 6.7, 6.8 and 6.9 respectively. Each table contains the instance name, the number of satellites the satellite distribution and the customer distribution in Columns 1, 2, 3 and 4. The mean accessibility as defined by (6.1) and the mean transportation cost of the satellites computed according to (6.2) are presented in Columns 5 and 6. Columns 7 and 8 contain the best solution and the lower bound computed by continuous relaxation of the model. Finally, the percentage gap between the best solution and the lower bound in presented in Column 9.

These results indicate that the gap is quite small up to 32 customers, while increases in the 50-customer tests. In particular, the gap is quite large in tests in set 2 involving 4 satellites.

The instances generated from the classical VRP instances present a distribution of the customers which is quite different from the distribution in realistic applications in urban and regional delivery. Moreover, the model is able to find solutions with an average gap of 12%. This is quite large, but understandable considering that the lower bounds come from the simple continuous relaxation of the model with cuts, and that the original 50-customer instance is still considered a difficult one for Branch & Cut and Branch & Bound algorithms developed for VRP.

The quality of the solutions diminishes as the number of satellites increases, even if this is probably due to the poor quality of the lower bound.

A better insight into the performance obtainable with 2E-CVRP and its model can be seen in the instances of set 4. These instances present different distributions, simulating different strategies. According to the results, the model seems to present an almost constant gap around 25%. This can be easily noticed considering the aggregated results of set 4 presented in Tables 6.10a, 6.10b and 6.10c. The tables contain in each cell the mean of the gaps between the best solution and the lower bounds grouped by satellites and customers distribution in Table 6.10a, number of satellites and customers distribution in Table 6.10b and number of satellites and satellites distribution in Table 6.10c. According to the tables, the best results are obtained when using the *Centroids* distribution for the customers and the *Forbidden* for the satellites. This result is not surprising, as it is easier for the model, in the case of the *Centroid* distribution, to

| Instance | Satellites | Satellite distribution | Customer distribution | Mean acc. | Mean transp. Cost | Final Solution | Best Bound | Gap |
|---|---|---|---|---|---|---|---|---|
| E-n22-k4-s6-17 | 2 | Random | E-n22-k4 | 76.04 | 40.96 | **417.07** | **417.07** | 0.00% |
| E-n22-k4-s8-14 | 2 | Random | E-n22-k4 | 35.33 | 31.38 | **384.96** | **384.96** | 0.00% |
| E-n22-k4-s9-19 | 2 | Random | E-n22-k4 | 36.29 | 30.03 | **470.60** | **470.60** | 0.00% |
| E-n22-k4-s10-14 | 2 | Random | E-n22-k4 | 42.06 | 24.91 | **371.50** | **371.50** | 0.00% |
| E-n22-k4-s11-12 | 2 | Random | E-n22-k4 | 50.62 | 52.98 | **427.22** | **427.22** | 0.00% |
| E-n22-k4-s12-16 | 2 | Random | E-n22-k4 | 38.21 | 64.25 | **392.78** | **392.78** | 0.00% |
| E-n33-k4-s1-9 | 2 | Random | E-n33-k4 | 29.70 | 40.96 | 731.21 | 706.80 | 3.45% |
| E-n33-k4-s2-13 | 2 | Random | E-n33-k4 | 20.04 | 31.38 | 742.64 | 681.80 | 8.92% |
| E-n33-k4-s3-17 | 2 | Random | E-n33-k4 | 32.05 | 30.03 | 744.21 | 656.95 | 13.28% |
| E-n33-k4-s4-5 | 2 | Random | E-n33-k4 | 29.78 | 24.91 | 860.47 | 722.68 | 19.07% |
| E-n33-k4-s7-25 | 2 | Random | E-n33-k4 | 31.05 | 52.98 | 766.11 | 736.37 | 4.04% |
| E-n33-k4-s14-22 | 2 | Random | E-n33-k4 | 19.59 | 64.25 | 780.17 | 754.63 | 3.38% |
| E-n51-k5-s2-17 | 2 | Random | E-n51-k5 | 44.42 | 40.55 | 636.45 | 543.66 | 17.07% |
| E-n51-k5-s4-46 | 2 | Random | E-n51-k5 | 51.88 | 17.95 | 570.10 | 512.31 | 11.28% |
| E-n51-k5-s6-12 | 2 | Random | E-n51-k5 | 40.62 | 17.98 | 592.78 | 517.86 | 14.47% |
| E-n51-k5-s11-19 | 2 | Random | E-n51-k5 | 35.97 | 47.34 | 639.52 | 551.87 | 15.88% |
| E-n51-k5-s27-47 | 2 | Random | E-n51-k5 | 42.21 | 15.54 | 570.33 | 506.72 | 12.55% |
| E-n51-k5-s32-37 | 2 | Random | E-n51-k5 | 33.86 | 28.35 | 692.77 | 521.51 | 32.84% |
| E-n51-k5-s2-4-17-46 | 4 | Random | E-n51-k5 | 48.15 | 29.25 | 674.34 | 502.95 | 34.08% |
| E-n51-k5-s6-12-32-37 | 4 | Random | E-n51-k5 | 37.24 | 23.16 | 763.97 | 497.30 | 53.63% |
| E-n51-k5-s11-19-27-47 | 4 | Random | E-n51-k5 | 39.09 | 31.44 | 675.68 | 497.45 | 35.83% |

Table 6.7. Summary of the computational results of Set 2

| Instance | Satellites | Satellite distribution | Customer distribution | Mean acc. | Mean transp. Cost | Final Solution | Best Bound | Gap |
|---|---|---|---|---|---|---|---|---|
| E-n22-k4-s14-15 | 2 | Random | E-n22-k4 | 39.28 | 31.47 | **526.15** | **526.15** | 0.00% |
| E-n22-k4-s14-17 | 2 | Random | E-n22-k4 | 42.11 | 28.77 | **521.09** | **521.09** | 0.00% |
| E-n22-k4-s14-18 | 2 | Random | E-n22-k4 | 70.71 | 24.64 | **496.38** | **496.38** | 0.00% |
| E-n22-k4-s15-20 | 2 | Random | E-n22-k4 | 31.52 | 15.71 | 498.80 | 480.42 | 3.83% |
| E-n22-k4-s15-20 | 2 | Random | E-n22-k4 | 62.94 | 8.87 | 512.81 | 497.68 | 3.04% |
| E-n22-k4-s20-22 | 2 | Random | E-n22-k4 | 38.97 | 0.70 | 520.42 | 501.69 | 3.73% |
| E-n33-k4-s17-23 | 2 | Random | E-n33-k4 | 23.61 | 44.26 | 1401.43 | 1310.16 | 6.97% |
| E-n33-k4-s17-25 | 2 | Random | E-n33-k4 | 25.34 | 39.06 | 1399.95 | 1329.23 | 5.32% |
| E-n33-k4-s20-27 | 2 | Random | E-n33-k4 | 20.52 | 38.49 | 1708.41 | 1667.80 | 2.43% |
| E-n33-k4-s23-27 | 2 | Random | E-n33-k4 | 23.29 | 39.16 | 1716.74 | 1645.54 | 4.33% |
| E-n33-k4-s25-29 | 2 | Random | E-n33-k4 | 28.21 | 28.14 | 1605.09 | 1544.29 | 3.94% |
| E-n33-k4-s26-29 | 2 | Random | E-n33-k4 | 35.91 | 24.48 | 1585.87 | 1560.13 | 1.65% |
| E-n51-k5-s14-20 | 2 | Random | E-n51-k5 | 38.34 | 16.50 | 861.60 | 652.46 | 32.05% |
| E-n51-k5-s14-43 | 2 | Random | E-n51-k5 | 39.07 | 19.35 | 795.78 | 642.75 | 23.81% |
| E-n51-k5-s14-45 | 2 | Random | E-n51-k5 | 36.66 | 25.35 | 860.43 | 633.98 | 35.72% |
| E-n51-k5-s41-43 | 2 | Random | E-n51-k5 | 31.54 | 9.02 | 811.50 | 651.23 | 24.61% |
| E-n51-k5-s42-43 | 2 | Random | E-n51-k5 | 29.40 | 15.98 | 851.80 | 648.27 | 31.39% |
| E-n51-k5-s42-45 | 2 | Random | E-n51-k5 | 27.00 | 21.98 | 819.59 | 649.22 | 26.24% |

Table 6.8. Summary of the computational results of Set 3

find the optimal assignment of the customers to the satellites. Moreover, this id the distribution which better represents the case of urban areas. In any case, the behavior of the model is quite good in the case of the *Quadrants* distribution. Indeed, in this case the mean gaps are almost constant independently of the distribution of the satellites.

