Vehicle Routing in Multi-Echelon Distribution Systems with Cross-Docking: A Systematic Lexical-Metanarrative Analysis
Jesus Gonzalez-Feliu

To cite this version:

HAL Id: halshs-00834573
https://halshs.archives-ouvertes.fr/halshs-00834573
Submitted on 16 Jun 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Vehicle Routing in Multi-Echelon Distribution Systems with Cross-Docking: A Systematic Lexical-Metanarrative Analysis

Jesus Gonzalez-Feliu

1 Laboratoire d’Economie des Transports, Institut des Sciences de l’Homme, Lyon, France

Correspondence: Jesus Gonzalez-Feliu, Laboratoire d’Economie des Transports, Institut des Sciences de l’Homme, 14 Avenue Berthelot, 69363 Lyon cedex 07, France. E-mail: jesus.gonzales-feliu@let.ish-lyon.cnrs.fr

Received: January 11, 2013    Accepted: April 26, 2013    Online Published: May 14, 2013
doi:10.5539/cis.v6n3p28    URL: http://dx.doi.org/10.5539/cis.v6n3p28

Abstract

Multi-echelon distribution systems and more precisely, optimization of LTL routes related to them is one of the most popular subjects in the last 5 years of vehicle routing research. Although a plethora of models, methods and visions is found, it is still difficult to compare them because they use different terminologies and some authors insist on the fact there are a multitude of close but different problems. This paper presents the main concepts of multi-echelon distribution with cross-docks and the notation standards for cost optimization in this field on an attempt of unification, in order to provide a guide to researchers and practitioners. A literature review is first presented, in order to list the main problems and methods that are found in the literature. Then, by a hybrid systematic analysis method combining a lexical and a meta-narrative analysis, the main concepts and standards of multi-echelon based vehicle routing optimization problems are presented. A theoretical model as well as a classification of solving methods, both exact and heuristic, is presented. Finally, research paths are proposed to support both scientists and outbound logistics practitioners.

Keywords: location-routing problems, multi-echelon distribution, cross-docking, combinatorial optimization, survey

1. Introduction

Vehicle routing optimization is one of the most popular subjects in operations research and transport sciences that have more than 50 years of existence (Laporte, 2009). From the first, basic version of Dantzig and Ramser (1959), a wide variety of sub-problem families and variants have been defined (Toth & Vigo, 2002; Golden et al., 2008). Due to logistics constraints and costs (Note 1), multi-echelon distribution systems have started to become popular (Gonzalez-Feliu, 2012a). A wide variety of fields have developed multi-echelon transportation approaches. The most representative examples are seen in press distribution (Gonzalez-Feliu & Morana, 2011), parcel and postal distribution (Gracin & Stipetíc, 2009) or intermodal transport (Lowe, 2005). Such systems are also found in the automotive industry, more precisely in spare parts supply distribution (Kargari et al., 2012) and in urban freight distribution (Crainic, 2008). Moreover, Grocery distribution presents also heterogeneous group chains which are based on distribution systems presenting several echelons, mainly after the development of new services for the end-consumer supply, like e-commerce and proximity deliveries (Gonzalez-Feliu et al., 2012).

Although multi-echelon transportation systems are very common in real cases, they are usually decomposed into an addition of single-echelon distribution cases. We find however several works in research dealing to system cost optimization in N-echelon systems, and although the problem was little studied in the decade 1980-1990, in the last five years we observe a renewal of this field, with tens of papers published since 2010 (Gonzalez-Feliu, 2011) and the growing trend has increased in the last two years.

The aim of this paper is to present a guide to scientists and practitioners about cost optimization for multi-echelon distribution systems with cross-docking facilities and a solid step onto unification. This work is the continuation of two precedent exploratory analyses, one more descriptive (Gonzalez-Feliu, 2012b) and the other representing a first systematic analysis approach (Gonzalez-Feliu, 2011). The main differences between this paper and previous work are the following. First is that previous works were exploratory and to this fact, and due to a lack of unification, several papers that set milestones (in terms of dates where different findings where announced) were not identified. To deal with this difficulty, a combination of lexical and meta-narrative approaches led to the identification of seldom cited papers which are prior to most known works. Second is that,
differently to classical systematic reviews, grey literature was included in the analysis. Third is that previous works focus mainly on heuristics, and exact methods and modeling frameworks, more popular in the last two years, were not published when the work was made. Moreover, tens of new works have emerged between the acceptance of previous work and end 2012. For those reasons, it was important to achieve the exploratory analysis on a mature systematic literature review. The paper is organized as follows. In section 2 we will make an overview on vehicle routing, focusing on approaches that are directly related to multi-echelon distribution, following a descriptive approach. Section 3 presents the systematic analysis method, which is an innovative combination of lexical and metanarrative approaches. Since those terms are not always familiar to computer science researchers, the main notions of lexical and metanarrative analyses are introduced before describing the method. Section 4 presents the results of the analysis, which are three. The first is the justification of a unification need, by proving that the different problems are subvariants of the same family and they are so closed that a unified modeling framework can be proposed. The second is the adaptation of a previous model to represent a wide variety of problems, all related to multi-echelon route optimization. The third is a classification of methods, which focuses on vehicle routing and the representation granularity of the observed reality. Finally, as a conclusion, research paths for further developments in the field of multi-echelon route optimization are addressed.

2. Literature Review

The Vehicle Routing Problem (VRP) is a popular subject of research in combinatorial optimization. The basic version of the problem is due to Dantzig and Ramser (1959), and since that, a plethora of problem variants, modeling approaches, exact and heuristic solving methods and applicative frameworks are found in the literature. The basic version consists in defining a set of routes for a fleet of vehicles based at one depot for a number of geographically dispersed points, called customers. 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.

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 small number of cases, sometimes in a simplified way but can be a suitable representation for strategic and tactical planning purposes (Crainic & Laporte, 1997). Depending on the situation and the important parameters that define the envisaged freight transport system, one or more groups or constraints can be added or modified to represent the different. In other cases, the objective function (mainly cost reduction) will be also modified (reduction of the number of used vehicles, crew management optimization, profit increasing, quality satisfaction, etc.) even proposing multi-objective approaches. Other variant arise when heterogeneity is considered. The most common problem with heterogeneity is obtained when the vehicles have different capacities. 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 (Gonzalez-Feliu, 2008).

Time constraints are the most used in literature (Braysy & Gendreau, 2005a, b). Such problems are known as VRP with Time Windows (VRP-TW). A Time Window (TW) is defined as the interval of time within which a vehicle must arrive to a node, and it is usually characterized by an early arrival time (EAT) and a late arrival time (LAT). Note than VRP-TW presents a wide variety of sub-variants, such as VRP with soft or semi-soft time windows, where a penalty on transportation cost is applied if the TW is not respected (Qureshi et al., 2010), dynamic VRP-TW (Taniguchi et al., 2010), where travel times are evolving with times or VRP with access time windows, where the time constraints are associated to a zone to access and not to a customer to serve (Quak & De Koster, 2009; Muñuzuri et al., 2012). Other variants are Multi-depot VRP (MD-VRP), where more than one starting depot is defined, Windy VRP, also known as Asymmetric VRP (A-VRP), where the road network presents asymmetric costs, split deliveries VRP (SD-VRP), where deliveries can be split in order to put each customer deliveries in more than one vehicle if necessary, or stochastic VRP (S-VRP), where customer demands are not known in advance and are represented using stochastic processes. Many works and surveys related to VRP, as well as more variants, can be found in literature (for more details about this problem, see Golden & Assad, 1988; Laporte, 1992; Toth & Vigo, 2002; Cordeau et al., 2007; Golden et al., 2008).

