Multi-Echelon Distribution Systems in City Logistics

Simona Mancini

1 Department of Control and Computer Engineering, Politecnico di Torino, Italy

Abstract

In the last decades, the increasing quality of services requested by the customer, yields to the necessity of optimizing the whole distribution process. This goal may be achieved through a smart exploitation of existing resources other than a clever planning of the whole distribution process. For doing that, it is necessary to enhance goods consolidation. One of the most efficient way to implement it is to adopt Multi-Echelon distribution systems which are very common in City Logistic context, in which they allow to keep large trucks from the city center, with strong environmental advantages. The aim of the paper is to review routing problems arising in City Logistics, in which multi-echelon distribution systems are involved: the Two Echelon Location Routing Problem (2E-LRP), the Two Echelon Vehicle Routing Problem (2E-VRP) and Truck and Trailer Routing Problem (TTRP), and to discuss literature on optimization methods, both exact and heuristic, developed to address these problems.

Keywords: City Logistics, Freight Distribution, Multi-Echelon systems

1. Introduction

The sudden change of habits in the modern society, the advance of progress, the achievement of welfare and prosperity and frenzy increase of life rhythm, yielded to the necessity to find new solutions for the management of freight distribution, to reach a higher level of efficiency. This goal may be achieved through a better exploitation of presently available resources, a clever planning of the whole distribution process, a smart network design and a strictly collaboration among shipping companies (Mancini 2013). In fact, in order to improve the efficiency of freight movements it is necessary to stop considering each firm, shipment or vehicles individually, starting to look at them as components of an integrated logistic system. Such kind of approach implies the consolidation of loads of different shippers and carriers on the same vehicle, or, more generally, on the same service, and an efficient coordination of the resulting transportation activities. One of the most efficient way to implement goods consolidation is to adopt Multi-Echelon distribution systems, which allow to split the transportation chains in different legs, in each one of which, goods are consolidated at facilities, where they are sorted and carried on other vehicles which perform the

* Corresponding author: Simona Mancini (simona.mancini@polito.it)
delivery to the customers or to another set of facilities. Depending on the context analyzed, facilities may vary from parks, where the loading/unloading operations are performed, to true intermodal exchanging centers (Verlinde et al., 2012). Beside the significant economic advantages yielded by this kind of distribution approach, Multi-Echelon systems may results more efficient, respect to standard approaches, also from an environmental point of view. The sustainability of the system, which in the past was not considered as important as the economic implications, nowadays plays a crucial role on the distribution operations planning. In the last decades, more attention has been given to the respect for the environment and to the reduction of the impact of nuisance factors (traffic congestion, pollution, exc.), in order to preserve the quality of life in large urban areas. In response to this problem a new area of transports planning, called City Logistic, has emerged. In a City Logistic context, Multi-Echelon, and, in particular, Two-Echelon systems may yield great advantages from the environmental point of view, in which freight, coming from external depots, is delivered, by trucks, to satellites placed around the urban area, where it is loaded in small environmental friend vehicles, performing delivery to customers inside the city. In this way, it is possible to keep large trucks from the city center, with strong advantages on the quality of the air, the pollution level and the traffic congestion. This chapter is organized as follows. Section 2 is devoted to urban freight distribution analysis. More in detail, the relation between distribution systems and City Logistics challenges is discussed. Section 3 deals with Multi-Echelon routing and location problems arising in the City Logistics context, while in Section 4 a systematic comparison of solution approaches is reported and a detailed discussion on differences and analogies on the approaches used to address these problems, is proposed. Finally, in Section 5 conclusions and future developments in this field are reported.

2. City Logistics and urban freight distribution

As it is well recognized, the urban freight transport, which plays a vital role in the sustainable development of cities, recently faces many challenging problems, including high levels of traffic congestion, negative environmental impacts and high energy consumption. Freight carriers are expected to provide higher levels of service at lower prices and the economic welfare joined to the awful performance of the public transports have provided an increase of private vehicles circulation with a consequent increase of congestion of the urban roads (Ambrosini and Routhier, 2004; Ville et al. 2012). In addition to the congestion of the traffic, that means an increasing of the time delivering and consequently of the delivering costs, a negative environmental effect should be considered. Large trucks produce a substantial amount of air pollution in urban areas, and truck crashes can be a trauma for the community. Energy consumption is also an important issue, not only because of the limited amount of natural resources available but also for reducing CO2 emissions and limit the global warming. Nowadays, there is an on-going trend toward urbanization in the world. Cities provide more attractive opportunities for employment, education, cultural and sport activities, etc. The concentration of population in urban areas is observed in most industrialized and developing countries. However, this leads to an expansion of the urban area, which generates freight transport problems due at the lack of appropriate urban logistic policies. The movement of goods within cities requires an high cost in terms of money
as well in terms of time. There are three main different types of urban goods movements (Gonzalez-Feliu et al., 2010b, 2012):

- **Inter-establishment movements** (IEM) or classical freight distribution flows in urban areas, which represent about 40% of the total distance covered.

- **End-consumer movements** (ECM) commonly identified with shopping trips. In the last decade, other flows have been included in this category, like those derived from home deliveries and reception points or other customer-to-customer movements. Their share with respect to the total number of the total distance covered is about 50%.

- **Urban management movements** (UMM), related to public infrastructure maintenance, building works, waste management and other urban space management functions. They represent about 8% of the total km covered.

