The Two-Echelon Capacitated Vehicle Routing Problem: Models and Math-Based Heuristics

Guido Perboli
Roberto Tadei
Daniele Vigo

December 2008

CIRRELT-2008-55

The Two-Echelon Capacitated Vehicle Routing Problem: Models and Math-Based Heuristics

Guido Perboli[1,2,*], Roberto Tadei[2], Daniele Vigo[3]

[1] Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation (CIRRELT)
[2] DAUIN, Politecnico di Torino, C.so Duca degli Abruzzi 24, I-10129, Torino, Italy
[3] DEIS, University of Bologna, Viale Risorgimento, 2, 40136 Bologna, Italy

Abstract. Multi-echelon distribution systems are quite common in supply-chain and logistic systems. They are used by public administrations in their transportation and traffic planning strategies as well as by companies to model their distribution systems. In the literature, most of the studies address issues related to the movement of flows throughout the system from their origins to their final destinations. A recent trend is to also focus on the management of the vehicle fleets required to provide transportation among the different echelons of the system. The aim of this paper is twofold. First, it introduces the family of Multi-Echelon Vehicle Routing Problems, a term which broadly covers such settings, where the delivery from one or more depots to customers is managed by routing and consolidating freight through intermediate depots. Second, it considers in detail the basic version of Multi-Echelon Vehicle Routing Problems, the Two-Echelon Capacitated Vehicle Routing Problem, which is an extension of the classical VRP where the delivery is compulsory delivered through intermediate depots, named satellites. A mathematical model for Two-Echelon Capacitated Vehicle Routing Problem, some valid inequalities and two math-heuristics based on the model are presented. Computational results up to 50 customers and 4 satellites show the effectiveness of the developed methods.

Keywords. Vehicle routing, multi-echelon systems, City Logistics.

Acknowledgements. The authors are grateful to Jesús Gonzáles Feliu for his contribution to a previous version of the paper. This project has been partially supported by the Ministero dell'Università e della Ricerca (MUR) (Italian Ministry of University and Research), under the Progetto di Ricerca di Interesse Nazionale (PRIN) 2007 "Optimization of Distribution Logistics".

---

* Corresponding author: Guido.Perboli@cirrelt.ca

Dépôt légal – Bibliothèque et Archives nationales du Québec,
Bibliothèque et Archives Canada, 2008

1 Introduction

The freight transportation industry is a major source of employment and supports the economic development of the country. However, freight transportation is also a disturbing activity, due to congestion and environmental nuisances, which negatively affects the quality of life, in particular in urban areas.

In freight transportation there are two main distribution strategies: direct shipping and multi-echelon distribution. In direct shipping, vehicles starting from a depot, transport their freight directly to the customers, while in the multi-echelon systems, freight is delivered from the origin to the customers through intermediate depots. Growth in the volume of freight traffic as well as the need to take into account factors such as the environmental impact and traffic congestion has led research in recent years to focus on multi–echelon distribution systems, and, in particular, two-echelon systems (Crainic et al., 2004). In two-echelon distribution systems, freight is delivered to an intermediate depot and, from this depot to the customers.

Multi-echelon systems presented in the literature are related to the movement of flows throughout the system from their origins to their final destinations and usually explicitly consider only the routing problem at the last level of the transportation system. While this relaxation may be acceptable if the dispatching at higher levels is managed with a truckload policy (TL), the routing costs of the higher levels are often underestimated and decision-makers can not directly use the solutions obtained from the models in the case of the less-than-truckload (LTL) policy (Ricciardi et al., 2002; Daskin et al., 2002; Shen et al., 2003; Verrijdt and de Kok, 1995).