Location Routing Problems (LRP) differ from classical VRP approaches in the fat they seek 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 (Min et al., 1998). The basic ones, analogously to basic VRP, are: (1) customer demands are satisfied without exceeding vehicle capacities, (2) facility capacities are not exceeded, (3) each route begins and ends at the same facility, and (4) the number of active vehicles does 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 (Madsen, 1983; Laporte, 1988; Min et al., 1998; Nagy & Sahli, 2007). Analogously to VRP, many variants are defined for these problems. Literature reviews on this subject are quite recent (Nagy & Sahli, 2007; Drexl, 2013).

Concerning multi-echelon transport cost optimization, we can several sets of problems, derived from those two families (VRP and LRP). The first set of problem is the N-echelon Location Routing Problems (NE-LRP), defined as the generalization to N-echelon distribution systems of classical LRPs (Min et al., 1997). They were firstly introduced by Jacobsen and Madsen (1980) in its two-echelon version and a multi-echelon variant therorised by Laporte (1988). Two-echelon models can be found in Zegordi and Nikbaksh (2009) and Nguyen et al. (2011, 2012) and approximated three and four echelon versions are respectively proposed by Hamidi et al. (2005) and Ambrosino and Scutellà (2005). The N-echelon problem has been formally modelled recently, with a set partitioning formulation proposed by Gonzalez-Feliu (2012b) giving some advices on how the problem can be solved.

Most works in literature concern heuristics for the two-echelon version of the problem (the 2E-LRP). Jacobsen and Madsen (1980) first introduce the problem and present three fast heuristics for solving a real case application from the press distribution field. The first method is the Tree Tour Heuristic (TTH), which is based on the Steiner Tree Problem (Francis & White, 1974), solved using 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-SA heuristic is a three stage procedure composed from the Alternate Location Allocation (ALA) presented in Rapp (1962) and the Savings algorithm (SAV) introduced by Clarke and Wright (1964). The third heuristic (SAV-DROP) is also a three stage procedure composed from the Clarke and Wright Savings algorithm and the DROP method (Feldman et al., 1966).

Lin and Lei (2009) formulate a two-echelon LRP with routing considerations at both echelons. In this problem, it is possible to deliver some customers directly from the depot, or make it passing through distribution centers (DCs). Three simultaneous decision problems are considered: 1) number and locations of DCs, 2) customer assignment to DCs, 3) first-echelon routing (from depot to DCs and customers served in direct shipping), and 4) second-echelon routing (between DCs and customers not included in first-echelon routing). The authors propose a hybrid method that consists on a genetic algorithm embedded with a local search procedure. Zegordi and Nikbaksh (2009) study the same problem, adding distance constraints, and proposing a derived model and a simulated annealing algorithm. In this case, new instances are also proposed and solved.

Boccia et al. (2010) define the structure of a 2E-LRP for a real application. The problem is then solved by a meta-heuristic combining a constructive procedure improved by Tabu Search post-optimisation that connects the two echelons by proposing moves that modify routes of both echelons at the same time. This method can solve real large size instances. Sterle (2011) proposes three models and a Tabu Search heuristic that splits the system into two separate problems. The two groups of methods (models and Tabu Search heuristics) are compared on a set of instances derived from Boccia et al.’s (2010) instances. Lee et al. (2010) formulated 3E-LRP to represent a supply chain including suppliers, manufacturers, distribution centers (DCs) and customers. Although there is a production stage at manufacturers, the problem consider all three echelons of transport and seeks to optimize them simultaneously, also including the location of both manufacturers and DCs, customers assignment to DCs and supplier assignment to manufacturers. The authors propose a mathematical formulation and a heuristic algorithm based on LP-relaxation to solve the problem.

Nguyen et al. (2010, 2011) propose four constructive heuristics and one meta-heuristic to solve the LRP-2E with capacity constraints on vehicles and satellites. The best heuristic builds giant tours over the set of customers, each tour being limited by the capacity of a 1st-echelon vehicle. The giant tours are then partitioned using a splitting procedure that inserts the satellite depots. A TSP is finally solved to visit the satellites selected. The proposed meta-heuristic is a greedy randomized adaptive search procedure (GRASP) reinforced by a learning process. The same authors also present a hybrid metaheuristic combining a GRASP procedure and an evolutionary/iterated local search (ELS/ILS) to solve a 2E-LRP. The GRASP procedures uses three constructive heuristics followed by local search to generate an initial solution, then an intensification strategy is carried out by a dynamic alternation between ELS and ILS, using Tabu Search (Nguyen et al., 2012). Schwenger et al. (2011) propose a Variable Neighborhood Search algorithm and compare the obtained results with the existing methods having been applied to instances used by Nguyen et al. (2010, 2011). Contardo et al. (2012) present a
branch-and-cut algorithm and an adaptive large neighborhood search algorithm (ALNS) to solve Nguyen et al.’s (2010) instances, solving to optimality several instances of 50 customers and some of 100 and 200 customers.

Finally, Gonzalez-Feliu (2012b) and Hamidi et al. (2012) attempt a unification of this problem by modeling respectively an N-echelon and a three-echelon LRP. The first is a theoretical work that proposes to solve the problem via branch-and-prize and obtain lower bounds by column generation without providing computational results. The author also lists some existing heuristics methods for two-echelon variants mainly and envisages the possibilities of solving N-echelon problems with them. The second aims to be applied to supply chain management, and the model is solved via a commercial tool. Only one two-echelon instance is solved to optimality, including 2 depots, 3 DCs and 5 customers.

The other popular family of problems is the N-echelon Capacitated Vehicle Routing Problem (NE-CVRP). In this family of problems only one depot is considered and the intermediary facilities allow a limited number of vehicles for transshipments (Gonzalez-Feliu, 2008). The satellites do not present location costs and constraints but only allocation costs and capacity limitations related to the number of vehicles that can host. Gonzalez-Feliu et al. (2007) present a first MIP formulation for the 2E-CVRP derived from multi-commodity network design to study the limits and the general behavior of the mathematical model. The model is tested on three sets of instances, available at OR Library website (Beasley, 1990). Optimal solutions are found for instances up to 21 customers (76 of 78 instances) and lower bounds are presented for all the instances. The authors also introduce some cuts which make the calculation time decrease. An alternative formulation is proposed by Gonzalez-Feliu (2008). The author compares this model to that of Gonzalez-Feliu et al. (2007) and proposes a method to obtain lower bounds by column generation using a set covering formulation. The method needs extremely high computational times because of the exponential multiplication of equivalent solutions when generating columns, due to the modeling choices that led to the proposed model, and the author concludes that the method needs to be revised or the formulation changed to decrease computational time (Gonzalez-Feliu, 2008). More lower bounds for this problem are proposed by Crainic et al. (2008). First, the authors split the problem into two simpler sub-problems, one for each level; then, in a second phase, a global lower bound is computed by addition of the sub-problems costs without considering the interactions between both levels. Perboli et al. (2010) propose several cuts on the commodity-based model by introducing several inequalities, some of them derived from the TSP-based model and other from commodity formulations. All instances up to 21 customers are solved to optimality via Branch-and-Cut, as well as some of the 32 customers’ subset. Santos et al. (2012) proposes a branch-and-prize algorithm from the two-echelon version from a set covering formulation for N-echelon problems, and estimates lower bounds using column generation. This method, tested on Gonzalez-Feliu et al.’s (2007) instances, obtains lower bounds of similar quality as Perboli et al. (2010). Baldacci et al. (2013) propose a bounding procedure based on dynamic programming and a decomposition algorithm that splits the problem into single and multi-depot CVRPs, obtaining exact optimums in 144 out of 153 instances. Jepsen et al. (2013) propose a three index formulation, then new cuts and a BC algorithm is proposed. With this method, all instances up to 32 customers are solved to optimality, as well as some of those with 50 customers.