This paper is focused on the analysis of the first type of movements, (IEM). Trucks carry most goods directly to the customer and road transport has become expensive due to the decreasing load factors of trucks. In response to this problem a new area of transports planning, called City Logistic, has emerged. Although different definitions of City Logistics have been proposed in the literature, the most broadly used one is those proposed by Taniguchi et al. (2001) which have been adopted also in Dablanc (2007) and Crainic (2008): City Logistic is the process of totally optimizing urban logistics activities by considering the social, economic, and environmental impact of urban freight movement and it provides an opportunity for the development of innovative solutions that allow to improve the quality of life in urban areas. This modelling approach is relatively recent and so it is not yet commonly used in city planning. Although several cities have already implemented some City Logistic initiatives, only a limited number of evaluation tools have been developed for predictin the consequences of such schemes.

Urban consolidations centers (UCCs) are a very popular type of facilities. Consolidation is broadly used in urban freight distribution and many case studies have been presented in literature, (Danielis et al., 2010, Allen et al., 2012). In Allen et al. (2007), a UCC has been defined as a logistics facility situated in relatively close proximity to the geographic area that it serves (be that a city centre, an entire town or a specific site such as a shopping centre), to which many logistics companies deliver goods destined for the area, from which consolidated deliveries are carried out within that area, in which a range of other value-added logistics and retail services can be provided. The first UCCs found in literature were private or semi-private initiatives developed on a viewpoint of economy and optimization (Ville et al., 2012). Later, environmental and social issues where the motivations of developing such systems for urban goods distribution (Gonzalez-Feliu, 2008).

Related to consolidation, we observe two main distribution schemes (Crainic et al., 2013), i.e. direct shipping schemes and advanced distribution schemes (Crainic et al.,
The main difference between these approaches is that direct schemes follow classical LTL strategies, directly related to single-echelon VRP, which are close to classical freight transport systems and need big efforts in terms of number of vehicles, if we consider that an increasing number of cities adopt policies which impose limits on the maximum size of trucks authorised to access the historical centre. Those systems can lead to an increase of the number of vehicles, which is in contrast with the objective of congestion reduction (Gonzalez-Feliu and Salanova, 2012). The second approach aims to rationalize the usage of vehicles by proposing a two-echelon system where vehicles start from the UCC, travel to a vehicle reception point (Boudouin et al., 2013) where good is consolidated into small environmental friend vehicles, like cargohoppers (Van Duin et al., 2013) or electric assisted trycicles (Gonzalez-Feliu et al., 2013). This approach has a potential and starts to be addressed by different cities (see for example, La Petite Reine, La tournée in France, or the Cargohopper project in The Netherlands), and new tactical planning tools need to be developed.

Figure 1: A 2EVRP scheme

From a physical point of view, a Two-Echelon transport system operates as follows (see also Figure 1):
- freight arrives at an external zone, the depot, where it is consolidated into the 1st-echelon vehicles;
- each 1st-echelon vehicle travels to a subset of satellites that will be determined by the model and then it will return to the depot;
- at each satellite, freight is transferred from 1st-echelon vehicles to smaller, environmental friendly vehicles, belonging to 2nd-echelon fleet;
- each 2nd-echelon vehicle performs a route to serve the designated customers, then travels back to its departure satellite.

In Crainic et al. (2009) and Crainic et al. (2004), authors focus on a City Logistics planning issue, the integrated short-term scheduling of operations and management of resources, for the general case involving a two-tier distribution structure, while in Benjelloun et al. (2010) is reported a taxonomy of City Logistics projects that provides
the means to explore similarities and differences in the elements characterizing various City Logistics initiatives. Although multi-echelon transportation systems are very common in real cases, they are not always related to urban consolidation (Gonzalez-Feliu, 2012b). Furthermore, literature on optimization methods addressing such systems in their wholeness is limited, because they are usually decomposed into a sequence of single-echelon distribution cases. Moreover, most of the optimization tools used for tactical and operational planning derive from methods for the Vehicle Routing Problem (VRP). This family of problems has been deeply studied, but refers essentially to single echelon systems (for detailed surveys, see Toth and Vigo, 2002 and Golden et al., 2008). In current planning practices, transportation cost optimization for a N-echelon system is usually made by splitting the system into N single-echelon problems, then optimizing them, but some authors have started to analyze the advantages of considering the global costs of the system in the optimization process (Crainic, 2008) and several studies dealing with multi-echelon distribution optimization using global vehicle routing based approaches may be found in the literature.

The main problem studied in literature is the Vehicle Routing Problem with Time Windows (VRPTW), with access restriction (Quak and De Koster, 2006, 2009; Deflorio et al., 2012; Munuzuri et al., 2012; Munuzuei and van Duin, 2013) and the dynamic nature of travel times (Taniguchi et al., 1999; 2001; Zeimpekis, 2009). In particular, it is of great interest to analyze the influence of time windows on the costs of urban freight distribution services in City Logistics applications, (Taniguchi et al., 2012). The new advances in the technologies have been a positive factor for the development of new markets and new consumer needs: the growth of e-commerce and postal shopping, (Gonzalez-Feliu et al., 2012), have reinforced the importance of just in time policies in freight distribution. For this reason, the time within which a shipping company may guarantee the delivery of the products, has become a relevant index of the service quality. The total travel time of a vehicle trip depends not only on travelled distance, but it is also affected by waiting and access time, congestion, deadlines, service features, etc. In addition the generalised economic and financial crisis yield to the necessity of a readjustment on the freight transportation strategies that have to be included in the main logistics tactical decisions. For these reasons, it is important for a distribution system to ensure the efficiency while maintaining a service quality defined by the time windows or other quality indices. These two factors are usually related: the higher the quality, the higher the cost incurred, but this relation is not trivial and these two factors are not considered in the same manner by the different shippers and transport carriers (Danielis et al., 2005). Moreover, an estimation of the cost level is useful to compare different service settings and support the decision, on the base of quantitative indicators (Taniguchi and Van Der Heijden, 2000, Pluvinet et al. 2012). Another important emergent issue arising in City Logistics, concerns collaborative distribution systems, (Gonzalez-Feliu and Salanova, 2012, Gonzalez-Feliu et al. 2010a, Thompson and Hassall 2012), which are not explicitly addressed in this paper.
3. Multi-Echelon location routing problems arising in City Logistics