Moreover, in the past decade multi-echelon systems with LTL dispatching policies have been introduced by practitioners in different areas as express delivery service companies (http://www.tntlogistics.com), grocery and hypermarkets product distribution , e-commerce and home delivery services (http://www.sears.com), Newspaper and press distribution (Jacobsen and Madsen, 1980), and city logistics (Crainic et al., 2004)

The main contribution of this paper is to introduce the *Multi-Echelon Vehicle Routing Problem*, a new family of routing problems where routing and freight management are explicitly considered at the different levels. The basic type of Multi-Echelon Vehicle Routing Problems, the *Two-Echelon Capacitated Vehicle Routing Problem* (2E-CVRP) is introduced and examined in detail. In 2E-CVRP, the freight delivery from the depot to the customers is managed by shipping the freight through intermediate depots. Thus, the transportation network is decomposed into two levels: the 1st level connecting the depot to the intermediate depots and the 2nd one connecting the intermediate depots to the customers. The objective is to minimize the total transportation cost of the vehicles involved in both levels. Constraints on the maximum capacity of the vehicles and the intermediate depots are considered, while the timing of the deliveries is ignored.

A flow-based model for the 2E-CVRP is introduced, as well as valid inequalities used to strengthen the continuous lower bound. Moreover, the same model is used to derive two fast math-heuristics.

The paper is organized as follows. In Section 2 we recall the literature related to Multi-Echelon Vehicle Routing Problems. In Section 3 we give a general description of Multi-Echelon Vehicle Routing Problems. Section 4 is devoted to introduce 2E-CVRP and give a mathematical model, which is strengthened by means of valid inequalities in Section 5, while Section 6 presents the two heuristics using different simplified variants of the base model to quickly find feasible solutions for the 2E-CVRP. Finally test instances for 2E-CVRP are introduced and some computational results are discussed in Section 7.

2 Literature review

Freight distribution and vehicle routing are playing, in the past decade, a central role not only in the supply chain and production planning, but also for their leading role in several environmental and politic aspects. Moreover, several transportation and production systems have been moved from a single-level to a multi-echelon distribution schema. As stated in the introduction, this paper focuses on the extension to multi-echelon systems of vehicle routing problems, which have been poorly studied form the routing point of view up to now. For this reason, in the following we present the literature review along two directions. First, the literature on multi-echelon systems is discussed. Second, a quick review of the large field of Vehicle Routing Problems is presented.

In the literature, the multi-echelon systems, and the two-echelon systems in particular, refer mainly to supply chain and inventory problems (Ricciardi et al., 2002; Daskin et al., 2002; Shen et al., 2003; Verrijdt and de Kok, 1995). These problems do not use an explicit routing approach for the different levels, but focus more on the production and supply chain management issues. In location problems, some studies deal with the location of intermediary facilities for a multi-echelon distribution systems (Ricciardi et al., 2002; Crainic et al., 2004).

The first application of a two-echelon distribution system with an explicit minimization of the total transportation costs can be found in (Jacobsen and Madsen, 1980). In this study, a comparison of several fast heuristics for solving a two-echelon Location Routing Problem in which transfer points are not known in advance is presented. The distribution system and input data are based on a real case, in which two newspaper editors combined their printing and transportation facilities to decrease transportation costs.

A more recent real application of two-tier distribution networks is due to Crainic et al. and is related to the freight distribution in a large urban area (Crainic et al., 2004). The authors developed a two-tier freight distribution system for congested urban areas, using small intermediate platforms, called satellites, as intermediate points for the freight distribution. This system is developed for a specific case study and a generalization of such a system has not been formulated.

Multi-echelon systems presented in the literature usually explicitly consider the routing problem at the last level of the transportation system, while at higher levels a simplified routing problem is considered. While this relaxation may be acceptable if the dispatching at higher levels is managed with a truckload policy (TL), the routing costs of the higher levels are often underestimated and decision-makers can not directly use the solutions obtained from the models in the case of the less-than-truckload (LTL) policy (Ricciardi et al., 2002; Daskin et al., 2002; Shen et al., 2003; Verrijdt and de Kok, 1995).