Concerning heuristics, Crainic et al. (2010) propose a two-phase heuristic algorithm based on a clustering first routing second procedure plus a classical local search procedure. This heuristic has been used to make a satellite location analysis in order to build instances up to 250 customers, providing also a first sub-optimal solution as a reference for further, methods for the 2E-CVRP. Those results are improved by the same authors using multi-start heuristics (Crainic et al., 2011) combining local search and tabu search algorithms. A hybrid heuristic ant colony algorithm for the same problem is proposed by Wang et al. (2011). The problem is divided into m+1 CVRP by a separation strategy (distance-based cluster). Then better feasible solutions are built by an improve ant colony optimization with multiple neighbourhood descent (IACO_MND). Computational results show the performance of the method, which obtain comparable results to those of Crainic et al. (2011) in shorter times.

Gonzalez-Feliu et al. (2010) propose a clustering-first route-second algorithm for a multi-stakeholder, heterogeneous fleet version of the problem. Several instances derived from Fisher’s (1995) ones are solved. Improvements on this algorithm are shown in Gonzalez-Feliu (2012a) who makes a practical analysis of logistics sharing and crossdocking schemes. Boissier (2011) adapts the Groër et al. (2011) library of VRP heuristics to solve the same problem. In that work, two construction heuristics are compared (both adapted from the ALA-SAV and the TTH of Jacobsen and Madsen, 1980) and obtain comparable results to precedent ones. Gonzalez-Feliu and Salanova (2012) propose a Semi-Greedy algorithm for a real-life application of the same problem, proposing a set of very large instances (up to 5 depots, 14 satellites and more than 1500 customers). Breuing (2009) proposes a Large Neighborhood Search method (LNS) for the same problem, comparing the algorithm’s performance to the current literature. Hemmelmayr et al. (2012) propose an Adaptive Large
Neighbourhood Search procedure and apply it to both LRP and 2E-CVRP instances. In Crainic et al. (2013) propose a Reactive GRASP with Path-Relinking, similar to that of Nguyen et al. (2011, 2012). The authors are able to solve instances up to 5 satellites and 50 customers. Finally, Xu et al. (2013) propose a memetic algorithm, using a genetic algorithm to identify the most suitable solution spaces to investigate then a local search procedure to find most suitable solutions within the solution spaces. As seen before, most works use instances of Gonzalez-Feliu et al. (2007) and Mancini (2011) (Note 2) to compare the performance of algorithms to current literature.

The road-train routing problem, also known as truck-and-trailer routing problem (TTRP), 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 pick up the trailer (Nagy & Salhi, 2007). The difference from the basic 2E-LRP variant is that in this case some customers can be served directly a 1-route. The first studies related to the TTRP were developed in the agricultural field, in order to optimize the milk collection operations, and were presented as applied solving methods for a real situation. The milk is collected at storage tanks on the farmyards 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. Wren (1971), Brunswicker (1986) and Vahrenkamp (1989) propose 3-stage heuristic procedures which work in a similar way. In the first phase (clustering), customers are grouped considering the number of vehicles and their characteristics (mostly the capacity of each lorry-trailer convoy). The second phase (allocation) consists in determining, for each convoy, one transshipment location. The third and last stage is the routing phase.

Semet and Taillard (1993) formulate the TTRP as a Mixed Integer Program (MIP), and propose an algorithm based on tabu search to solve it. An initial solution is obtained by a sequential algorithm and improved by tabu search, when customers are reallocated. This method does not distinguish between locational and routing moves. Semet (1995) proposes a clustering first routing second solution method. First customers are allocated to roots then the resulting routing problems are solved via Lagrangian relaxation. Gerdessen (1996) 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 (2002) 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 (2006) presents two new construction heuristics: a clustering-based sequential insertion procedure and an adaptation of the well-known sweep algorithm by Gillett and Miller (1974) and a tabu search improvement procedure. Moreover, the author adapts these procedures to the multi-depot and the multi-period version of the problem. Hoff and Løkketangen (2007) present 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. Lin et al. (2009) propose a simulated annealing (SA), then computational tests on the instances proposed by Chao (2002) are presented to compare the proposed SA algorithm to Chao’s and Scheuerer’s procedures.

Drexel (2007) defines the VRP with Trailers and Transshipments (VRPTT) as the generalization of the TTRP when trailers can be pulled by different lorries on its itinerary and several transshipments are allowed. The author introduces also the notion of support vehicles, which are connecting the depots with the transshipment points without the possibility to visit any customers (1st-echelon vehicles). The problem is solved using branch-and-cut and branch-and-prize, which are the first exact methods developed for a variant of 2E-LRP. The author extends the algorithms (Drexel, 2011) to provide heuristics solutions for instances up to 25 truck customers and 25 trailer customers. Tan et al. (2006) applied a hybrid evolutionary algorithm (HMOEA) to the multi-objective variant of the VRPTT. The algorithm uses specialized genetic operators, a variable-length representation and a local search method. The authors compare it with a non-hybrid genetic algorithm (using only the proposed genetic operators) and a hybrid genetic algorithm where standard operators are used instead of those proposed by HMOEA.

Caramia and Guerriero (2009) propose an alternative formulation from the problem, and solve it heuristically by dividing the problem into two sub-problems: a customer-vehicle assignment problem (CAP), and a route definition problem (RDP), similar to a TSP. A heuristic method is then constructed by combining both formulations then the initial solutions obtained in this way are improved using tabu search. The algorithm is tested on Chao’s (2002) instances. The authors adapt the algorithm to a realistic case for milk collection by adding a multiple-restart mechanism to improve the solution (Caramia & Guerriero, 2010). In this case, only
small instances are solved, but they correspond to real vehicle dispatching problems. Villegas et al. (2010) propose two metaheuristics based on greedy randomized adaptive search procedures (GRASP), variable neighborhood descent (VND) and evolutionary local search (ELS) to solve that problem. To evaluate these metaheuristics the authors propose a set of 32 randomly generated problems. The computational experiment shows that a multi-start evolutionary local search obtain better results that GRASP/VND. A synthesis on such problems has recently been made by Drexl (2013).