Although Multi-Echelon distribution systems are commonly used in practical applications, from an academic point of view they have been rarely considered in their wholeness. Both the location and routing aspects of the problem, which are strictly connected and correlated to each other, play a crucial role. The literature related to multi-echelon routing problems is very poor. The Multi-Echelon system, and the two-echelon system in particular, refer mainly to supply chain and inventory problems, (Daskin et al. (2002) and Verrijdt and De Kok (1995)). These problems do not use an explicit routing approach for the different levels, focusing more on the production and supply chain management issues. In this section, both multi-echelon location routing problems and multi-echelon routing problems are addressed. The first ones arise at the tactical and strategic levels, in which decision about how many satellites or platforms should be open and where they should be located, must be taken, while the second ones are used at the operational level, in which the satellites location is already known, and the focus of the problem is more on the routing phase. A complete review on problems and issues related to each one of the main decision-making levels (tactical, strategic and operational) is given Crainic and Laporte (1997), while in Gonzalez-Feliu (2013a) an analysis of routing problems arising in Multi-Echelon distribution systems is reported.

3.1. The Two-Echelon location routing problem (2E-LRP)

Two-Echelon Location-Routing Problems arises in several logistics context, as freight distribution in urban areas, express delivery services, large distribution to grocery and stores, and all the activities in which products available at primary facilities pass through secondary facilities before to be delivered to the customers. It is important to remark that, for a correct and efficient use of this kind of systems, it is of crucial importance to properly locate the facilities in the most strategic point, to open only the necessary number of facilities, saving, in this way, opening and management costs of under used or useless facilities, and to properly assign customers to facilities in order to reduce routing costs necessary to perform the delivery to the customers.

The first paper addressing two-level routes optimization is Jacobsen and Madsen (1980), in which a comparison among heuristics methods is reported, followed by Madsen (1983), in which realistic size instances are solved. Multi-Echelon location routing problems have been introduced in Laporte (1988) but formally presented in Nguyen et al. (2011). Some hints on multi-echelon systems in the context of location routing problems may be found also in Min et al. (1997) and Nagy et al. (2007). A unified notation has been proposed in Gonzalez-Feliu (2012b), where the main concepts of multi-echelon distribution with cross-docks are also discussed. In this section, the attention is focused on the Two-Echelon location routing problem (2E-LRP), which consists in defining number and location of primary and secondary facilities, performs the allocation operations, i.e. assign each final customers to an open secondary facility, and each secondary facility to an open primary facility, satisfying capacity facility constraints, and solve the resulting routing problem, identifying how many vehicles, for each fleet, are used, by which vehicle each customer is served, and in which order the vehicle performs its deliveries.
This problem have been addressed in Boccia et al. (2010), where the authors propose a Tabu Search heuristic based on the integration of the nested and iterative approach. The problem is decomposed in its two main components, i.e. two location routing problems. Each component, in turn, is decomposed in a capacitated facility location problem (CFLP) and a multi-depot vehicle routing problem (MDVRP). A bottom-up approach is used, i.e. first echelon solution is built and optimized on second echelon solution. The Tabu Search operates on each echelon in two coordinated and integrated phases (location and routing). In Nguyen et al. (2010) a hybrid metaheuristic is presented, that is composed by the interaction of a greedy randomized adaptive search procedure (GRASP) and an evolutionary/iterated local search (ELS/ILS), using a Tabu Search to solve the 2E-LRP. The GRASP uses, alternatively, three constructive heuristics followed by local search to generate the initial solutions. From a solution of GRASP, an intensification strategy is carried out by a dynamic alternation between ELS and ILS. In this phase, each child is obtained by mutation and evaluated through a splitting procedure of giant tour followed by a local search. Moreover, a GRASP combined with a path-relinking strategy and a multi-start iterated local search combined with Tabu and Path-Relinking have been presented in Nguyen et al. (2012b) and Nguyen et al. (2011a), respectively, while a Variable Neighborhood Search (VNS) approach has been proposed in Schwengerer et al. (2012). A branch and cut and a new formulation for the problem have been proposed in Contardo et al. (2012), while a fast heuristic and a Simulated Annealing (SA) are proposed in Zegordi and Nikbakhsh (2009).

3.2. The Two-Echelon vehicle routing problem (2E-VRP)

In some logistics problems, especially at the operational level, facilities number and positions have been decided in a preliminary phase, which means that in this case the location part of the problem is already solved. In these cases, the Multi-Echelon location routing problem, (ME-LRP), becomes a Multi-Echelon vehicle routing problem, (ME-VRP), in which there are two still open issues, the customers allocation (decide to which facilities each customer must be assigned) and the resulting routing problems. The most common version of Multi-Echelon Vehicle Routing Problem used in practice is the Two-Echelon Vehicle Routing Problem (2E-VRP), where just two levels are considered. For a survey on two-echelon transportation optimisation the reader may refer to Gonzalez-Feliu (2011).