The increasing gap between best solution and lower bound is still remarkable with the *Random* and the *Quadrants* distributions of the customers (see Table 6.10c). Even in the case of the satellite distribution, the *Random* distribution is the most sensitive to the number of satellites, while the other two distributions present a better behavior.

| Instance | Satellites | Satellite distribution | Customers distribution | Mean acc. | Mean transp. cost | Final Solution | Best Bound | Gap |
|---|---|---|---|---|---|---|---|---|
| Instance50-1 | 2 | Random | Random | 41.58 | 49.86 | 1639 | 1528.70 | 7.22% |
| Instance50-2 | 2 | Random | Random | 16.73 | 46.83 | 1561 | 1353.79 | 15.31% |
| Instance50-3 | 2 | Sliced | Random | 41.75 | 50.86 | 1642 | 1519.42 | 8.07% |
| Instance50-4 | 2 | Sliced | Random | 30.86 | 73.25 | 1542 | 1358.93 | 13.47% |
| Instance50-5 | 2 | Forbidden | Random | 58.17 | 99.46 | 2386 | 2092.61 | 14.02% |
| Instance50-6 | 2 | Forbidden | Random | 30.04 | 55.82 | 1387 | 1245.03 | 11.40% |
| Instance50-7 | 2 | Random | Centroids | 60.63 | 48.83 | 1643 | 1401.18 | 17.26% |
| Instance50-8 | 2 | Random | Centroids | 56.65 | 46.79 | 1427 | 1235.68 | 15.48% |
| Instance50-9 | 2 | Sliced | Centroids | 46.33 | 49.47 | 1568 | 1411.62 | 11.08% |
| Instance50-10 | 2 | Sliced | Centroids | 62.83 | 75.04 | 1462 | 1290.29 | 13.31% |
| Instance50-11 | 2 | Forbidden | Centroids | 86.13 | 97.60 | 2124 | 1963.24 | 8.19% |
| Instance50-12 | 2 | Forbidden | Centroids | 61.74 | 56.99 | 1250 | 1134.10 | 10.22% |
| Instance50-13 | 2 | Random | Quadrants | 32.54 | 50.63 | 1608 | 1433.29 | 12.19% |
| Instance50-14 | 2 | Random | Quadrants | 53.23 | 51.44 | 1481 | 1272.76 | 16.36% |
| Instance50-15 | 2 | Sliced | Quadrants | 48.24 | 50.08 | 1550 | 1447.92 | 7.05% |
| Instance50-16 | 2 | Sliced | Quadrants | 35.09 | 75.79 | 1457 | 1311.28 | 11.11% |
| Instance50-17 | 2 | Forbidden | Quadrants | 40.82 | 99.48 | 2233 | 2004.98 | 11.37% |
| Instance50-18 | 2 | Forbidden | Quadrants | 34.29 | 60.85 | 1285 | 1163.96 | 10.40% |
| Instance50-19 | 3 | Random | Random | 70.88 | 49.81 | 1757 | 1525.44 | 15.18% |
| Instance50-20 | 3 | Random | Random | 35.46 | 55.16 | 1329 | 1180.33 | 12.60% |
| Instance50-21 | 3 | Sliced | Random | 73.99 | 40.12 | 1701 | 1478.39 | 15.06% |
| Instance50-22 | 3 | Sliced | Random | 36.74 | 56.50 | 1399 | 1236.39 | 13.15% |
| Instance50-23 | 3 | Forbidden | Random | 71.53 | 73.41 | 1898 | 1725.37 | 10.01% |
| Instance50-24 | 3 | Forbidden | Random | 33.38 | 48.64 | 1420 | 1234.92 | 14.99% |
| Instance50-25 | 3 | Random | Centroids | 84.50 | 48.79 | 1884 | 1439.01 | 30.92% |
| Instance50-26 | 3 | Random | Centroids | 41.19 | 56.24 | 1210 | 1105.15 | 9.49% |
| Instance50-27 | 3 | Sliced | Centroids | 69.69 | 35.86 | 1776 | 1354.17 | 31.15% |
| Instance50-28 | 3 | Sliced | Centroids | 56.81 | 57.76 | 1287 | 1126.92 | 14.21% |
| Instance50-29 | 3 | Forbidden | Centroids | 42.49 | 70.61 | 2018 | 1628.20 | 23.94% |
| Instance50-30 | 3 | Forbidden | Centroids | 68.21 | 48.84 | 1386 | 1082.39 | 28.05% |
| Instance50-31 | 3 | Random | Quadrants | 53.07 | 49.95 | 1710 | 1361.36 | 25.61% |
| Instance50-32 | 3 | Random | Quadrants | 45.20 | 59.19 | 1322 | 1124.25 | 17.59% |
| Instance50-33 | 3 | Sliced | Quadrants | 67.47 | 42.94 | 1684 | 1417.86 | 18.77% |
| Instance50-34 | 3 | Sliced | Quadrants | 44.46 | 60.81 | 1297 | 1156.82 | 12.12% |
| Instance50-35 | 3 | Forbidden | Quadrants | 49.63 | 74.10 | 2003 | 1488.72 | 34.55% |
| Instance50-36 | 3 | Forbidden | Quadrants | 41.98 | 52.61 | 1306 | 1121.60 | 16.44% |
| Instance50-37 | 5 | Random | Random | 96.85 | 47.44 | 2164 | 1407.79 | 53.72% |
| Instance50-38 | 5 | Random | Random | 95.10 | 37.92 | 1181 | 932.09 | 26.70% |
| Instance50-39 | 5 | Sliced | Random | 106.74 | 57.66 | 2153 | 1415.89 | 52.06% |
| Instance50-40 | 5 | Sliced | Random | 63.08 | 54.30 | 1361 | 1049.37 | 29.70% |
| Instance50-41 | 5 | Forbidden | Random | 99.67 | 69.42 | 2761 | 1527.51 | 80.75% |
| Instance50-42 | 5 | Forbidden | Random | 61.19 | 52.31 | 1273 | 1116.03 | 14.07% |
| Instance50-43 | 5 | Random | Centroids | 40.95 | 46.47 | 2218 | 1311.02 | 69.18% |
| Instance50-44 | 5 | Random | Centroids | 45.21 | 40.44 | 1181 | 932.09 | 26.70% |
| Instance50-45 | 5 | Sliced | Centroids | 101.69 | 56.45 | 2129 | 1322.65 | 60.96% |
| Instance50-46 | 5 | Sliced | Centroids | 58.16 | 57.34 | 1217 | 937.46 | 29.82% |
| Instance50-47 | 5 | Forbidden | Centroids | 80.44 | 66.47 | 1837 | 1462.65 | 25.59% |
| Instance50-48 | 5 | Forbidden | Centroids | 93.72 | 52.15 | 1146 | 1002.98 | 14.26% |
| Instance50-49 | 5 | Random | Quadrants | 81.41 | 47.58 | 2235 | 1340.42 | 66.74% |
| Instance50-50 | 5 | Random | Quadrants | 78.21 | 41.52 | 1260 | 961.93 | 30.99% |
| Instance50-51 | 5 | Sliced | Quadrants | 96.07 | 57.92 | 2038 | 1279.08 | 59.33% |
| Instance50-52 | 5 | Sliced | Quadrants | 61.94 | 57.34 | 1182 | 969.35 | 21.94% |
| Instance50-53 | 5 | Forbidden | Quadrants | 84.17 | 70.21 | 2145 | 1439.85 | 48.97% |
| Instance50-54 | 5 | Forbidden | Quadrants | 69.05 | 53.59 | 1301 | 1026.49 | 26.74% |

Table 6.9. Summary of the computational results of Set 4

| Average (%) | Customers | | |
|---|---|---|---|
| Satellites | Centroids | Quadrants | Random |
| Forbidden | 18.38% | 22.79% | 24.21% |
| Random | 28.17% | 28.25% | 21.79% |
| Sliced | 26.75% | 21.72% | 21.92% |

(a)

| Average (%) | Satellites | | |
|---|---|---|---|
| Num Sats | Forbidden | Random | Sliced |
| 2 | 10.93% | 13.97% | 10.68% |
| 3 | 18.68% | 18.56% | 17.41% |
| 5 | 35.06% | 45.67% | 42.30% |

(b)

| Average (%) | Customers | | |
|---|---|---|---|
| Num Sats | Centroids | Quadrants | Random |
| 2 | 12.59% | 11.41% | 11.58% |
| 3 | 22.96% | 18.11% | 13.50% |
| 5 | 37.75% | 42.45% | 42.83% |

(c)

Table 6.10. Summary of the mean gap between the best solution and the lower bound in Set 4

# Chapter 7

# Solving the 2E-VRP

In this chapter, we will describe the procedure for solving the medium 2E-CVRP instances where the full model is not able to finish the optimization process in less than three hours of calculation, focusing on the difficulties which are added to the problem complexity because of the connexion between the two levels. To deal with these aspects, we chose to develop a Column Generation procedure to obtain Lower Bounds for the problem, reformulating it to obtain a set covering formulation which variables are related to routes and not to each single arc. In a first time, the approach and the models will be presented. Then, the procedure will be described, and the problems due to the connexion constraints will be analyzed, presenting the limits of the chosen methodology and the need of new alternative approaches to solve the problem. A last section will present the guidelines to develop alternative Column Generation approaches, and describe future work directions for constructing a Branch-and-prize exact algorithm.