In the case where a less-than-truckload policy with vehicle trips serving several customers is applied only at the second level, the problem is close to a multi-depot VRP. However, since the most critical decisions are related to which satellites will be used and in assigning each customer to a satellite, more pertinent methods will be found in multi-depot Location Routing Problems (LRP). In these problems, the location of the distribution centers and the routing problem are not solved as two separate problems but are considered as a more complex problem (for a more detailed survey of LRP, see Nagy and Salhi, 2007; Albareda-Sambola et al., 2005; Prins et al., 2007). Although many of these studies refer to direct shipping strategies (i.e., single echelon), some heuristics have been developed for some specific multi-echelon problems (Jacobsen and Madsen, 1980), even if no extension of a general multi.echelon routing scheme has been developed.

Vehicle Routing has become a central problem in the fields of logistics and freight transportation. In some market sectors, transportation costs constitute a high percentage of the value added of goods. Therefore, the use of computerized methods for transportation can result in savings ranging from 5% to as much as 20% of the total costs (Toth and Vigo, 2002). Unfortunately, to our knowledge, only the single-level version of the Vehicle Routing Problem has been studied.

The main contributions in the area are presented below.

The case modeled by the VRP, also known as Capacitated VRP (CVRP), considers a fleet of identical vehicles. The objective is the minimization of the transportation costs under the constraint of the maximum shipping capacity of each vehicle. When an additional constraint on the maximum distance that each vehicle can cover is combined, 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 variant of VRP is the most commonly studied, and recent studies have developed good heuristic methods. Exact algorithms can solve relatively small instances and their computational effort is highly variable (Cordeau et al., 2005). For this reason, exact methods are mainly used to determine optimal solutions of the test instances, while heuristic methods are used in practical applications.

Cordeau et al. proposed a Tabu Search algorithm, called Unified Tabu Search Algorithm (UTSA) (Cordeau et al., 2001), to different versions of VRP problems, including CVRP. It tolerates intermediate unfeasible solutions through the use of a generalized objective function containing self-adjusting coefficients. This feature permits a decrease in the average deviation from the best known solution without any further computational effort. The Granular Tabu Search (GTS) by Toth and Vigo is based on the idea that removing the nodes unlikely to appear in an optimal solution could considerably reduce the neighborhood size and thus the computational time (Toth and Vigo, 2003).

These results have been recently improved by different approaches based on

Hybrid and Evolutionary Algorithms (Perboli et al., 2008; Prins, 2004; Mester and Bräysy, 2005). For a detailed survey of the exact and heuristic methods see (Cordeau et al., 2007; Toth and Vigo, 2002; Cordeau et al., 2005).

In real world applications, the problem is often different and many variants of VRP have been developed. The most well known variants are VRP with time windows (VRP-TW), multi-depot VRP (MDVRP) and VRP with pick-ups and deliveries (VRP-PD) (for a survey, see(Cordeau et al., 2007; Toth and Vigo, 2002)). We note only one variant of VRP where satellites facilities are explicitly considered, the VRP with Satellites facilities (VRPSF). In this variant, the network includes facilities that are used to replenish vehicles during a route. When possible, satellite replenishment allows the drivers to continue the deliveries without necessarily returning to the central depot. This situation arises primarily in the distribution of fuels and some other retail applications; the satellites are not used as depots to reduce the transportation costs (Crevier et al., 2007; Angelelli and Speranza, 2002; Bard et al., 1998).

3 The Multi-Echelon Vehicle Routing Problems

Freight consolidation from different shippers and carriers associated with some kind of coordination of operations is among the most important ways to achieve a rationalization of the distribution activities. Intelligent Transportation Systems technologies and operations research-based methodologies enable the optimization of the design, planning, management, and operation of City Logistics systems (Crainic and Gendreau, forthcoming; Taniguchi et al., 2001).