It is also possible to define such systems by approximating one part of the transport costs and considering the routing of the others. Gendron and Semet (2009) formulated the N-echelon location-distribution problem (NE-LDP), which differs from LRP in the cost estimation. In fact, no routing construction is made in the NE-LDP, but an estimated cost is defined to deliver a customer from a depot through a satellite. The authors propose a MIP formulation for the two-echelon variant of the family, and solve it using CPLEX. They also present a heuristic method based on variable neighborhood search. Following an initial greedy construction procedure, this method iterates over three phases. In the first phase (node+arc-based neighborhood descent) the authors propose four types of neighborhoods: (1) close one depot; (2) close one satellite; (3) serve one customer through a different satellite; (4) serve one satellite through a different depot. In the second phase (open depot+satellite neighborhood descent), two types of neighborhoods are used: (1) open one depot and (2) open one satellite. To evaluate the impact of each move of type (1), the authors first serve through the open depot every satellite connected to it, and then we perform phase 1, the node+arc-based neighborhood descent, on the resulting network. The evaluation of each move of type (2) is performed similarly. Then, a diversification phase is made using an adaptive memory that stores the best solutions found so far, selects some of these solutions and perturbs them by performing random moves in a large-scale neighborhood, defined by all moves that consist in closing k paths, replacing them by k alternative paths. Crainic (2008) developed a theoretical route optimization methodology for the two-echelon freight distribution system in congested urban areas defined in Crainic et al. (2004). The problem is similar to the 2E-LDP and can be considered an extension of Gendron and Semet (2009) model to city logistics. The methodology is a support for day-before planning operations in a system where satellites are used as intermediate transshipment points for the freight distribution, and the synchronization between the vehicles of both echelons is one of the most constraining aspects of this problem. Those works are expanded and validated in Crainic et al. (2009). In the proposed method, the system is split into two connected sub-problems that are solved separately. However, when solving each sub-problem, constraints that take into account the complementary echelon are defined to keep a systemic approach. Two classes of models are proposed: a service network design model for the 1st-echelon vehicles, which approximates the second echelon routing costs, gives a first estimation of the overall costs; then a second model optimizes the 2nd-echelon trip costs considering the first-echelon vehicle movements.

Ambrosino and Scutellà (2005) propose a mathematical programming formulation for a number of static and dynamic scenarios based on the general multi-echelon LRP. 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.

Another cost approximation approach is the vehicle routing problem with satellite facilities (VRP-SF). In this variant, the network includes facilities that are used to replenish vehicles during a route (Bard et al., 1998a, 1998b; Agnelelli & Speranza, 2002). When possible, satellite replenishment allows the drivers to continue the deliveries without necessarily returning to the central depot. However, the transportation cost of 1st-echelon vehicles is not considered in the system, i.e. only the second-echelon cost is taken into account in the objective function. Crevier et al. (2007) studied a variant of this problem where inter-satellite trips are allowed for satellite replenishment, presenting a model which considers the inter-satellite transportation costs and a heuristic method to solve the problem. This method combines an adaptive memory principle and a tabu search procedure.

Another approach is that of multi-echelon facility location problems (NE-FLP). They deal with the location of depots and intermediary facilities. In such problems, costs are generally approximated, mainly by giving a fixed
cost to assign each customer to each intermediary facility and in similar way to relate intermediary facilities among them or to a depot (Aikens, 1985). This problem does not contain VRP-related approaches and the main goals are related to location and allocation, but not to tactical actions as vehicle routing. For this reason, we will not develop them here. However, detailed surveys including mathematical formulations and solving methods can be found in Aikens (1985), Klose and Drexl (2004), Dullaert et al. (2006) and Melo et al. (2009).

Another approach when defining cross-docking, transshipment and echelon exchanges is to define those exchange points as special customers that first receive freight, then it is collected by different vehicles to deliver other facilities or customers. This approach consists then to define the problem as a pickup and delivery problem (PDP) at some nodes, and a classical VRP at other nodes. PDP are more problems but present several variants because of the different possibility to consider the pickups and the deliveries in route organization. Savelsberg and Sol (1985) formulated the general problem for transportation services with pickups and deliveries. Moreover, two other recent works make different classifications showing the difficulty to follow a unified lecture grid (Berbeglia et al., 2007; Parragh et al., 2008). An explicit multi-echelon variant of the PDP is the VRP with cross-docking operations, described in Wen et al. (2009). Two types of vehicles are defined: one type make the delivery operations at a cross-docking platform and the other type is charged of picking up the freight and shipping it to customers. At each cross-docking platforms, delivery vehicles must arrive before pickup vehicles. The authors propose a mathematical formulation for instances with only one cross-docking platform and a Tabu Search heuristic procedure which uses two neighborhoods and is finally embedded within an Adaptative Memory Procedure (AMP), in order to reach better and most robust solutions. This problem is interesting because decisions for cross-docking operations (in terms of loading, unloading and consolidation) are taken into account. However, no generalization to a multi-platform problem has been formulated. Although young, this problem presents some interesting works, that are not detailed here (mainly Wen et al., 2009; Liao et al., 2010; Tarantilis, 2012; Dondo & Cerdà, 2013, but a similar problem is proposed in Yang & Xiao, 2008). Another variant consists on allow pickup and deliveries on all nodes of the graph, but consider transshipment points as special nodes where the pickup and the delivery commodities are equal for each point (in nature and quantity) and freight needs to be first delivered to be then picked up. Some examples, not detailed here, are described in Drexl (2013).

3. Methodology

From the literature review, we observe heterogeneous groups of works and approaches, with different notations and methodologies. In order to unify them into a general framework to guide vehicle routing for multi-echelon (or multi-stage) transport systems, it is important to examine them in-depth. To do this, we observe two main categories of approaches: text mining and lexical analyses and meta-narrative analyses. The first category (Feldman & Sanger, 2006) is very common in transport and logistics research, whereas the second is almost exclusive of medical and biological science (Greenhalgh et al., 2005), with only few applications on other fields.

In a precedent work (Gonzalez-Feliu, 2011), a meta-narrative analysis was pursuit to define the first unification criteria, although the work was exploratory and did to lead to a general unified framework.

Here, we aim to combine both to find a unified framework. To do this, it is important to first identify the similarities and differences between both approaches. Lexical analyses are text-mining techniques that aim to identify common words and expression in order to define classifications and clusters using techniques similar to those of data mining but applied to textual data (Feldman & Sanger, 2006). Metanarrative analyses are related not to words but to concepts. According to Greenhalgh et al. (2005), “a meta-narrative embraces a shared set of concepts, theories and preferred methods” and “is sited within a particular scientific discipline and should to be regarded not as the unified voice of a community of scholars but as the unfolding of what they are currently discussing about”. Indeed, researchers and scholars of similar communities tend to cite one another’s work, not only to agree but also to contest it, they attend the same scientific events, publish in the same journals, and use broadly similar criteria for judging validity and rigor.

The basic methods are similar and based on the following framework, which is an iterative procedure working as follows:

1) **Informal search phase.** With a view to identify a first set of meta-narratives, the main actions were exploratory (mainly browsing and asking colleagues), in order to find a first set of keywords. Then, using snowballing techniques (i.e. searching references of references and using citation-tracking databases), we were able to identify key sources. After that, seminal sources were identified by asking what were cited as key original and scholarly contributions by other researchers in the same field. The main differences is that lexical methods follow a database strict search (mainly form Web of Science and Scopus), completing them
with complementary sources, and meta-narrative methods combine database search with citation of citation tracking and interview-based methods (mainly during conferences, seminars and other scientific events) to track a large set of works.

2) **Mapping phase.** From these sources, the main concepts are extracted. Lexical methods use data mining (mainly clustering-based) techniques to identify the most common words or short expressions in order to define distances, proximities and similarities between works, although metanarrative methods search for methodological schemes that formed meta-narratives in a rigorous way.