## 7.1 Notation and formulations

In this section we present the set-covering and set-partitioning formulations for the 2E-CVRP. These formulations derive from those defined for the VRP, which were first introduced by (6) and have been adapted to many VRP variants. In these two formulations, the decision variables represent the usage or not of a route. We will then refer to these models as route formulations.

The 2E-CVRP route formulations are an extension to 2-echelon distribution of the CVRP route formulations. This optimization problem can be defined in the following way. Consider an oriented graph $G$ defined as follows. Let $V_0 = \{v_0\}$ be the depot, $V_s = \{v_{s_1},v_{s_2},...,v_{s_{n_s}}\}$ the set of satellites and $V_c = \{v_{c_1}, v_{c_2}, ..., v_{c_{n_c}}\}$ the set of customers. Associated to customer $c_i \in V_c$ is the demand $d_i$ which represents the load which must be delivered to customer $c_i$'s location. Two classes of vehicles are defined: 1st-level and 2nd-level vehicles. The number of 1st-level vehicles available at the depot is $m_1$. These vehicles have the same given capacity $K^1$. The total number of 2nd-level vehicles available is equal to $m_2$. We do not introduce a constraint on the number of vehicles available at each single satellite. The 2nd-level vehicles have the same given capacity $K^2$. From a physical point of view, a 2E-CVRP system operates as follows:

- The freight that have to be distributed to the customers is supposed to start at the depot, where it is consolidated into 1st-level vehicles.
- Each 1st-level vehicle travels to a subset of satellites that will be determined by the model and then it will return to the depot;
- At a satellite, freight is transferred from 1st-level vehicles to 2nd-level vehicles;
- Each 2nd-level vehicle performs a route to serve the designated customers, and then travels to a satellite for its next cycle of operations. The 2nd-level vehicles return to their departure satellite.

Let $t_{ij}$ denote the length of arc $(i,j)$ of the graph $G$. It is assumed that the costs $t_{ij}$satisfy the triangular inequality. These represent two types of arc costs:

- costs of the arcs traveled by 1st-level vehicles, i.e. arcs connecting the depot to the satellites and the satellites between them;
- costs of the arcs traveled by 2nd-level vehicles, i.e. arcs connecting the satellites to the customers and the customers between them.

Another cost that can be used is the cost of loading and unloading operations at the satellites. Supposing that the number of workers in each satellite $v_{s_k}$ is fixed, we consider only the cost due to the management of the freight and we define $S_k$ as the unit cost of freight handling at the satellite $v_{s_k}$. For the purpose of this study we will not consider these costs.

Enumerate all feasible 1st-level routes and all 2nd-level routes. Let the index set of each level feasible routes be respectively $R^1 = \{1,2,...,R^1\}$ and $R^2 = \{1,2,...,R^2\}$. Let $c_i$ be the cost of 1st-level route $i$ and $c_k j$ the cost of route $j$ which starts on satellite $k$. Define:

- $x_i$ is a {0,1} variable that represents the usage of 1st-level route $i$. $x_i$ is equal to 1 if route $i$ is used and 0 otherwise.

- $y_{jk}$ is a {0,1} variable that represents the usage of 2st-level route $j$, which starts on satellite $k$. $y_k j$ is equal to 1 if route $j$ is used and 0 otherwise.

In the same way, we define two sets of {0,1} attributes:

- $\delta^1_{ikc}$: indicates if the 1st-level route i uses the satellite k to serve the customer c

- $\delta^2_{jkc}$: indicates if the 2nd-level route j that starts from the satellite k serves the customer c

For ensuring each customer is serve by only one satellite, these attributes have to respect the following constraints:

$$\sum_k \delta^1_{ikc} \leq 1; \forall i \in R^1, c \in V_c \tag{7.1}$$

$$\sum_k \delta^2_{jkc} \leq 1; \forall j \in R^2, c \in V_c \tag{7.2}$$

| | |
|---|---|
| $V_0 = \{v_0\}$ | Depot |
| $V_s = \{v_{s_1}, v_{s_2}, ..., v_{s_{n_s}}\}$ | Set of satellites |
| $V_c = \{v_{c_1}, v_{c_2}, ..., v_{c_{n_c}}\}$ | Set of customers |
| $R^1 = \{1, 2, ..., R^1\}$ | Set of 1st-level routes |
| $R^2 = \{1, 2, ..., R^2\}$ | Set of 2nd-level routes |
| $n_s$ | number of satellites |
| $n_c$ | number of customers |
| $m_1$ | number of the 1st-level vehicles |
| $m_2$ | number of the 2nd-level vehicles |
| $K^1$ | capacity of the vehicles for the 1st level |
| $K^2$ | capacity of the vehicles for the 2nd level |
| $d_i$ | demand required by customer i |
| $t_{ij}$ | cost of the arc $(i,j)$ |
| $c_i$ | cost of the 1st-level route $i$ |
| $c_j$ | cost of the 2nd-level route $j$ |
| $S_k$ | cost for loading/unloading operations of a unit of freight in satellite $k$ |

Table 7.1. Definitions and notations

### 7.1.1 Set partitioning formulation

We can build a set partitioning formulation for the 2E-CVRPextending the model originally proposed by (6) . The model is:

$$\min \sum_i c_i x_i + \sum_k \sum_j c_{jk} y_{jk} \tag{7.3}$$

s.t.

$$\sum_{i,k} \delta^1_{ikc} x_i = 1 \qquad \forall c \tag{7.4}$$

$$\sum_{j,k} \delta^2_{jkc} y_{jk} = 1 \qquad \forall c \tag{7.5}$$

$$\sum_i \delta^1_{ikc} x_i - \sum_j \delta^2_{jkc} y_{jk} = 0 \qquad \forall k,c \tag{7.6}$$

$$x,y = \{0,1\} \tag{7.7}$$

Constraints 7.4 impose that customer $c$ is served by only one of the selected 1st-level routes, and constraints 7.5 impose that customer $c$ is served by only one of the selected 2st-level routes. Constraints7.6 require that a customer have to be served from the same satellite by first level and second level routes.

### 7.1.2 Set covering formulation

As seen in (13) , if the transportation costssatisfy the triangular inequality we can substitute the equality constraints 7.4 and 7.5 by inequality constraints, obtaining an equivalent model. The resulting set covering formulation for the 2E-CVRPis:

$$\min \sum_i c_i x_i + \sum_k \sum_j c_{jk} y_{jk} \tag{7.8}$$

s.t.

$$\sum_{i,k} \delta^1_{ikc} x_i \geq 1 \qquad \forall c \tag{7.9}$$

$$\sum_{j,k} \delta^2_{jkc} y_{jk} \geq 1 \qquad \forall c \tag{7.10}$$

$$\sum_i \delta^1_{ikc} x_i - \sum_j \delta^2_{jkc} y_{jk} = 0 \qquad \forall k,c \tag{7.11}$$

$$x,y = \{0,1\} \tag{7.12}$$

Constraints 7.9 require that customer $c$ is served at least one of the selected 1st-level routes, and constraints 7.10 impose that customer $c$ is served at least one of the selected 2st-level routes. Constraints7.11 require that a customer have to be served from the same satellite by first level and second level routes. Since in the linear relaxation, set covering models are in general easier to work for implementation than set partitioning formulations, we will realize our methodologies for the set covering model already defined.