This problem has been formally introduced in Gonzalez-Feliu et al. (2007) and Perboli et al. (2011), where several variants of the problem are also described. More in details, variants may be grouped in three different kinds:

- Basic variants without time dependence

- Two-echelon capacitated VRP (2E-CVRP). This is the simplest version of multiechelon VRPs. At each level, all vehicles belonging to that level have the same fixed capacity. The size of the fleet of each level is fixed and known in advance, and there exists an upper bound on the number of
vehicle which can start from the same satellite. The objective is to serve customers by minimizing the total transportation cost, satisfying the capacity constraints of the vehicles. There is a single depot and a fixed number of capacitated satellites. All the customer demands are fixed, known in advance, and must be compulsorily satisfied. Moreover, no time window is defined for the deliveries and the satellite operations. For the 2nd level, the demand of each customer is smaller than each vehicles capacity and cannot be split in multiple routes of the same level.

- **Basic variants with time dependence**
  - Two-echelon VRP with time windows (2E-VRPTW). This problem 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, if they are violated a penalty cost is due.
  - Two-echelon VRP with satellites synchronization (2E-VRP-SS). In this problem, time constraints on the arrival and the departure of vehicles at the satellites are considered. In fact, the vehicles arriving at 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 first case, every time a 1st-level vehicle unloads its freight, 2nd-level vehicles must be ready to load it (this constraint is formulated through a very small hard time window). In the second case, if 2nd-level vehicles are not available, a penalty is paid. If the satellites are capacitated, constraints on loading/unloading operations are incorporated, such that in each time period the satellite capacity in not violated. For a complete survey on synchronization in vehicle routing problem the reader may refer to Drexl (2012).

- **Other variants**
  - Two-Echelon Multidepot problem. In this problem the satellites are served by more than one depot.
  - 2E-CVRP with pickup and deliveries (2E-VRPPD) In this case satellites are considered as intermediate depots, where both the freight that has been picked up from the customers and that which must be delivered to the customers are stored.

Although all these variants have been introduced by the authors, only the 2E-CVRP has been explicitly treated from an optimization point of view. More in detail, we observe MIP formulations for the problem (Gonzalez-Feliu et al. 2007; Gonzalez-Feliu, 2008; Perboli et al., 2011; Santos et al., 2012, Jepsen et al., 2013), valid inequalities
(Perboli et al., 2010; Jepsen et al., 2012). Exact methods have been presented in Gonzalez-Feliu (2008), Perboli et al., (2010), Santos et al. (2012 and Jepsen et al.(2013), while fast lower bounds have been proposed in Mancini (2012), Mancini (2011) and Crainic et al. (2008b). Several metaheuristics, able to find highly accurate solution in very short computational time can also be find in literature. In Crainic et al. (2011) the authors proposed a Multi-Start heuristic, while a Path-Relinking a GRASP with Path-Relinking and a Reactive GRASP with Path-Relinking have been proposed respectively, in Crainic et al. (2010a), Crainic et al. (2013) and Crainic et al. (2012a), while an Adaptive Large Neighbourhood Search procedure have been presented in Hemmelmayr et al. (2012). In Wang et al. (2011) an hybrid ant colony heuristic is proposed. A fast clustering based local search algorithm, able to solve instances up to 250 customers and 10 satellites, has been proposed in Crainic et al. (2008a). This algorithm has been used for a satellite location analysis and a comparison between two-echelon and single echelon approach in Crainic et al. (2010b) and for an analysis of the impact of generalized travel costs on the 2E-VRP in Crainic et al. (2012b). A real application of 2E-VRP addressing milk collection in western Norway has been presented in Hoff and Løkketangen (2008) where the authors propose a Tabu Search heuristic to solve the problem.

A problem closely relate to the 2E-VRP is the so-called vehicle routing problem with cross-docking (VRPCD), introduced for the first time in Lee et al. (2006). The VRPCD can be defined as the problem of transporting products from a set of suppliers (pickup nodes) to a set of customers (delivery nodes) via a single cross-dock. Products are picked-up from the suppliers by a fleet of homogeneous vehicles, then it is consolidated at the cross-dock, and immediately delivered to customers by the same set of vehicles, without intermediate storage. While the 2E-VRP deals with a single depot and a set of intermediary facilities where consolidation is performed, the VRPCD consider a set of depot and just one intermediary facility. Lee et al. (2006) addressed a simplified version of the problem in which all the vehicles are supposed to arrive simultaneously at the cross-dock. Dondo et al. (2009) proposed an MILP model that relies on a continuous-time representation, which have been extended in Dondo et al. (2011) where multiple types of products are handled, and goods delivery may be performed either via the cross-dock or via direct shipping.