3) **Formal search, clustering and appraisal iterative phases.** This is an iterative phase that aims to refine search from the first identified expressions (in lexical works) or metanarratives, in order to complete the systematic literature search. In this way, once the basic expressions or metanarratives are defined, new searches can be carried out more in-depth.

4) **Synthesis phase.** After several iterations (when no more works related to the chosen field, i.e., multi-echelon freight transport cost optimisation, are found), a synthesis of the chosen works is made. The most used keywords are identified and a literature refinement leads to the definition of the final set of expressions or meta-narratives, from which an in-depth quantitative and qualitative analysis is carried out.

The basic idea of this methodology is to follow the common structure, by combining both approaches. Although systematic reviews and other text and meta-narrative mining techniques are usually applied to empirical works, it has been recently shown that they can also be related to theoretical, modeling and computational studies when the aims of the systematic analysis are the same, i.e. to identify and unify concepts on a field were only incomplete synthesis have been made. To straighten the method, we gave great weight to studies that had been flagged as “high quality” by other scholars in combinatorial optimization. As stated by Greenhalgh et al. (2005) and Gonzalez-Feliu (2011), this type of methods are iterative and can sometimes lead to several false steps in the classification scheme, as well as underline uncertainties concerning the quality and relevance of papers in traditions when some of them are unfamiliar to us. However, the number of total references in multi-stage transport management is smaller than in other fields of logistics management and operations research such as inventory management, demand forecasting or time-dependent direct shipping optimization.

### 4. A Unified Vehicle Routing Framework for Multi-Echelon Distribution

In supply chain management, one of the first aspects that takes place when relating freight transport to other operations is the definition of one or more shipping strategies. The various strategies in practice can be related to vehicle usage and to hierarchical configuration of the system (Gonzalez-Feliu, 2011). Concerning vehicle usage, and focusing on road transport, two main approaches are seen, related to the way trucks are used. When vehicles do direct routes with only one destination stop, independently of the loading factor, we speak about Truck-Load (TL) or Full Truck Load (FTL), referring to the fact that full truck load is delivered to only one destination. When a vehicle delivers two or more destinations, i.e. makes routes, we speak about Less-than-Truck Load (LTL). In this paper, we focus on LTL transport.

Focusing on the hierarchical configuration of the transport system, three predominant shipping strategies can be found in outbound logistics (Gonzalez-Feliu, 2012a):

- **Direct shipping** consists in delivering freight directly from the origin to the destination. It is important to distinguish vehicle usage and hierarchical configuration. Direct shipping means that freight is not changing vehicle from origin to destination, but not necessarily that it travels following an FTL scheme (in real applications, we observe many direct shipping routes, so LTL transport without cross-docking or warehousing).

- **Multi-echelon distribution with warehousing** refers to systems in supply chains containing one or more manufacturers, a set of warehouses, and the final destination of freight. Freight requests are made to warehouses, which have a stock of freight. These warehouses order freight in big quantities from factories.

- **Multi-echelon transport with cross-docking** differs from the warehousing strategy in the fact that cross-docking platforms do not have the possibility to stock, but allow the consolidation and transshipment operations, and the orders are made directly to the origin of the freight, which is in general a factory or a warehouse.

- **Mixed schemes** can also be defined in supply chains mixing warehousing and cross-docking, and also in cases where the same destination combines multi-echelon supply systems with supply at manufacturers using direct shipping to ship their commodities.
All those concepts seem well established but we observe several divergences in scientific literature. More precisely, the concepts related to multi-echelon logistics are well defined but we observe a lack concerning multi-echelon transport and cost optimisation issues that take into account the whole transport system and not only a part of it. The main difficulty that researchers have when making literature reviews on this field is to identify the suitable works because such systems are not defined using a unified terminology (Gonzalez-Feliu, 2012). Moreover, the confusion increases when dealing with supply chain systems: for example, Ambrosino and Scutellà (2005) define a 4-echelon supply chain which is connected by a 3-echelon transport system. For this reason, it is important here to precise what a multi-echelon transport system is, taking into account the context of our study. In a supply chain, an echelon can be defined as the elementary organisation unit of this chain. This means that a supply chain, which results from a complex aggregation of operations, can be divided into various sets of homogeneous operations, mainly related to raw material collection, production, assembling, warehousing and retailing. Transport represents the links between those elementary organisations, and is often externalised in supply chain management. The proposed definition derives then from transport planning and not from supply chain management. So, a multi-echelon transport system is defined as a transport system composed of at least two distinct transport stages, with a connection between them. In that connection, transhipment is made (Drexl, 2007). Such systems can be applied to both personal and goods transport (Ortuzar & Willumsen, 2001). In other words, multi-echelon freight transport systems are multi-echelon logistics systems involving at least two transport schemes and cross-docking operations (Gonzalez-Feliu, 2011). In several works, the word echelon is privileged; other common terms are echelon, level or tier (Gonzalez-Feliu, 2011). Regarding tactical planning and optimization, we observe that VRP and LRP are the base of vehicle routing methods for n-echelon transport systems. Several authors insist in the fact they are different problems (Perboli et al., 2010; Baldacci et al., 2013; Jepsen et al., 2013). Their main arguments are three. We aim to present them and present an argumentation to each statement.

1) The NE-LRP derives from the LRP and the NE-VRP from the VRP (Mancini, 2013).

This is only partially verified. Since NE-LRP derives directly from LRP (if we take a 2E-LRP with only 1 depot and a null distance between the depot and the satellites, we obtain an LRP), we can also obtain a VRP by simplification: the NE-LRP is easily seen to be NP-Hard via a reduction to VRP, which is a special case of NE-LRP arising when just one echelon and one satellite (with no travel cost from the depot to the satellite) is considered. Moreover, LRP derives also from VRP, so the statement is not exact. Indeed, both NE-LRP and NE-VRP derive from LRP then from VRP. To demonstrate it, take the basic 2E-VRP as defined in Gonzalez-Feliu et al. (2007), also taken as reference in Perboli et al. (2010): one depot, \( n_s \) satellites and \( n_c \) customers. The aim of this problem is to both allocate freight to satellites and define both echelon routes to deliver the customer by making cross-docking operations at satellites. If we consider a cost of 0 (null cost) between the depot and the satellites and only one first echelon vehicle of unlimited capacity , we obtain a problem where the aim is to allocate freight to the satellites (and then choose which of them will be activated) and deliver the customers from the satellites. This problem is not a multi-depot VRP (in such problems, the allocation is not a choice and unlimited availability of goods is supposed at each depot) but a classical LRP.

2) In the NE-VRP, the number of satellites is smaller and limited (Perboli et al., 2010).

This is true, but is not an enough condition to affirm that NE-LRP and NE-VRP are different problems. In this case, and only if the difference arises on the number of satellites and nothing more, the NE-VRP and the NE-LRP do not have any difference, since the mathematical formulations are the same and the only difference is a limit on the number of nodes of a subset of the entire graph. Moreover, the computational effort is similar, because it arises only of a number of nodes. In this case the argument is not valid to affirm that both problems are different (no differences on mathematical formulations can be stated if we take the number of satellites as the only differentiation condition).

3) The NE-LRP includes the activation cost of satellites in its objective function and NE-VRP not (Nguyen et al., 2011).