## 7.2 Solving the linear relaxation of P

To solve the linear relaxation of P without enumerating all the routes, we can use the Column Generation technique. The general idea of the Column Generation is the following: a portion of all possible routes is enumerated, and the linear relaxation with this partial route set is solved. This problem in know as restricted Master Problem. The solution to this linear program is then used to determine if there are any routes not included in the formulation that can further reduce the objective function value. To do it, we use the values of the dual variables corresponding to the optimal solution of the restricted Master Problem to solve a simpler optimization problem, called subproblem, where we identify if there are one or more routes that should be included in the formulation. This is also called Column Generator. We add this columns to the restricted Master Problem and resolve it. This is continued until no additional routes are found that can reduce the objective function value. In that case, we can show that an optimal optimal solution to the linear program is found, it is also optimal for the complete route set. The most important part of the method arises in building a good Column Generator.

The linear relaxation of the problem, called P', is formulated as follows:

$$\min \sum_i c_i x_i + \sum_k \sum_j c_{jk} y_{jk} \tag{7.13}$$

s.t.

$$\sum_{i,k} \delta^1_{ikc} x_i \geq 1 \qquad \forall c \tag{7.14}$$

$$\sum_{j,k} \delta^2_{jkc} y_{jk} \geq 1 \qquad \forall c \tag{7.15}$$

$$\sum_i \delta^1_{ikc} x_i - \sum_j \delta^2_{jkc} y_{jk} = 0 \qquad \forall k{,}c \tag{7.16}$$

$$x_i \geq 0; \forall i \in R^1 \tag{7.17}$$

$$y_j \geq 0; \forall j \in R^2 \tag{7.18}$$

$$\tag{7.19}$$

We can associate a dual model to P'. Let the dual variables associated to the primal model be:

- $\alpha_c$: variables associated to constraints 7.14
- $\mu_c$: variables associated to constraints 7.15
- $\lambda_{kc}$: variables associated to constraints 7.16

The dual model can then be formulated as

$$\max \sum_c (\alpha_c + \mu_c) \tag{7.20}$$

s.t.

$$\sum_{k,c} \delta^1_{ikc} \alpha_c + \sum_{k,c} \delta^1_{ikc} \lambda_{kc} \leq c_i \forall i \tag{7.21}$$

$$\sum_c \delta^2_{jkc} \mu_c - \sum_c \delta^2_{jkc} \lambda_{kc} \leq c_{jk} \forall j \forall k \tag{7.22}$$

$$\alpha{,}\lambda{,}\mu \text{ free} \tag{7.23}$$

## 7.3 First method: decomposition

In a set covering formulation each column represents a feasible route. Our model has two types of variables: 1st-level routes and 2-level routes. One med for generating routes is to solve two separate subproblems, one for each level, and add the columns obtained in this way to the Master Problem.

From the Primal Model, using a subset R' of feasible routes, we obtain the restricted Master Problem(MP). Using an MP solver we can obtain the optimal solution to the restricted Master Problem and the associated dual variables. We know that given a Linear Program and a solution, can calculate the reduced costs. If all the reduced costs are non-negative, the solution is optimal. We can define the reduced costs in different ways. Having the dual variables we can write the reduced costs respectively as:

$$\bar{c}_i = c_i - \sum_{k,c} \delta^1_{ikc}(\alpha_c + \lambda_{kc}) \tag{7.24}$$

$$\bar{c}_{jk} = c_{jk} - \sum_{c} \delta^2_{jkc}(\mu_c - \lambda_{kc}) \tag{7.25}$$

If the minimum reduced cost is non negative, the solution is the optimum. If not, we add the columns corresponding to the routes that have a negative cost. We will iterate until we find the optimum. To obtain these negative reduced costs, we define the following subproblems:

***Subproblem 1***

The subproblem 1 is used to generate 1st-level columns. The objective function is to find the route with minimum reduced cost. A feasible route is a route that starts from the depot, goes to one or more satellites and return to the depot. The load of the considered 1st-level vehicle cannot exceed its capacity. The subproblem 1 can be defined as follows:

$$\min \bar{c}_i = c_i - \sum_{k,c} \delta_{ikc}(\alpha_c + \lambda_{kc}) \tag{7.26}$$

s.t.

$$\sum_{k,c} \delta^1_{ikc} d_c \leq K^1 \forall i \tag{7.27}$$

Constraint 7.27 is a Knapsack constraint which represents the capacity limitations of each route.

***Subproblem 2***

In the same way, we can define the subproblem 2 in order to generate 2nd-level columns. This subproblem is analogous to subproblem 1, but if formulated for each satellite k. Given a satellite $\bar{k}$, subproblem 2 can be formulated as follows:

$$\min \bar{c}_{j\bar{k}} = c_{j\bar{k}} - \sum_c \delta^2_{j\bar{k}c}(\mu_c - \lambda_{\bar{k}c}) \tag{7.28}$$

s.t.

$$\sum_c \delta^2_{i\bar{k}c} d_c \leq K^2 \forall j \tag{7.29}$$

Constraint 7.29 is a Knapsack constraint which represents the capacity limitations of each route.

***Methodology***

Consider a route r. The route cost $c_r$ can be decomposed in arc costs as already seen. For each subproblem, we have Knapsack constraints. Decomposing the problem into 2 separate subproblems, we do not have, at the Column Generation step, the constraint which connects both levels, but two separated subproblems which are similar to those we can obtain on CVRP problems. Similarly to what have been presented in (37) , we can solve solve our subproblems decomposing each route into shortest path problems with resource constraints and using Dynamic Programming to solve these shortest path problems. With this method the Column Generator step can be realized in pseudo-polinomial time. The method also allows to find more than one route with negative reduced cost.

For the 1st-level problem, the variables for the dynamic programming are the usage of an arc (and its relative satellite) in the route j (for $c_{ij}$: cost of the arc ij) and the assignment satellite-customer (dual variables $\mu c$ and $\lambda_{kc}$ and knapsack constraint). In this problem, the graph can be represented as follows. Each satellite is connected to another satellite by a macro-node defined by all the customers connected between them. The costs which are considered will be those corresponding to each customer associated to the 2nd-level routes which start at the considered satellite. We can see an example of how this graph is realized for one satellite in Figure 7.1. For the second level, the representation from the original graph is more direct: we create $n_s$ graphs, one for each satellite. On each of them, we duplicate the considered satellite into a starting and ending point. The starting point is connected to each customer, which is also connected to the rest of customers, and to the ending point, as shown in Figure 7.2

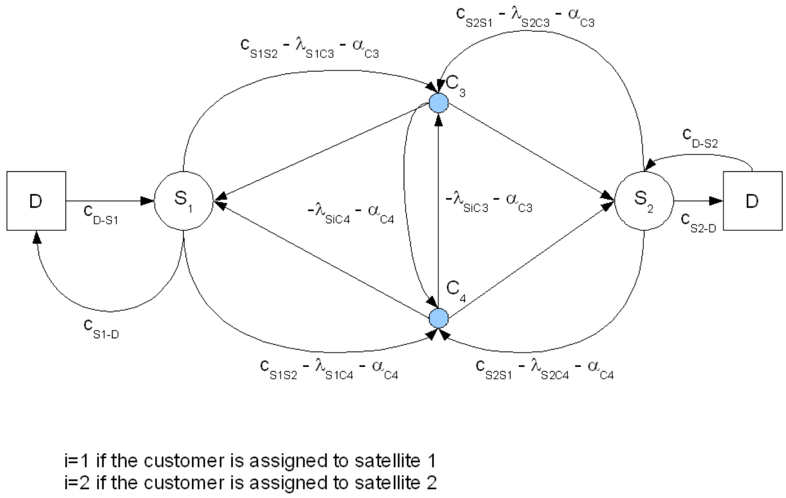


Figure 7.1. Example of graph for one satellite

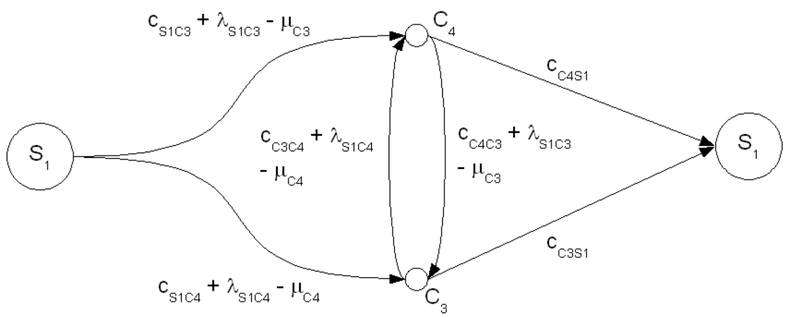


Figure 7.2. Example of graph for one 2nd-level route

This problem becomes a shortest path problem with resource limits (the vehicle's capacity is associated to each node). The resource constraints are Knapsack constraints. In both subproblems we have the costs which can be calculated as the sum of the corresponding arc costs, and the constraints are knapsack constraints. So, in order to enumerate a big number of columns with negative reduced cost, we can use shortest-path approaches to find routes with negative reduced cost.