3.3. The Truck and Trailer routing problem (TTRP)

Another problem arising in urban distribution is the Truck and Trailer Problem (TTRP), which can be seen as a variant of 2EVRP with certain specific features. Each vehicle is compose by a truck and a trailer. Some customers can be served directly by a 1st level trip while in areas having limited accessibility, the trailer must be detached at satellites, and second level trips are performed by the truck alone to reach customers. These constraints exist in many practical situations. This problem is quite similar to our problem to the 2EVRP, because it involves two routing levels strictly interconnected. The main difference is that, while in the 2EVRP freight must pass through the satellites, because it must be delivered to the customer only by second level vehicles, in the TTRP the delivery of certain customers can be directly performed by first level vehicle (truck and trailer) without passing through satellites. In Lin et al. (2009) a simulated annealing based heuristic is presented. Its main characteristics are the combination of a two-level
solution representation with the use of dummy depots/roots, and the random neighborhood structure which utilizes three different types of moves: two classical operators like insertion and swap, and the change of vehicle service type, an operator created ad hoc for the problem. In Scheuere (2006) two simple, but efficient, construction heuristics and a tabu search heuristic with a variable number of sub-tours for further improvement, are introduced, while a tabu search is presented in Chao (2002). For a survey on the applications of TTRP, the reader may refer to Drexl (2013), while a survey on formulations, exact and heuristic methods can be find in Drexl (2007).

In Villegas et al. (2010), the Single Truck and Trailer Routing Problem with Satellite Depots(STTRPSD), a particular version of the TTRP, is introduced. In the STTRPSD a vehicle composed of a truck with a detachable trailer serves the demand of a set of customers reachable only by the truck without the trailer. This accessibility constraint implies the selection of locations where to park the trailer before performing the trips to the customers. This version of the problem is the most similar to the 2EVRP while all the deliveries must be performed by the same kind of vehicle (track without the trailer), even if in this case, only one vehicle is considered. The authors propose two metaheuristics based on greedy randomized adaptive search procedures (GRASP), variable neighborhood descent (VND) and evolutionary local search (ELS), to solve this problem. In Tan et al. (2006) a multi-objective version of the TTRP is presented and it is solved by an evolutionary algorithm hybridized with a local search algorithm.

4. Solving methods classification

This section is devoted to a classification of the algorithm proposed in literature, based on the type of approach (decomposition, systemic and approximation) and, for heuristic methods, on the class which they belong to. Furthermore, a discussion on differences and analogies among the three addressed problem (2E-LRP, 2E-VRP and TTRP). First of all, some classes of heuristic are defined, and a brief description is given for each class:

1. **Local Search (LS):** local search based heuristics, iteratively starts from a current solution S’ and analyse a neighbourhood surrounding S’ in the solution search space. (Aarts and Lenstra, 1997). The exploration of the neighbourhood could be exhaustively carried out (Best Improvement) and after that the best solution in the neighbourhood is taken as current best and the algorithm is restarted or the exploration can be interrupted as soon as an improving solution is found and immediately restarted from the new current best, (First Improvement). The first strategy focus more on intensification of the search on a limited area, while the second one allows to explore a wider area of the search space. There is not a strategy which overcome the other, instead both approaches could be profitable depending on the problem addressed. LS is commonly used in several applications dealing with different classes of combinatorial optimization problems, because it is easy to implement, and it requires very short computational times. On the other
hand LS algorithm often remain trapped into local minima, which means that the quality of the solution obtained may be sensibly vary according the starting point of the search, and that the accuracy level reached is generally low. Nevertheless, LS is broadly and usefully applied as an intensification tool into a metaheuristic framework as Multi Start heuristic (Marti et al. 2010) or Guided Local Search (Voudoris et al. 2010) or combined with Evolutionary Algorithms in hybrid or memetic heuristics, (Moscato and Cotta, 2010).

2. **Tabu Search (TS):** Tabu search is a higher level heuristic procedure for solving optimization problems, designed to guide other methods (or their component processes) to escape the trap of local optimality. Tabu search has been applied with promising results to a wide variety of classical and practical optimization problems. It uses flexible structures memory (to permit search information to be exploited more thoroughly than by rigid memory systems or memoryless systems), conditions for strategically constraining and freeing the search process (embodied in tabu restrictions and aspiration criteria), and memory functions of varying time spans for intensifying and diversifying the search (reinforcing attributes historically found good and driving the search into new regions). One major issue with Tabu Search is that it is only effective in discrete search spaces. It is rare that a search would visit the same real-value point in space multiple times, making a tabu list worthless. Another problem with Tabu Search is that if the search space is very large or of high dimensionality, it remains within a small area of the search space, partially but not completely overcoming LS characteristics to be remained trapped into local minima. For a complete survey on TS, the reader may refer to Glover (1990) and Glover and Laguna (1993).

3. **Simulated Annealing (SA):** Simulated annealing is a local search algorithm (meta-heuristic) capable of escaping from local optima. Its ease of implementation, convergence properties and its use of hill-climbing moves to escape local optima have made it a popular technique over the past two decades. It is typically used to address discrete, and to a lesser extent, continuous optimization problems. In its original form SA is based on the analogy between the simulation of the annealing solids and the problem of solving large combinatorial optimization problems. For this reason the algorithm became known as simulated annealing. In condensed matter physics, annealing denotes a physical process in which a solid in a heat bath is heated up by increasing the temperature of the heat bath to a maximum value at which all particles of the solid randomly arrange themselves in the liquid phase, followed by cooling through slowly lowering the temperature of the heat bath. In this way, all particles arrange themselves in the low energy ground state of a corresponding lattice, provided the maximum temperature is sufficiently high and the cooling is carried out sufficiently slowly. (Nikolaev and Jacobson, 2010) This procedure has been resumed by operation researcher, creating this meta-heuristic, in the following way. At the beginning of the search, even slightly worsening solution (within a given
threshold) are accepted, in order to allow the algorithm to explore a wider area of the search space, (representing the liquid phase of the matter, in which the particles are free to arrange themselves); increasing the time of the search, accepting rules for worsening solutions become more and more selective, until only improving solution are accepted by the algorithm (representing the achievement of the solid state of the matter). In this way, the algorithm integrates a diversification phase in which the goal is to explore a wide part of the search space and an intensification phase in which the search is focused on the most promising region. This method has been applied with relevant results to several combinatorial optimization problems. Nevertheless, to obtain high quality performances, SA parameters, determining the solution accepting rule, must be carefully tuned, and for this reason, this technique is less frequently used respect to more intuitive algorithms, like LS based heuristics.