In this case, we observe a difference between both models, and it appears only on the objective function. Since the same problem can present different objective functions (Perboli et al., 2012) and the constraints remain essentially the same (the way they are modeled can change, as shown in Drexl, 2007; Gonzalez-Feliu, 2008; Nguyen et al., 2012; Jepsen et al., 2013), the complexity of operating the different models will depend on the way the constraints are defined, but since they are equivalent in their way to represent the observed reality (Zbilut & Giuliani, 2007), the complexity can be seen as similar. The formal representations are the same, but they represent the same “observed reality” and then the same organizational phenomena. It is then difficult to
affirm that both problems are different only because one adds only a new cost with respect to the other. It is more appropriate, according to Nagy and Salhi definition, to affirm that NE-VRP and NE-LRP are two different names for the same problem. Moreover, this distinction is not always clear to all authors, as shown in Drexel (2013). Indeed, 2E-VRP derives from the works of Jacobsen and Madsen (as said by Gonzalez-Feliu, 2008; Perboli et al., 2012; Jepsen et al., 2013), which is defined as a 2E-LRP but is the basis of several 2E-VRP approaches; also TTRP is defined as a VRP variant (Semet & Taillard, 1993) and as an LRP variant (Nagy & Salhi, 2007). The hypotheses of being separate and different problems are not justified, since they both refer to the same work as the first problem of this type, and confirm Ackoff’s (1977) theories on the scientific rigor and opt-out nature of operations research modern visions. We will then use the 2E-LRP notation since it is the first adopted in literature and consensuated by a larger scientific community (Nagy & Salhi, 2007).

We can then define the N-echelon problem analogously to LRP and VRP literature as follows. Consider an N-echelon distribution system composed by N stages. The depots are the starting points of the distribution chain. We can then define the N-echelon problem analogously to LRP and VRP literature as follows. Consider an N-echelon distribution system composed by N stages. The depots are the starting points of the distribution chain. Customers are defined as the final destinations of the freight (in many real applications they are the stores or retailers, but also households in some home-delivery services). Using this definition, depots constitute the set of 0-satellites and customers the set of N-nodes on the graph. The overall transportation network can then be decomposed into N echelons. The 1st echelon connects the depots to the 1st echelon intermediary facilities. Then, N-2 intermediate echelons are defined to link the different intermediary facilities, or satellites. Finally, the Nth echelon, where the freight is delivered from the (N-1)th echelon intermediary facilities to the final destinations.

To deliver the freight, a number of vehicle fleets are defined. Each echelon e usually has its own fleet of vehicles, defined by different characteristics (capacity, dimensions, speed), and can be heterogeneous or homogeneous. An e-echelon vehicle is a vehicle that travels from an e-1-satellite to an e-satellite.

Let us denote the system as a graph G with three types of nodes. The depots will be represented by the set \( V_0 = \{ s^0_1, ..., s^0_n \} \), the set of e-satellites will be denoted by \( V^e = \{ s^e_1, ..., s^e_n \} \) (for each echelon e) and the customers by the set \( V = \{ c_1, ..., c_n \} \). The e-satellites can be capacitated and this capacity can be noted as the maximum number \( ms^e_k \) of e+1-th-echelon vehicles allowed to receive freight at the e-satellite \( s^e_k \). Each depot is also capacitated, defined in terms of maximum freight quantity stored at depot \( k \), \( K^0_k \). Moreover, each customer \( i \) has associated a demand \( d_i \) to be delivered. The demand of each customer cannot be split among different vehicles only for the Nth echelon, but split deliveries are allowed in the other echelons (Gonzalez-Feliu, 2012b). In other echelons, we assume that each e-echelon can be served by more than one e-1-echelon vehicle, that each e-1-echelon vehicle can serve more than one e-1-satellite in the same route, and that each e-satellite can be served by any number of e-th-echelon vehicles, or by none. Moreover, each vehicle has a specific capacity, i.e. we define a heterogeneous fleet of vehicles. We define as e-th-echelon route \( y^e_k \) a Hamiltonian circuit made by a e-th-echelon vehicle which starts from e-1-th-echelon satellite \( k \), serves one or more e-th-echelon satellites and ends at the departure node.

We can then define a Mixed Integer Programming Model derived from set-partitioning. We define two types of variables. First are the route variables \( y^e_k \), a \( \{0, 1\} \) variable that shows if the route \( y^e_k \) starting at e-1-satellite \( i \) is used or not. To each route \( y^e_k \) are associated its cost, the nodes that serve, respecting capacity and length constraints and the order they are visited. The attribute \( \delta^e_{ik} \) indicates if e-node \( h \) is served by e-route \( i \) starting at (e-1)-node k. Second, we define the facility assignment variables \( I^e_k \), defined as a \( \{0, 1\} \) variable that represent the activation (\( I^e_k = 1 \)) or not (\( I^e_k = 0 \)) of e-satellite \( k \).

Three types of costs are considered: the cost of each route \( y^e_k \) is noted \( c^e_k \), and the activation cost of satellite \( S^e_k \) and the overall cost of transshipment and other operations at this satellite are respectively noted \( SL^e_k \) and \( S^e_k \). The model is defined as:

\[
\min \sum_{e=1}^{N} \sum_{k \in R^e} c^e_k y^e_k + \sum_{e=1}^{N} \sum_{k \in R^e} S^e_k + \sum_{e=1}^{N} \sum_{h \in R^e} \sum_{i \in R^e} \delta^e_{ik} y^e_k + \sum_{e=1}^{N} \sum_{k \in R^e} SL^e_k I^e_k
\]

\[
\sum_{k \in R^e} \sum_{i \in R^e} \delta^e_{ik} y^e_k = 1 \quad \forall e \in \{1, ..., N\} \forall i \in R^e
\]

\[
K^e_h \geq \sum_{k \in R^e} \delta^e_{hk} y^e_k \quad \forall h \in V_0
\]
The objective function to minimize (1) is the total cost resulting on the addition of transportation costs and the satellite’s activation to those derived from the different operations at the satellite. Constraints (2) show that each customer is served by only one route. Constraints (3) and (4) show respectively the limitation capacity of each satellite. Constraints (5) and (6) ensure the demand’s conservation. The nature of the decision variables is formulated in (7) and (8).

We can attempt to propose a classification of solving methods for vehicle routing optimization in n-echelon transport schemes. Most of them are related to the two-echelon version of this problem. In the following we propose to synthesize three main approaches that take into account the multi-echelon nature of such systems and the connection among echelons: the first is that of exact methods, which seek an exact optimum estimating routes for all echelons; the second is that of multi-echelon routing heuristics, which do not find an exact optimum but take into account routing at all echelons. The last one is that of approximation approaches, which take into account the multi-echelon nature of systems but approximate them by avoiding route construction in some echelons (in general, only one echelon has its routes optimized, whereas the other ones use pre-constructed routes or approximate route costs in the combinatorial optimization processes. Although we find several literature reviews on the subject, they often focus on heuristics (Min et al., 1998; Nagy & Salhi, 2007; Gonzalez-Feliu, 2011, 2012b; Mancini, 2013). Indeed, exact methods are recent and few comparisons are provided, and approximation approaches are difficult to identify. For that reason, although all three categories are shown, we focus on exact methods and approximation approaches.

### 4.1 Exact Methods

Exact methods aim at solving the problem in an exact way, i.e., finding solutions and improving them until it is not proved that they are optimal. We observe three types of methods: branch-and-bound, branch-and-cut, and branch-and-prize. The interest of exact methods is not seen in real applications but in finding exact solutions to provide a set of instance in order to calibrate and test heuristics methods (this second category being more suitable to solve problems with real characteristics). However we observe that most methods are not able to solve problems with more than 50 customers (only Contardo et al. find some optima for instances with 100 and 200 customers).