We chose to adapt the algorithm of Desrochers, Desrosiers and Solomon (1994) for the second-level subproblem. Consider the graph in Figure 7.2. To be able to solve Subproblem 2 to use dynamic programming in pseudo-polinomial time, we modify it to allow routes which visit the same customer more than once. The Lower Bounds obtained with this modified problem are Lower Bounds of the not modified problem. Given a path $\Pi = \{0,u_1,u_2,...,u_l\}$, the total load of this path is defined as $q = sum_{i=1}^{l} d_i$. Consider a satellite $k$. Let $\{\bar{c}_{ij} : i,j \in V_c \cup k\}$ be the marginal cost between two nodes (i and j) of $V_c \cup k$. Let $f_j^k(q)$ be the cost of the least-cost path that starts at satellite $k$ and finishes at a customer $j$ with a total load of $q$ (this is a q-path). This can be calculated using the following recursion:

$$f_j^k(q) = \min_{j \neq i}[f_j^k(q - d_i) + \bar{c}_{ij}^{2k} | q \neq C^2], \tag{7.30}$$

$$\text{for all } j,q \text{ such that } j \in V_c, d_i \leq q \leq C^2. \tag{7.31}$$

$$\tag{7.32}$$

Let $f_j^{kr}(q) = f_j^k(q) + c_{jk}$.

To eliminate 2-cycles, we define $p_j^k(q)$ as the predecessor of $j$ in the path of cost $f_j^k(q)$. We also define $g_j^k(q)$ as the cost of the least-cost path that starts at satellite $k$ and finishes at a customer $j$ with a total load of $q$ and not having $p_j^k(q)$ as the last customer visited before $j$.

The algorithm works as follows:

- Initialization:

  $\forall k \in V_s, f_k^k(0) = 0 \; \forall k \in V_s, g_k^k(0) = 0$

  for $(e = 1; e \leq n_c; e++)$

- Search for a state (q) to be treated $(1 \leq q \leq C^2)$

- Treatment of state q: Calculation of all the $f_j^k(q)$, $g_j^k(q)$, $f_j^k r(q)$

  For all $f_j^k r(q) \leq 0$ add the corresponding column in the Restricted Master Problem.

***Implementation and limits of the methodology***

We have implemented the Column Generation method for solving the Linear relaxation of the Set Covering formulation, in C++ using CPLEX 10.1.1 as the LP solver for the Master problems. On the first iteration, The Master Problem is solved using the LP solver. The columns of $MP^0$ are obtained from a feasible solution of the problem (one of the first solutions obtained solving the Flow Model, but not the optimal solution) and we calculate the dual variables corresponding to the optimal solution of $MP^0$. Then, for each iteration, we proceed as follows:

- Subproblem 1 and Subproblem 2 are solved
- If the minimum reduced costs for each subproblems are non-negative, then the optimum is found and the algorithm stops. Otherwise, we identify the minimum reduced costs for each subproblems (or a group of negative reduced costs) and we add the corresponding columns to the Master Problem
- If we added columns, the new MP is solved using CPLEX, and we repeat all the process until no negative reduced costs are found.

When no negative reduce costs are found, the solution we obtain in optimal for P'. The optimal solution of the linear relaxation P' is a lower bound of the 2E-CVRP problem we presented on section 7.1.

We realized the first tests on instances of 21 customers, for which we know the optimal solution value. After the first tests we observe that on the first iterations a lot of columns are added, and used, then only few column per iteration are found. We also observe that because of our method, both subproblems are not connected, which makes the method introduce groups of columns that sometimes cannot be used (for example, if a group of 2nd-level columns are introduced but there is no 1st-level column which allow a combination of these routes to be used, these routes will be set to 0 until such a 1stevel route is introduced). This c be observed by a non variation of the 1st-level solution for some iterations. Another question is that, because of the connexion constraints, introducing a 1st-level column with non negative reduced cost can make possible to use 2ndlevel columns with negative reduced cost and the total reduced cost of the route configuration (first and second level) should be negative.

For allowing to solve this problem we need to introduce not separate routes but combinations of first and second level routes in order to reduce the number of routes that, given a route subset, we are not able to use because we do not have complementary routes to them. Different approaches have been proposed, but most of them generate columns in a heuristic way.

## 7.4 Second method: integral path

Because the decomposition method generates unconnected routes, we need a method which produces a set of complementary routes in order to accelerate the calculation times and avoid situations similar to what we have explained.

The general methodology will be the same as already explained, but the subproblems will be changed. In order to produce a group of complementary routes instead of separate routes, we will reformulate the graph to consider paths that contain both levels. In this modified graph the depot will be decomposed into a node Depot-departure and a node Depot-arrival. The nodes customer will remain the same but the nodes satellite are multiplied by the number of 2nd-level vehicles. These new nodes satellite will also be duplicated into an origin and a destination. The new graph, which is oriented, is organized as follows. The depot-departure is connected to the origin satellites $s_{iv}^{o}$". Each origin satellite is connected to the customers, which are connected among them and also connected to the destination satellites. Each destination satellite is connected to the other origin satellites and to the Depot-arrival. We can see an example on Figure 7.3. The routes of this modified graph are an aggregation of a 1st-level route and a number of second level routes compatible with the 1st-level route. We will call this aggregation mega-route.

htb

Once the modified graph is defined, we can use the labeling algorithm of (37) already described to find the mega-route of minimum reduced cost. This mega-route represents the aggregation of first and second level routes. The resulting subproblem is similar to the subproblem obtained for CVRP cases, with the difference that the reduced cost is more complex. The capacity constraints are Knapsack constraints which indicate that the sum of the loads of all the customers served by the mega-route have to be less or equal to the 1st-level capacity. In a first moment we will not add a 2nd-level capacity constraint to this subproblem because it will

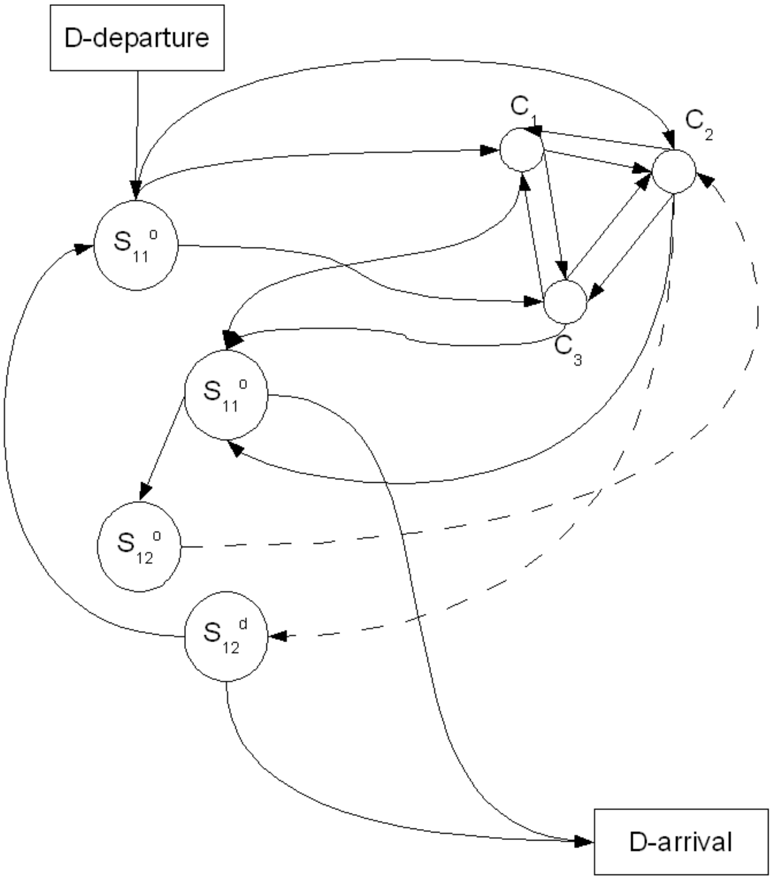


Figure 7.3. Example of graph for one satellite

increase considerably its complexity.