4. Variable Neighbourhood Search (VNS): Variable Neighbourhood Search is a heuristic method which propose an innovative approach: the change of neighborhood in the search. (Hansen and Mladenovic, 2001, Hansen et al. 2010). Contrary to other metaheuristics based on local search methods, VNS does not follow a trajectory but explores increasingly distant neighborhoods of the current incumbent solution, and jumps from this solution to a new one if and only if an improvement has been made. In this way often favorable characteristics of the incumbent solution, e.g., that many variables are already at their optimal value, will be kept and used to obtain promising neighbouring solutions. Moreover, a local search routine is applied repeatedly to get from these neighbouring solutions to local optima. This routine may also use several neighbourhoods. Therefore, to construct different neighbourhood structures and to perform a systematic search, one needs to have a way for finding the distance between any two solutions, i.e., one needs to supply the solution space with some metric (or quasi-metric) and then induce neighbourhoods from it. The definition of the neighbourhoods and of the order in which they should be applied is not a trivial issue to be carried out, and this process may be highly problem-sensitive, i.e. a neighbourhood structures highly performing on a problem, could not be able to reach satisfactory results on another problem.

5. Greedy Randomized Adaptive Search Procedure (GRASP):
GRASP is a multistart two-phase metaheuristic for combinatorial optimization. (Resende e Ribeiro, 2010) The first phase is a construction phase that builds an initial solution using a greedy randomized procedure, whose randomness allows solutions in different areas of the solution space to be obtained. The second phase is a local search phase that improves these solutions. The etymology of GRASP derives from the characteristics of the algorithm. The greediness is given by the fact that, at each iteration, the solution is created following a greedy procedure, i.e. it is constructed step by step. A random component is considered, in fact, at each iteration the algorithm randomly choose among different alternatives (i.e. including arc a,b,c or d in the solution, or opening facility 1,2,3 or 4, etc..), each one of
which characterized by a probability of being chosen. Finally the procedure is adaptive, because, it is able to learn from the past, and to take advantage from the information coming out from the search process.

6. **Evolutionary Algorithms (EA):** EAs are computer programs that attempt to solve complex problems by mimicking the processes of Darwinian evolution. In an EA a number of artificial creatures search over the space of the problem. They compete continually with each other to discover optimal areas of the search space. It is hoped that over time the most successful of these creatures will evolve to discover the optimal solution. The artificial creatures in EAs, known as individuals, are typically represented by fixed length strings or vectors. Each individual encodes a single possible solution to the problem, under consideration. The EA is started with an initial population of size \( \mu \) comprising random individuals. Every individual is then assigned a fitness value. To generate a fitness score the individual is decoded to produce a possible solution to the problem. The value of this solution is then calculated using the fitness function. Population members with high fitness scores therefore represent better solutions to the problem than individuals with lower fitness scores. Following this initial phase the main iterative cycle of the algorithm begins. Using mutation (perturbation) and recombination operators, the \( \mu \) individuals in the current population produce children. The children are assigned fitness scores. A new population of \( \mu \) individuals is then formed from the \( \mu \) individuals in the current population and the children. This new population becomes the current population and the iterative cycle is repeated. At some point in the cycle evolutionary pressure is applied. That is, the Darwinian strategy of the survival of the fittest is employed and individuals compete against each other. This is achieved by selection based on fitness scores, with fitter individuals more likely to be selected. One of the most commonly used type of EA are the Genetic Algorithms (GA), (see Holland, 1975 and Reeves, 2010).

Heuristic and Meta-Heuristic methods are capable to obtain high quality solutions in reasonable computational time, but they do not guarantee the optimality of the obtained solutions, and they are not able to prove the optimality, even if they reach it. An alternative approach is to develop exact methods, which can solve problems to the optimality. To this class belong both mathematical models, Linear Programming (LP) models, which can be solved through commercial solvers or ad-hoc developed methods. Unfortunately these kind of methods generally have computational time exponential increasing with the size of the problems, and therefore they may cannot be used to address large size instances.

In the following are reported tables containing a list of the algorithms proposed and their characteristics, as size of instances addressed (n° of depots, n° of intermediate facilities, and n° of customers), type of approach (decomposition or systemic), and if the work has been applied in a real context, for each one of the main problem addressed in
More in details, Table 1 deals with solving methods for the 2E-LRP, Table 2 for 2E-VRP and Table 3 for TTRP.