#### Table 1. Main exact methods

<table>
<thead>
<tr>
<th>Authors</th>
<th>Type of algorithm</th>
<th>N_D Max</th>
<th>N_S Max</th>
<th>N_C Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drexl (2007)</td>
<td>Branch-and-Cut and Branch-and-Price</td>
<td>1</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>Gonzalez-Feliu et al. (2007)</td>
<td>Mixed integer formulation with commercial tool</td>
<td>1</td>
<td>5</td>
<td>50 (Note 3)</td>
</tr>
<tr>
<td>Crainic et al. (2008)</td>
<td>Branch-and-cut</td>
<td>1</td>
<td>5</td>
<td>50 (Note 4)</td>
</tr>
<tr>
<td>Gonzalez-Feliu (2008)</td>
<td>Branch-and-bound, column generation lower bounds</td>
<td>1</td>
<td>5</td>
<td>50^d</td>
</tr>
<tr>
<td>Perboli et al. (2010)</td>
<td>Branch-and-cut</td>
<td>1</td>
<td>5</td>
<td>50</td>
</tr>
<tr>
<td>Contardo et al. (2012)</td>
<td>Branch-and-cut</td>
<td>1</td>
<td>10</td>
<td>200 (Note 5)</td>
</tr>
<tr>
<td>Santos et al. (2012)</td>
<td>Branch-and-prize</td>
<td>1</td>
<td>5</td>
<td>50^l</td>
</tr>
<tr>
<td>Baldacci et al. (2013)</td>
<td>Branch-and-bound</td>
<td>1</td>
<td>5</td>
<td>50</td>
</tr>
<tr>
<td>Jepsen et al. (2013)</td>
<td>Branch-and-cut</td>
<td>1</td>
<td>5</td>
<td>50</td>
</tr>
</tbody>
</table>
4.2 “Freight Allocation and Vehicle Routing” Approaches

System route optimization proposed to simultaneously optimize all the routes belonging to the various echelons, as well as the demand assignment to each intermediary platform. They are of different nature and complexity, but as shown in Gonzalez-Feliu (2013) we find three main types of heuristics: constructive heuristics, tabu search methods and advanced local search meta-heuristics. Constructive heuristics, sometimes combined with classical local search post-optimization, are fast methods which result less performing than other heuristics in terms of gap to exact optima, but present the advantages to be adaptable to many contexts, able to be used in large and very large instances, and easy to understand and reproduce. Tabu search heuristics are very popular, mainly due to the generalization of the Unified Tabu Search and the Granular Tabu Search algorithms (Toth & Vigo, 2002) in tools like ILOG Dispatcher (IBM, 2009) and the COIN optimization suite (Groer et al., 2011). They are able to find good solutions (near 3% to exact optima) in medium instances (up to 200 customers) but have difficulties with real size networks (more than 1000 customers and 10 satellites, like in Gonzalez-Feliu and Salanova, 2012). Finally, some advanced techniques, mainly derived from the adaptive and Large Neighborhood Search (Breuing, 2009) or path relinking (Nguyen et al., 2012) start to be used, and have become popular. However, in current practices, only constructive heuristics with local search are used, and in few cases, related to two-echelon distribution systems (Madsen, 1983; Gonzalez-Feliu, 2012a).

Table 2. Main approaches and solving methods for category 1

<table>
<thead>
<tr>
<th>Authors</th>
<th>Type of algorithm</th>
<th>ND Max</th>
<th>NS Max</th>
<th>NC Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wren (1971)</td>
<td>Clustering-allocation-routing constructive heuristics</td>
<td>1</td>
<td>Multiple</td>
<td>200</td>
</tr>
<tr>
<td>Jacobsen and Madsen (1980)</td>
<td>Three route construction and satellite allocation procedures</td>
<td>1</td>
<td>3</td>
<td>4510</td>
</tr>
<tr>
<td>Brunswicker (1986)</td>
<td>Clustering-allocation-routing using constructive heuristics and local search post-optimization</td>
<td>1</td>
<td>52</td>
<td>739</td>
</tr>
<tr>
<td>Vahrenkamp (1989)</td>
<td>Savings algorithm, clustering first routing second algorithm for multi-depot problems</td>
<td>Multiple</td>
<td>Multiple</td>
<td>200</td>
</tr>
<tr>
<td>Semet and Taillard (1993)</td>
<td>Two-step “initial solution” procedure improved by tabu search, with customers reallocation</td>
<td>1</td>
<td>9</td>
<td>45</td>
</tr>
<tr>
<td>Semet (1995)</td>
<td>Lagrangian relaxation-based heuristic algorithm</td>
<td>1</td>
<td>50</td>
<td>100</td>
</tr>
<tr>
<td>Gerdessen (1996)</td>
<td>Combination of sequential heuristics, then improved by a selection of local search heuristics</td>
<td>1</td>
<td>200</td>
<td>200</td>
</tr>
<tr>
<td>Chao (2002)</td>
<td>Cluster first route second “initial solution” and tabu search post-optimisation with customer reallocation</td>
<td>1</td>
<td>150</td>
<td>199</td>
</tr>
<tr>
<td>Scheuerer (2006)</td>
<td>Clustering-based sequential insertion procedure with tabu search post-optimisation</td>
<td>1</td>
<td>150</td>
<td>199</td>
</tr>
<tr>
<td>Tan et al. (2006)</td>
<td>Evolutionary procedure combining genetic operators, variable neighbourhood search and local search</td>
<td>1</td>
<td>150</td>
<td>199</td>
</tr>
<tr>
<td>Lin et al. (2009)</td>
<td>Simulated annealing</td>
<td>1</td>
<td>150</td>
<td>199</td>
</tr>
<tr>
<td>Zegordi and Nikbaksh (2009)</td>
<td>Simulated annealing</td>
<td>10</td>
<td>50</td>
<td>100</td>
</tr>
<tr>
<td>Crainic et al. (2010)</td>
<td>Clustering first routing second with local search</td>
<td>1</td>
<td>5</td>
<td>250</td>
</tr>
<tr>
<td>Bocci et al. (2010)</td>
<td>Tabu Search</td>
<td>5</td>
<td>20</td>
<td>200</td>
</tr>
<tr>
<td>Gonzalez-Feliu et al. (2010)</td>
<td>Clustering-first routing-second semi-greedy algorithm</td>
<td>3</td>
<td>7</td>
<td>310</td>
</tr>
<tr>
<td>Nguyen et al. (2010)</td>
<td>Four constructive heuristic algorithms; GRASP algorithm with learning processes</td>
<td>1</td>
<td>10</td>
<td>200</td>
</tr>
<tr>
<td>Perboli et al. (2011)</td>
<td>Math heuristics and Branch-and-cut hybridated with Tabu Search</td>
<td>1</td>
<td>5</td>
<td>50</td>
</tr>
<tr>
<td>Boisier (2011)</td>
<td>Savings algorithm, local search post-optimization</td>
<td>3</td>
<td>7</td>
<td>310</td>
</tr>
<tr>
<td>Crainic et al. (2011)</td>
<td>Multi-start heuristics, local search post-optimization</td>
<td>1</td>
<td>5</td>
<td>250</td>
</tr>
</tbody>
</table>

39
4.3 Approximation Approaches

In these problems, the main goal is not to precisely design each route plan but to give a general detailed definition of the two-echelon transport system. For this reason, costs are approximated, creating groups of customers that are then assigned to routes. We observe from Table 3 that this approximation allows developing exact methods and using linear programs solved by commercial tools. The number of customers remains in many problems small.