After solving the subproblem, we have to convert the mega-route in 1st-level routes and 2nd-level routes, and add the corresponding columns to the Master Problem. The procedure to separate the mega-route into routes that we can add to the Master Problem work as follows. In a first time, we build the 1st-level route, which will be only one for each selected mega-route. For doing it, we consider only the satellites of the mega-route, and we unify all the satellites from the modified graph that were created from the same satellite. All customers between an origin and destination satellite are associated to this satellite, defining its corresponding $\delta^1_{ikc}$. Then we obtain a route that starts from the depot and visit a succession of satellites, and we can calculate its cost. Associated to each satellite of the route we define the set of customers which be served from this satellite, $PC_k$ In a second time, for each satellite we have the associated customers and some routes that can be obtained from the mega-route by extrapolation. Each sub-path from $S^o_{iv}$ to $S^d_{iv}$ represents a possible 2nd-level route. These routes are not always respecting the 2nd-level capacity constraint.In the case a route respects this constraint, we can add it to the Master Problem. When this constraint is not respected, we have to convert this route into a group of feasible routes. In order to obtain an exact solution, we want to obtain the group of 2nd-level routes that minimize the total reduced cost. For doing it, we can solve consecutive Shortest Path Problems as already described. In a first time, we solve the following subproblem. Given a satellite $\bar{k}$, the subproblem to be solved can be formulated as:

$$\min \bar{c}_{j\bar{k}} = c_{j\bar{k}} - \sum_{c \in PC_k} \delta^2_{j\bar{k}c}(\mu_c - \lambda_{\bar{k}c}) \tag{7.33}$$

s.t.

$$\sum_{c \in PC_k} \delta^2_{i\bar{k}c} d_c \leq K^2 \forall j \tag{7.34}$$

Constraint 7.34 is a Knapsack constraint which represents the capacity limitations of each route. After finding the route with minimum reduced cost, we take out the customers corresponding to the route obtained and define a new subproblem with the set of the remaining customers. We solve it and repeat the process until all the customers of set $PC_k$ are assigned to a route. We add the routes obtained with this procedure to the Master problem.

In this method, the column generation is made in two steps:

- subproblem solving: the labeling algorithm of Desrochers, Desrosiers and Solomon (1994) is applied to the modified graph, in order to obtain a mega-route with minimum modified cost.

- route conversion: the obtained mega-route can then be converted into a combination of one 1st-level route and a set of compatible 2nd-level routes. In this procedure, a 1st-level route is directly obtained from the mega-route. The other information that can be obtained from it are the assignments satellite-customer corresponding to customers with which compatible 2nd-level routes can be obtained. With this information, the labeling algorithm can be applied to obtain a set of 2nd-level routes. Then, the 1st-level route and the set of 2nd-level routes are added to the Master Problem, and these steps are repeated until no negative reduced costs are found.

# Chapter 8

# Conclusions

In this thesis, the importance of multi-echelon distribution in real transport applications, and more specifically in City Logistics, have been presented. Nowadays, many combinatorial optimization problems deal with freight distribution, but no general methods which consider the entire system without separating it in a number of independent smaller problems have been presented. It's why we introduced a new family of combinatorial optimization problems, the Multi-echelon VRP.

In a first time we defined a basic city logistics system based using a hub and spoke strategy (based on UDCs) on a urban context, by observing the existing experiences in this domain. As we presented, in some cities, this approach can be and in fact is used, but other important aspects have to be considered, as the economic, environment (pollution and sound) and social costs, which are difficult to model in a mathematical way. Also the diseases produced by freight transportation have to be considered.

After defining a general multi-echelon distribution system and giving a synthetic survey con multi-echelon systems and route optimization problems, we have presented the new family of decision problems known as Multi-Echelon VRP, giving their generalities and main complex aspects, which are in general found at the connexion constraints between levels. Then, we focused on the 2-Echelon Capacitated VRP, which have been defined in detail, and two vehicle flow MIP formulations have been developed. One of them derives from multi-commodity network design, and is called Flow model because it uses the concept of freight flow (which represent the vehicle load) that travels each art. The other model derives from the classical vehicle flow formulation for the VRP, which is based on TSP concepts and uses subtour elimination constraints (which number depends exponentially from the number of satellites and customers). Also valid inequalities have been defined, specially for the flow model.

The flow model and the inequalities have been tested on new benchmarks derived from the CVRP instances of the literature, showing a good behavior of the model for small and medium sized instances. Moreover, a first attempt to find a priori conditions on the solution quality of 2E-CVRP has been performed, enabling the introduction of a classification of the instances according to the combination of easy-to-compute instance parameters, such as satellite accessibility and mean transportation cost.

The model limitations are evident, and for real applications, a quick and simple heuristic should be the best solution, due to the number of customers involved in these kind of application (in general from 150 to 500, or even a higher number of customers), even if the number of satellites will not be in general greater than 5 for city logistics applications. However, in other distribution applications (multimodal transport, Grocery and journal distribution) the number of satellites and customers can increase. In some applications, meta-heuristics can be used. To test these future heuristic methods, it is important to obtain exact solutions for the proposed instances. In order to build an exact method, which could be a Branch-and-Prize, we have developed a Column Generation method to obtain lower bounds. The main aspect of this method is that it allows to separate both levels in the column generation step. The problem is that the columns which are generated are not necessarily connected, and the calculation times will increase considerable. In order to deal with this aspect, another method, which considers a mega-route containing the routes of the first ans the second levels, is proposed, but has not yet been tested.

Future developments of this research can be grouped into two different areas. The first of them is the elaboration of an exact method to solve this problem, and will consist on building and testing an algorithm which applies the second method for obtaining Lower Bounds using Column Generation. Then, a Branch and Prize procedure can be realized. The second area is focused on heuristic methods. In a fist time, it seems more interesting for real applications to focus on quick local search heuristics for different reasons. The first of them is that they are more flexible than meta-heuristics, and even if the results they obtain are worse than those obtained with meta-heuristics, they can be applied to large-scale problems, and then can be adapted easily to deal with real time applications where the calculation times are more important than to obtain a near-optimal solution (to obtain a quite good solution is enough in these cases). Such heuristics can also be used in a second time to give initial solutions for exact and meta-heuristic methods.

# Bibliography

[1] A. Alfa, S. Heragu, M. Chen, A 3-opt based simulated annealing algorithm for vehicle routing problem, Computers and Industrial Engineering 21 (1991) 635–639.

[2] K. Altinkemer, B. Gavish, Parallel savings based heuristics for the delivery problem, Operations Research 39 (1991) 456–469.

[3] D. Ambrosino, M. Scutellà, Distribution network design: new problems and related models, European Journal of Operational Research 165 (2005) 610–624.

[4] E. Angelelli, M. Speranza, The periodic vehicle routing problem with intermediate facilities, European Journal of Operational Research 137 (2002) 233–247.

[5] B. C. Arntzen, G. G. Brown, T. P. Harrison, L. Trafton, Global supply chain management at digital equipment corporation, INTERFACES 25 (1995) 69–93.

[6] M. Balinski, R. Quandt, On an integer program for a delivery problem, Operations Research 12 (1964) 300–304.

[7] J. Bard, L. Huang, M. Dror, P. Jaillet, A branch and cut algorithm for the vrp with satellite facilities, IIE Transactions 30 (1998) 821–834.

[8] J. Bardi, J. Coyle, J. Langley, The Management of Business Logistics: A Supply Chain Perspective, Cengage Learning, 2000.

[9] B. M. Beamon, Supply chain design and analysis: Models and methods, International Journal of Production Economics 55 (3) (1998) 281–294.

[10] J. E. Beasley, Or-library: distributing test problems by electronic mail, Journal of the Operational Research Society 41 (11) (1990) 1069–1072.

[11] BESTUFS, Best urban freight solutions http://www.bestufs.net.

[12] J. Bramel, D. Simchi-Levi, A location based heuristic algorithm for general vehicle routing problems, Operations Research 43 (1995) 649–660.

[13] J. Bramel, D. Simchi-Levi, Set-covering based algorithms for the capacitated vrp, in: P. Toth, D. Vigo (Eds.), The Vehicle Routing Problem, Vol. 9 of SIAM Monographs on Discrete Mathematics and Applications, 2002 85–108.