### Table 1. Solving methods for the 2E-LRP

<table>
<thead>
<tr>
<th>Authors</th>
<th>Type of algorithm</th>
<th>Size</th>
<th>Type of approach</th>
<th>Real context</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jacobsen and Madsen (1980)</td>
<td>Construction heuristics</td>
<td>One depot, three IF and 4510 customers</td>
<td>Decomposition</td>
<td>Yes</td>
</tr>
<tr>
<td>Madsen (1983)</td>
<td>Construction heuristic with systemic LS post-optimization</td>
<td>One depot, three IF and 4510 customers</td>
<td>Systemic</td>
<td>Yes</td>
</tr>
<tr>
<td>Nguyen et al. (2010)</td>
<td>EA with systemic LS post-optimization</td>
<td>One depot, 10 IF and 250 customers</td>
<td>Systemic</td>
<td>No</td>
</tr>
<tr>
<td>Boccia et al. (2010)</td>
<td>Construction heuristic with systemic TS post-optimization</td>
<td>Five depots, 20 IF and 200 customers</td>
<td>Systemic</td>
<td>No</td>
</tr>
<tr>
<td>Nguyen et al. (2012a)</td>
<td>Construction heuristic with systemic TS post-optimization</td>
<td>One depot, 10 IF and 250 customers</td>
<td>Systemic</td>
<td>No</td>
</tr>
<tr>
<td>Nguyen et al. (2012b)</td>
<td>Construction heuristic with systemic GRASP post-optimization</td>
<td>One depot, 10 IF and 250 customers</td>
<td>Systemic</td>
<td>No</td>
</tr>
<tr>
<td>Schwengerer et al. (2012)</td>
<td>Construction heuristic with systemic VNS post-optimization</td>
<td>Test cases from Nguyen et al. (2010)</td>
<td>Systemic</td>
<td>No</td>
</tr>
<tr>
<td>Contardo et al. (2012)</td>
<td>Mathematical formulation solved by an exact method</td>
<td>Test cases from Nguyen et al. (2010)</td>
<td>Systemic</td>
<td>No</td>
</tr>
<tr>
<td>Zegordi and Nikbakhsh (2009)</td>
<td>Construction heuristic with systemic SA post-optimization</td>
<td>10 depots, 50 IF and 100 customers</td>
<td>Systemic</td>
<td>No</td>
</tr>
</tbody>
</table>

With those tables we extend the work of Gonzalez-Feliu (2013b) which used a similar classification focusing in the practical aspects of optimization, whereas the present paper focuses on the mathematics and computational frameworks.
### Table 2: Solving Methods for the 2E-VRP

<table>
<thead>
<tr>
<th>Authors</th>
<th>Type of algorithm</th>
<th>Size</th>
<th>Type of approach</th>
<th>Real context</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gonzalez-Feliu et al. (2007)</td>
<td>Mathematical Formulation solved by a commercial solver</td>
<td>One depot, 4 IF and 50 customers</td>
<td>Systemic</td>
<td>No</td>
</tr>
<tr>
<td>Gonzalez-Feliu (2008)</td>
<td>Mathematical Formulation solved by a commercial solver</td>
<td>One depot, 4 IF and 50 customers</td>
<td>Systemic</td>
<td>No</td>
</tr>
<tr>
<td>Perboli et al. (2011)</td>
<td>Mathematical Formulation solved by a commercial solver</td>
<td>One depot, 5 IF and 50 customers</td>
<td>Systemic</td>
<td>No</td>
</tr>
<tr>
<td>Santos et al. (2012)</td>
<td>Mathematical Formulation solved by a commercial solver</td>
<td>One depot, 5 IF and 50 customers</td>
<td>Systemic</td>
<td>No</td>
</tr>
<tr>
<td>Jepsen et al. (2013)</td>
<td>Mathematical Formulation solved by a commercial solver and valid inequalities</td>
<td>One depot, 5 IF and 50 customers</td>
<td>Systemic</td>
<td>No</td>
</tr>
<tr>
<td>Perobli et al. (2010)</td>
<td>Exact method</td>
<td>One depot, 5 IF and 50 customers</td>
<td>Systemic</td>
<td>No</td>
</tr>
<tr>
<td>Crainic et al. (2008)</td>
<td>Construction heuristic with systemic LS post-optimization</td>
<td>One depot, 10 IF and 250 customers</td>
<td>Decomposition</td>
<td>No</td>
</tr>
<tr>
<td>Crainic et al. (2010)</td>
<td>Construction heuristic with systemic LS post-optimization</td>
<td>One depot, 5 IF and 50 customers</td>
<td>Decomposition</td>
<td>No</td>
</tr>
<tr>
<td>Crainic et al. (2011)</td>
<td>GRASP</td>
<td>One depot, 5 IF and 50 customers</td>
<td>Decomposition</td>
<td>No</td>
</tr>
<tr>
<td>Crainic et al. (2013)</td>
<td>GRASP</td>
<td>One depot, 5 IF and 50 customers</td>
<td>Decomposition</td>
<td>No</td>
</tr>
<tr>
<td>Hemmelmayr et al. (2012)</td>
<td>Construction heuristic with systemic LS post-optimization (Adaptive Large Neighborhood Search)</td>
<td>One depot, 10 IF and 200 customers</td>
<td>Systemic</td>
<td>No</td>
</tr>
<tr>
<td>Hoff and Lokketangen (2008)</td>
<td>Construction heuristic with systemic TS post-optimization</td>
<td>Real size instances</td>
<td>Systemic</td>
<td>Yes</td>
</tr>
<tr>
<td>Wang et al. (2011)</td>
<td>EA with systemic LS post-optimization</td>
<td>One depot, 4 IF and 50 customers</td>
<td>Systemic</td>
<td>No</td>
</tr>
</tbody>
</table>
### Table 3. Solving Methods for the TTRP