Table 3. Main approaches and solving methods for the category 3

<table>
<thead>
<tr>
<th>Authors</th>
<th>Type of algorithm</th>
<th>ND Max</th>
<th>NS Max</th>
<th>NC Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sterle (2011)</td>
<td>Tabu Search</td>
<td>5</td>
<td>20</td>
<td>200</td>
</tr>
<tr>
<td>Schwenger et al. (2011)</td>
<td>Variable Neighborhood Search</td>
<td>1</td>
<td>10</td>
<td>200</td>
</tr>
<tr>
<td>Wang et al. (2011)</td>
<td>Simulated annealing</td>
<td>1</td>
<td>5</td>
<td>50</td>
</tr>
<tr>
<td>Hamidi et al. (2012)</td>
<td>Mixed integer formulation solved with commercial tool</td>
<td>2</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>Lin and Lei (2009)</td>
<td>Hybrid procedure: genetic algorithm and local search</td>
<td>1</td>
<td>5</td>
<td>50</td>
</tr>
<tr>
<td>Caramia and Guerriero (2009)</td>
<td>MIP-based heuristics with local search</td>
<td>1</td>
<td>150</td>
<td>199</td>
</tr>
<tr>
<td>Caramia and Guerriero (2010)</td>
<td>MIP-based heuristics, local search and multiple-restart</td>
<td>1</td>
<td>4</td>
<td>40</td>
</tr>
<tr>
<td>Crainic et al. (2013)</td>
<td>GRASP with Path Relinking</td>
<td>1</td>
<td>5</td>
<td>50</td>
</tr>
<tr>
<td>Hemmelmayr et al. (2012)</td>
<td>Adaptive Large Neighbourhood Search procedure</td>
<td>1</td>
<td>5</td>
<td>50</td>
</tr>
<tr>
<td>Contardo et al. (2012)</td>
<td>Adaptive Large Neighbourhood Search procedure</td>
<td>1</td>
<td>5</td>
<td>200</td>
</tr>
<tr>
<td>Crainic et al. (2013)</td>
<td>GRASP with Path-Relinking</td>
<td>1</td>
<td>5</td>
<td>50</td>
</tr>
<tr>
<td>Crainic et al. (2004)</td>
<td>Mixed integer formulation solved with commercial tool, 1st echelon approximated</td>
<td>1</td>
<td>12</td>
<td>51</td>
</tr>
<tr>
<td>Ambrosino and Scutellà (2005)</td>
<td>Mixed integer formulation solved with commercial tool, some echelons approximated</td>
<td>1</td>
<td>5</td>
<td>25</td>
</tr>
<tr>
<td>Gendron and Semet (2009)</td>
<td>Mixed integer formulation solved with commercial tool, 1st echelon approximated</td>
<td>93</td>
<td>320</td>
<td>722</td>
</tr>
<tr>
<td>Semet et al. (2012)</td>
<td>Sequential procedure (constructive “initial solution” heuristic and tabu search post-optimisation )</td>
<td>1</td>
<td>5</td>
<td>50</td>
</tr>
<tr>
<td>Dondo et al. (2012)</td>
<td>Mixed integer formulation solved with commercial tool, 1st echelon being TL</td>
<td>1</td>
<td>2</td>
<td>20</td>
</tr>
<tr>
<td>Gianessi et al. (2012)</td>
<td>Mixed integer formulation approximated and relaxed to solve the problem in a heuristic way</td>
<td>1</td>
<td>5</td>
<td>50</td>
</tr>
</tbody>
</table>

5. Conclusion

In this paper we have overviewed the main optimization problems for multi-echelon distribution with cross-docking for tactical planning. After an overview of the scientific literature, we proposed, via a systematic analysis method combining both lexical and meta-narrative methods, a standard definition and notation for the route optimization problem, as well as adapted Gonzalez-Feliu’s (2012b) modeling framework to the general multi-depot case. We have also synthesized the main solving method, making a classification on three sets: exact, system heuristics and approximation-based method. Since nowadays three reference instances sets are used respectively for TTRP, 2E-CVPR and 2E-LRP (Chao, 2002; Gonzalez-Feliu et al., 2007; Nguyen et al., 2012), it is important to discuss and affirm the similarities and differences of those problems. As stated by some authors (Nagy & Salhi, 2007; Gonzalez-Feliu, 2012b; Drex1, 2013) and formally proven in this paper, all those problems
are similar and need to be taken into account as subvariants of the same problem. Although publication requirements motivate an explosion of variants and subproblems, it is scientifically important, in the sense of Descartes and Poisson (1724), to think simply and find a unified discourse, in order to canalize the way method is exposed. This joins Ackoff’s (1977) considerations about derive that operations research and having increased in the last 30 years.

In any case, practitioners’ needs and economic context pressure have incited researchers to think more applied, and start to work pluridisciplinarily and rationally, in order to achieve operability and applicability objectives. Research on multi-echelon routing approaches is starting to be unified, and as shown by several recent works, those needs will lead to the development of real-life edged operable tools, in line with the last trends of management such as collaboration or sustainability (Gonzalez-Feliu & Morana, 2010, 2011). The main research directions we identify are three. First is that of developing exact methods, not for the sake of beating precedent works, but with an aim of providing exact solutions for medium-sized instances (like in Contardo et al., 2012), in order to build a set of realistic test cases that can be used to compare heuristics. Second is that of heuristics methods, which need to be easy to adapt to different constraints and variants related to real LTL transport, but also robust and performing to show practitioners the interest of using them in their tactical and/or operational plans. Last but not least is that of applications and tools, mainly for decision support, which are less popular in operations research (Note 6) but become a necessity in practice.

Acknowledgements
The author also aims to thank anonymous referees and the Editor-in-Chief for their valuable comments that helped to improve the paper. The author also liked to apologize if readers find some works are missing, the unification exercise was difficult and if some works are missing it was involuntary. If this is the case, author asks and thanks in advance anyone who found missing works to communicate him, in order to update the literature review.

References


Brunswicker, J. (1986). *Optimale Standort- und Tourenplanung für die Rohmilcherfassung eines Molkereibetriebes Lit, Munster*.


Notes

Note 1. It is known that nowadays inventory costs are higher than those related to manufacturing or transportation (Brewer et al., 2001; Lambert, 2008).

Note 2. Note that those instances have been developed in the context of authors PhD. Theses, which explore the same problem and are consecutive (Gonzalez-Feliu, 2008; Mancini, 2011). Five sets are developed, the first three derived from Christofides and Eilon’s (1969) CVRP instances (Gonzalez-Feliu, 2008), the other two being specifically constructed (Mancini, 2011).

Note 3. Able to find exact optimums for instances up to 2 satellites and 21 customers.

Note 4. Able to find exact optimums for instances up to 2 satellites and 32 customers.

Note 5. Able to find exact optimums for instances up to 5 satellites and 100 customers.

Note 6. This fact could be seen as a «denaturation» of operations research, which aims to be applied (Winston and Goldberg, 2004).