[14] O. Braysy, M. Gendreau, Vehicle routing problem with time windows, part i: Route construction and local search algorithms, Transportation Science 39 (1) (2005) 104–118.

[15] O. Braysy, M. Gendreau, Vehicle routing problem with time windows, part ii: Metaheuristics, Transportation Science 39 (1) (2005) 119–139.

[16] A. V. Breedam, An analysis of the behavior of heuristics for vehicle routing problem for a selection of problems with vehicle-related, customer-related, and time-related constraints, Ph.D. thesis, University of Antwerp (1984).

[17] Bridgestone Corporation website http://www.bridgestone.com.

[18] J. Brunswicker, Optimale Standort- und Tourenplanung fur die Rohmilcherfassung eines Molkereibetriebes Lit, Munster, 1986.

[19] I. Chao, A tabu-search method for the truck and trailer routing problem, Computers and Operations Research 29 (2002) 33–51.

[20] N. Chistofides, A. Mingozzi, P. Toth, The vehicle routing problem, in: N. Chistofides, A. Mingozzi, P. Toth, C. Sandi (Eds.), Combinatorial Optimization, Wiley, Chichester, UK, 1979.

[21] N. Christofides, S. Elion, An algorithms for the vehicle dispatching problem, Operational Research Quarterly 20 (1969) 309–318.

[22] G. Clarke, J. Wright, Scheduling of vehicles from a central depot to a number of delivery points, Operations Research 12 (4) (1964) 568–581.

[23] J. Constantin, W. Hudson, Motor Transportation: Principles and Practices, The Ronal Press Company, New York, 1958.

[24] J. Cordeau, G. Desaulniers, J. Desrosiers, M. Solomon, F. Soumis, Vrp with time windows, in: P. Toth, D. Vigo (Eds.), The Vehicle Routing Problem, Vol. 9 of SIAM Monographs on Discrete Mathematics and Applications, 2002 157–193.

[25] J. Cordeau, M. Gendreau, A. Hertz, G. Laporte, J. Sormany, New heuristics for the vehicle routing problems, Tech. Rep. Les Cahiers du GERARD, G-2004-33 (2004).

[26] J. Cordeau, M. Gendreau, G. Laporte, A tabu search heuristic for periodic and multi-depot vehicle routing problems, Networks 30 (1997) 105–119.

[27] J. Cordeau, G. Laporte, A. Mercier, A tabu search heuristic for vehicle routing problems with time windows, Journal of the Operations Research Society 52 (2001) 928–936.

[28] J. Cordeau, G. Laporte, A. Mercier, Improved tabu search heuristic for the handling of route duration constraints in vehicle routing problems with time windows, Journal of the Operations Research Society 55 (2004) 542–546.

[29] J. Cordeau, G. Laporte, M. Savelsbergh, D. Vigo, Vehicle Routing, in: C. Barnhart, G. Laporte (Eds.), Transportation, North Holland Elsevier, 2007 367–428.

[30] T. Crainic, G. Laporte, Planning models for freight transportation, European Journal of Operational Research 97 (1997) 409–438.

[31] T. Crainic, N. Ricciardi, G. Storchi, Advanced freight transportation systems for congested urban areas, Transportation Research part C 12 (2004) 119–137.

[32] B. Crevier, J. Cordeau, G. Laporte, The multi-depot vehicle routing problem with inter-depot routes, European Journal of Operational Research 176 (2007) 756–773.

[33] B. Dalla Chiara, D. Marigo, G. Benzo, Interporti e terminali intermodali, Hoepli, Milan, 2002.

[34] M. S. Daskin, C. Coullard, Z.-J. M. Shen, An inventory-location model: Formulation, solution algorithm and computational results, Annals of Operations Research 110 (2002) 83–106.

[35] F. Deflorio, B. Dalla Chiara, J. Gonzalez-Feliu, Autostrada ferroviaria e trasporto stradale lungo la direttrice del frejus: analisi di scenari mediante un modello di rete, Ingegneria Ferroviaria 10.

[36] M. Desrochers, T. Verhoog, A matching based savings algorithm for the vehicle routing problem, Tech. Rep. Les Cahiers du GERARD, G-89-04 (1989).

[37] J. Desrosiers, F.Soumis, M. Desrochers, Routing with time windows by column generation, Networks 14 (1984) 545–565.

[38] M. Dror, G. Laporte, P. Trudeau, Vehicle routing with split deliveries, Discrete Applied Mathematics 50 (1994) 239–254.

[39] G. Dueck, New optimization heuristics: An analog approach to the traveling salesman problem using an elastic method, Journal of Computational Physics 104 (1993) 86–92.

[40] G. Dueck, T. Scheurer, Threshold acceptance: A general purpose optimization algorithm, Journal of Computational Physics 90 (1990) 161–175.

[41] W. Dullaert, O. Braysy, M. Goetschalckx, B. Raa, Supply chain (re)design: Support for managerial and policy decisions, European Journal of Transport and Infrastructure Research 7 (2) (2007) 73–92.

[42] European Commission, Cost 321 action - urban goods http://cordis.europa.eu/cost–transport/src/cost–321.htm.

[43] European Commission - Transportation General Directorade, COST 321. Urban Goods Transport. Final report of the Action, European Commission, Bruxelles, Belgium, 1998.

[44] European Commission - Transportation General Directorade, Energy and Transportation report 2000-2004, European Commission, Bruxelles, Belgium, 2004.

[45] E. Feldman, F. Lehrer, T. Ray, Warehouse location under continuous economies of scale, Management Science 12 (1966) 670–684.

[46] A. Fernandez, B. Arrunada, M. G. Diez, Quasi-integration in less-than-truckload trucking, in: Institutions, Contracts and Organizations, Edward Elgar, Cheltenham and Northampton, editor = C. Ménard,, 2000 293–312.

[47] M. Fisher, R. Jaikumar, A generalzed assignment heuristic for the vehicle routing problem, Networs 11 (1981) 109–124.

[48] B. Foster, L. Miller, An integer programming approach to the vehicle scheduling problem, Operations Research 27 (1976) 367–384.

[49] P. Frosini, Huntingford, G. Ambrosino, Urban mobility and freight distribution service: best practices and lessons learnt in the merope interreg iii b project, European Transport 28 (2005) 44–56.

[50] M. Gendreau, A. Hertz, G. Laporte, A tabu search heuristic for the vehicle routing problem, Management Science 40 (1994) 1276–1290.

[51] M. Gendreau, G. Laporte, J. Potvin, Metaheuristics for the capacitated vrp, in: P. Toth, D. Vigo (Eds.), The Vehicle Routing Problem, SIAM Monographs of Discrete Mathematics and Applications, Philadelphia, 2002 129–154.

[52] J. Gerdessen, Vehicle routing problem with trailers, European Journal of Operational Research 93 (1996) 135–147.

[53] B. Gillett, L. Miller, A heuristic algorithm for the vehicle dispatch problem, Operations Research 22 (1974) 340–349.

[54] M. Goetschalckx, C. J. Vidal, K. Dogan, Modeling and design of global logistics systems: A review of integrated strategic and tactical models and design algorithms, European Journal of Operational Research 143 (2002) 1–18.

[55] B. Golden, E. Wasil, J. Kelly, I. Chao, Metaheuristics in vehicle routing, in: T. Crainic, G. Laporte (Eds.), Fleet Management and Logistics, Kluwer, Boston, 1998 33–56.

[56] T. Goldman, R. Gorham, Sustainable urban transport: Four innovative directions, Technology in Society 28 (2006) 261–273.

[57] E. Hadjiconstantinou, N. Cristofides, A. Mingozzi, A new exact algorithm for the vehicle routing problem based on q-paths and k-shortest paths relaxations, Annals of Operations Research 61 (1995) 21–43.

[58] W. G. Hansen, How accessibility shapes land use, Journal of the American Institute Planners 25 (1959) 73–76.

[59] Y. Hinojosa, J. Puerto, F. R. Fernández, A multiperiod two-echelon multicommodity capacitated plant location problem, European Journal of Operational Research 23 (2003) 271–291.

[60] A. Hoff, A. Lokketangen, A tabu search approach for milk collection in western norway using trucks and trailers, Tech. Rep. Technical Report, Molde University College (2006).

[61] Institute of City Logistics http://www.citylogistics.org.

[62] S. Jacobsen, O. Madsen, A comparative study of heuristics for a two-level routing-location problem, European Journal of Operational Research 5 (1980) 378–387.