<table>
<thead>
<tr>
<th>Authors</th>
<th>Type of algorithm</th>
<th>Size</th>
<th>Type of approach</th>
<th>Real context</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chao (2002)</td>
<td>Construction heuristic with systemic TS post-optimization</td>
<td>One depot, 150 IF and 199 customers</td>
<td>Systemic</td>
<td>No</td>
</tr>
<tr>
<td>Drexl (2007)</td>
<td>Mathematical formulation solved by exact methods</td>
<td>One depot, eight IF and eight customers</td>
<td>Systemic</td>
<td>No</td>
</tr>
</tbody>
</table>

Decomposition approaches deal with the complexity by splitting the system into two or more sub-systems, then solving them separately but respecting the satellite connection and flow conservation constraints. In general, a first commodity assignment to satellites is made then to each satellite a 2nd stage VRP is associated and solved. In this context, such problems are classical VRPs able to be solved with heuristics and exact methods well established in the literature. Once all the 2nd stage routes have been determined, keeping the resulting information, 1st-stage routes can be defined using also a classical VRP solving method. The advantages of such approaches are that the reality representation they give, is close to the current practices, i.e. to the logical strategies of splitting the system into a set of easily understandable and controllable subsystems.

The main disadvantages of such systems derive from the fact the systemic nature of multi-stage transport is not really integrated into the solving method. However, they correspond to a current practice philosophy in terms of optimization and are very popular in practice, although little diffused in scientific publications.

Systemic route optimization approaches address the system in its wholeness and not by splitting it into sub-problems. This group of methods takes into account the systemic nature of multi-stage transport. However, most works remain on the domain of research since practical applications of this kind of approach are very rare. This may be due to the greater complexity of the systemic approach respect to the decomposition one, which results to be more intuitive and comprehensible for practitioners.

For what concerns solving methods, what comes out from Tables 1, 2 and 3 is that LS based metaheuristics are the most popular methods. In fact, LS is very intuitive, easy to implement and could be a powerful tool to be inserted in a metaheuristic framework (like Multi-Start, Iterated Local Search, exc..) or to be hybridized with evolutionary algorithms. Tabu Search is also very popular, because it puts together the easiness and immediateness of LS with the concept of memory, i.e. it avoid to visit points of the search space which have been already visited during the search process. This strategy limits the possibility of remaining trapped into local minima, which is the greatest
disadvantages of LS. Nevertheless, also TS, when working with search spaces of high dimensionality, tends to remain within a small area of the search space.

Methods which are able to overcome this problem are GRASP and VNS, which popularity is increased in the last years. These metaheuristics are a very powerful tool, but they require an accurate parameters tuning phase. Last but not least, Evolutionary Algorithms are also a high performing tool but they need parameters calibration too and furthermore, encoding a multi-echelon routing (or location routing) problem solution is not a trivial issue. Several exact methods have been developed, consisting of both mathematical formulation and ad-hoc developed algorithms. These methods are able to address only small-medium size instances and within large computational times, therefore they are not suited for application on real instances, although they play a crucial role in the scientific research field because they can be used as a benchmark to measure metaheuristic performances.

For what concern the portability of algorithms developed for a specific problem to the other problems, what can be said is that the 2E-LRP shows a similar structure respect to the 2E-VRP. In fact, in the 2E-VRP two main issues must be carried out: the assignment of customers to a facility, and the corresponding routing problem, while in 2E-LRP, there is a third decision level, higher than the others, which deal with the facilities opening. Therefore, the 2E-VRP, may be seen as a particular case of 2E-LRP where opening costs are null, and then, methods developed for the 2E-LRP can be certainly applied to the 2E-VRP. On the contrary, methods developed for the 2E-VRP cannot manage facilities opening costs, and so, even if they can provide feasible solutions for the 2E-LRP, the quality of these solutions cannot be guaranteed (especially in cases in which the most convenient facilities from a routing point of view have a high opening cost), and algorithm performance may be sensibly vary from instance to instance, and may be strongly influenced by instance layout and characteristics. The TTRP is slightly different from the other two problems, because deal with customers with different requests: some of them may be served by either a truck with trailer or a single truck, while some others only by a single truck. This feature does not appear in the other two problems, and thus, portability of algorithms between TTRP and 2E-VRP and 2E-LRP is very limited.

4. Conclusions and Future Developments

In this chapter, Multi-Echelon distribution systems and the role they play in City Logistics. These systems allow to split the transportation chains in different legs, in each one of which, goods are consolidated at facilities, where they are sorted and carried on other vehicles, which perform the delivery to the customers or to another set of facilities. Depending on the context analysed, facilities may vary from parks, where the loading/unloading operations are performed, to true intermodal exchanging centers. More in detail the aim of the paper is to review routing problems arising in City Logistics, in which multi-echelon distribution systems are involved: the Two Echelon Location Routing Problem (2E-LRP), the Two Echelon Vehicle Routing Problem (2E-
VRP) and Truck and Trailer Routing Problem (TTRP), to discuss literature on optimization methods, both exact and heuristic, developed to address these problems. Analogies and differences in the approach to each one of the three problems are reported and portability of the algorithms among the problems is discussed. Furthermore an analysis on the adaptability of the most common class of metaheuristics to these problems is carried out. Future development in this field could address the integration of Multi-Echelon systems with shippers collaboration strategies, in order to further increment the efficiency of the delivery process.

Acknowledgement
Part of this research was made during the author’s PhD. studies, under the responsibility of Guido Perboli and Roberto Tadei from Politecnico di Torino. The author would like to thank Professor Teodor Gabriel Crainic, From UQAM and CIRRELT, Montréal, Canada, for his support and advice as well as the Special Issue invited editors and the two anonymous referees for their valuable suggestions to improve this paper.

References