[63] V. Jayaraman, H. Pirkul, Simulated annealing methodology to distribution network design and management, European Journal of Operational Research 144 (2003) 629–645.

[64] U. Kohler, How to change urban traffic with city logistics, schnieder e., in becker u.(ed.), In: Control in Transportation Systems 2000 - Proceedings 9th IFAC Symposium, Braunschweig, Germany, vol. 1 (1994) 199–201.

[65] G. Laporte, Location-routing problems, in: A. A. B.L. Golden (Ed.), Vehicle routing: methods and studies, North-Holland: Amsterdam, 1988 163–197.

[66] G. Laporte, The vehicle routing problem: an overview of exact and approximate algorithms, European Journal of Operational Research 59 (1992) 345–358.

[67] G. Laporte, Y. Nobert, P. Pelletier, Hamiltonian location problems, European Journal of Operational Research 12 (1983) 82–89.

[68] G. Laporte, Y. Norbert, M. Desrochers, Optimal routing under capacity and distance restrictions, Operations Research 33 (1985) 1050–1073.

[69] G. Laporte, F. Semet, Classical heuristics for the capacitated vrp, in: P. Toth, D. Vigo (Eds.), The Vehicle Routing Problem, SIAM Monographs of Discrete Mathematics and Applications, Philadelphia, 2002.

[70] F. Li, B. Golden, E. Wasil, Very large-scale vehicle routing: new test problems, algorithms and results, Computers and Operations Research 32 (2005) 1165–1179.

[71] S. Lin, Computer solution of the traveling salesman problem, Bell System Technical Journal 44 (1965) 2245–2269.

[72] A. Marin, Lower bounds for the two-stage uncapacitated facility location problem, European Journal of Operational Research 179 (2007) 1126–1142.

[73] D. Mester, O. Bräysy, Active guided evolution strategies for the large scale vehicle routing problems with time windows, Computers and Operations Research 32 (6) (2005) 1593–1614.

[74] H. Min, V. Jayaraman, R. Srivastava, Combined location-routing problems: A synthesis and future research directions, European Journal of Operational Research 108 (1997) 1–15.

[75] Ministeére de l'Équipement, Division transport des marchandises en ville, France http://www.tmv.transports.equipement.gouv.fr.

[76] P. Moscato, C. Cotta, A gentle introduction to memetic algorithms, in: F. Glover, G. Kochenberger (Eds.), Handbook of Metaheuristics, Kluwer, Boston, 2003 105–144.

[77] G. Nagy, S. Salhi, Heuristic algorithms for simple and multiple depot vehicle routing problems with pickup and deliveries, European Journal of Operational Research 162 (2005) 126–141.

[78] G. Nagy, S. Salhi, Location routing: Issues, models and methods, European Journal of Operational Research 177 (2007) 649–672.

[79] Y. Nobert, D. Arpin, An exact algorithm for solving a capacitated location-routing problem, Annals of Operations Research 6 (1986) 293310.

[80] J. Ortúzar, L. Willmussen, Modeling Transport, Wiley, London, UK, 2001.

[81] I. Osman, Metastrategy simulated annealing and tabu search algorithm for the vehicle routing problem, Annals of Operations Research 41 (1993) 421–451.

[82] D. Patier-Marque, L'intégration des marchandises dans le systme des dplacements urbains–Actes des Treizimes Entretiens Jacques Cartier, Montral, octobre 2000, Laboratoire d'conomie des transports, Lyon, France, 2001.

[83] G. Perboli, F. Pezzella, R. Tadei, Eve-opt: an hybrid algorithm for the capability vehicle routing problem, DAUIN - Politecnico di Torino - OR/05/06 (2006).

[84] G. Pinna, S. Beretta, Italiamondo - logistica e Intermodalit: Speciale City Logistics, Italmondo Servizi ed Editoria, Milan, Italy, 2005.

[85] H. Pirkul, V. Jayaraman, A multi-commodity, multi-plant, capacitated facility location problem:next term formulation and efficient heuristic solution, Computers and Operations Research 25 (1998) 869–878.

[86] C. Prins, A simple and effective evolutionary algorithm for the vehicle routing problem, Computers and Operations Research 31 (12) (2004) 1985–2002.

[87] P.Toth, D.Vigo, The vehicle routing problem, SIAM Society for Industrial and Applied Mathematics, Philadelphia, USA, 2002.

[88] Y. Rapp, Planning of exchange locations and boundaries, Ericsson Technics 2 (1962) 1–22.

[89] F. Reeves, Genetic algorithms, in: F. Glover, G. Kochenberger (Eds.), Handbook of Metaheuristics, Kluwer, Boston, 2003 55–82.

[90] N. Ricciardi, R. Tadei, A. Grosso, Optimal facility location with random throughput costs, Computers and Operations Research 29 (6) (2002) 593–607.

[91] F. Robusté, C. Daganzo, R. Souleyrette, Implementing vehicle routing problems, Transportation Research part B 24 (1990) 263–286.

[92] W. Ruske, City logistics-solutions for urban commercial transport by cooperative operations management, In: OECD Seminar on Advanced Road Transport Technologies, Omiya, Japan.

[93] M. Savelsberg, M. Sol, The general pickup and delivery problem, Transportation Science 29 (1) (1995) 17–29.

[94] S. Scheuerer, Neue tabusuche-heuristiken fur die logistische tourenplanung bei restringierendem anhangereinsatz, mehreren depots und planungsperioden, Ph.D. thesis, Wirtschaftswissenschaftliche Fakultat, Universitat Regensburg (2004).

[95] L. Schmitt, An empirical computational study of a genetic algorithms to solve order based problems: an emphasis on tsp and vrptc, Ph.D. thesis, Fogelman College of Business and Economics, University of Memphis, TN (1994).

[96] SEARS website http://www.sears.com.

[97] F. Semet, A two-phase algorithm for partial accessibility constrained vehicle routing problem, Annals of Operations Research 61 (1995) 45–65.

[98] F. Semet, E. Taillard, Solving real-life vehicle routing problems efficiently using tabu search, Annals of Operations Research 41 (1993) 469–488.

[99] Z.-J. M. Shen, C. Coullard, M. S. Daskin, A joint location-inventory model, Transportation Science 37 (2003) 40–55.

[100] A. Svoronos, P. Zipkin, Evaluation of one-for-one replenishment policies for multiechelon inventory systems, Management Science 37 (1) (1991) 68–83.

[101] E. Taniguchi, R. Thompson, T. Yamada, J. van Duin, City Logistics: Network Modeling and Intelligent Transport Systems, Pergamon, Amsterdam, 2001.

[102] TNT Logistics website http://www.tntlogistics.com.

[103] P. Toth, D. Vigo, An overview of vehicle routing problem, in: P. Toth, D. Vigo (Eds.), The Vehicle Routing Problem, Vol. 9 of SIAM Monographs on Discrete Mathematics and Applications, 2002 1–51.

[104] P. Toth, D. Vigo, The granular tabu search an its applications to the vehicle routing problem, INFORMS Journal of Computing 15 (2003) 333–346.

[105] United Nations - Economic commission for Europe, European Conference of Ministers of Transport, EUROSTAT, Intermodal transport, in: U. S. d. EUROSTAT, ECMT (Ed.), Glossary of transport statistics, Transport and communication, vol. 2 N 1, 1997 35–39.

[106] R. Vahrenkamp, Transportation logistic in rural settingthe case of milk collection, Tech. Rep. Technical Report 5/1989, Fachbereich Wirtschaftswissenschaften, Gesamthochschule Kassel (1989).

[107] J. Verrijdt, A. de Kok, Distribution planning for a divergent n-echelon network without intermediate stocks under service restrictions, International Journal of Production Economics 38 (1995) 225–243.

[108] P. Wark, J. Holt, A repeated matching heuristic for the vehicle routing problem, Journal of Operational Research Society 45 (1994) 1156–1167.

[109] A. Wren, Computers in Transport Planning and Operations, Ian Allan, London, 1971.

[110] A. Wren, A. Holliday, Computer scheduling of vehicles from one or more depots to a number of delivery points, Operations Research 23 (1972) 333–344.

[111] T. Wu, C. Low, J. Bai, Heuristic solutions to multi-depot location-routing problems, Computers and Operations Research 29 (2002) 13931415.

[112] T. Zunder, J. Ibez, Urban freight logistics in the european union, European Transport 28 (2005) 77–84.