Consolidation activities take place at so-called *Distribution Centers* (DCs). When such DCs are smaller than a depot and the freight can be stored for only a short time, they are also called *satellite platforms*, or simply *satellites*. Long-haul transportation vehicles dock at a satellite to unload their cargo. Freight is then consolidated in smaller vehicles, which deliver them to their final destinations. Clearly, a similar system can be defined to address the reverse flows, i.e., from origins within an area to destinations outside it.

As stated in the introduction, in the Multi-Echelon Vehicle Routing Problems the delivery from the depot to the customers is managed by rerouting and consolidating the freight through different intermediate satellites. The general goal of the process is to ensure an efficient and low-cost operation of the system, while the demand is delivered on time and the total cost of the traffic on the overall transportation network is minimized. Usually, capacity constraints on the vehicles and the satellites are considered.

More precisely, in the Multi-Echelon Vehicle Routing Problems the overall transportation network can be decomposed into $k \geq 2$ levels:

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

In real applications two main strategies for vehicle assignment at each level can be considered. Given a level, the corresponding vehicles can be associated with a common parking depot, from where they are assigned to each satellite depending on the satellite demand. If the number of vehicles is not known in advance, a cost for each available vehicle is considered; this usually depends on the traveling costs from the parking depot to the satellites. Another strategy consists in associating with each satellite a number of vehicles which start and end their routes at the considered satellite. In our study we will consider the first strategy, considering similar costs for the assignment of each vehicle to a satellite. Thus, each transportation level has its own fleet to perform the delivery of goods and the vehicles assigned to a level can not be reassigned to another one.

The most common version of Multi-Echelon Vehicle Routing Problem arising in practice is the *Two-Echelon Vehicle Routing Problem*. From a physical point of view, a Two-Echelon Capacitated Vehicle Routing system operates as follows:

- freight arrives at an external zone, the depot, where it is consolidated into the 1st-level vehicles, unless it is already carried in a fully-loaded 1st-level truck;
- Each 1st-level vehicle travels to a subset of satellites and then it will return to the depot;
- At a satellite, freight is transferred from 1st-level vehicles to 2nd-level vehicles;
- Each 2nd-level vehicle performs a route to serve the designated customers, and then travels to a satellite for its next cycle of operations. The 2nd-level vehicles return to their departure satellite.

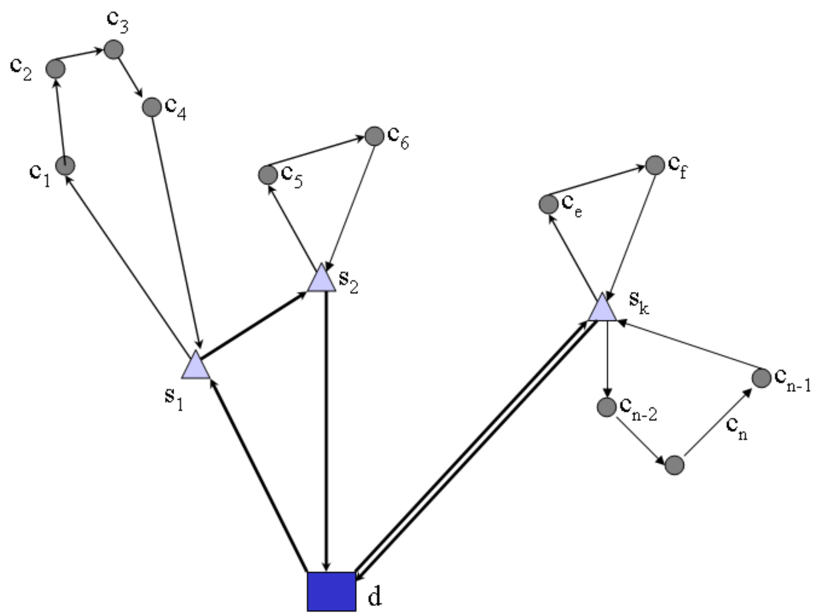

Figure 1: Example of 2E-CVRP transportation network

In the following, we will focus on Two-Echelon Vehicle Routing Problems, using them to illustrate the various types of constraints that are commonly defined on Multi-Echelon Vehicle Routing Problems. We can define three groups of variants:

Basic variants with no time dependence:

- Two-Echelon Capacitated Vehicle Routing Problem (2E-CVRP). This is the simplest version of Multi-Echelon Vehicle Routing Problems. At each level, all vehicles belonging to that level have the same fixed capacity. The size of the fleet of each level is fixed, while the number of vehicles assigned to each satellite is not known in advance. 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 vehicle's capacity and can not be split in multiple routes of the same level. For the 1st level we can consider two complementary distribution strategies. In the first case, each satellite is served by just one 1st-level vehicle and the aggregated demand passing through the satellite can not be split into different 1st-level vehicles. This strategy is similar to the classical VRP, and the capacity of 1st-level vehicles has to be greater than the demand of each satellite. In the second case, a satellite can be served by more than one 1st-level vehicle. This strategy has some analogies with the VRP with split deliveries and allow 1st-level vehicles with capacity which is lower than each satellite demand. If also the satellites are capacitated, constraints on the maximum number of 2nd-level vehicles assigned to each satellite are imposed. No information on loading/unloading operations is incorporated.

Basic variants with time dependence:

- Two-Echelon VRP with Time Windows (2E-VRP-TW). 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 can not be violated, while in the second, if they are violated a penalty cost is paid.

- 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 hard 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, the demand is lost and 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.

Other 2E-CVRP variants are:

- Multi-depot problem. In this problem the satellites are served by more than one depot. A constraint forcing to serve each customer by only one 2nd-level vehicle can be considered. In this case, we have a Multi-Depot Single-Delivery Problem.
- 2E-CVRP with Pickup and Deliveries (2E-VRP-PD). In this case we can consider the satellites as intermediate depots to store both the freight that has been picked-up from or must be delivered to the customers.
- 2E-CVRP with Taxi Services (2E-VRP-TS). In this variant, direct shipping from the depot to the customers is allowed if it helps to decrease the

cost, or to satisfy time and/or synchronization constraints.

4 The Two-Echelon Capacitated Vehicle Routing Problem

As stated in Section 3, 2E-CVRP is the two-echelon extension of the well known VRP problem. In this section we describe in detail the 2E-CVRP and introduce a mathematical formulation able to solve small and medium-sized instances. We do not consider any time windows or satellite synchronization constraints.

Let us denote the depot by v_0, the set of intermediate depots called satellites by V_s and the set of customers by V_c. Let n_s be the number of satellites and n_c the number of customers. The depot is the starting point of the freight and the satellites are capacitated. The customers are the destinations of the freight and each customer i has associated a demand d_i, i.e. the quantity of freight that has to be delivered to that customer. The demand of each customer can not be split among different vehicles at the 2nd level. For the first level, we consider that each satellite can be served by more than one 1st-level vehicle, so the aggregated freight assigned to each satellite can be split into two or more vehicles. Each 1st level vehicle can deliver the freight of one or more customers, as well as serve more than one satellite in the same route.

The distribution of the freight can not be managed by direct shipping from the depot to the customers. Instead the freight must be consolidated from the depot to a satellite and then delivered from the satellite to the desired customer. This implicitly defines a two-echelon transportation system: the 1st level interconnecting the depot to the satellites and the 2nd one the satellites to the customers (see Figure 1).

Define the arc (i, j) as the direct route connecting node i to node j. If both nodes are satellites or one is the depot and the other is a satellite, we define the arc as belonging to the 1st-level network, while if both nodes are customers or one is a satellite and the other is a customer, the arc belongs to the 2nd-level network.

We consider only one type of freight, i.e. the volumes of freight belonging to different customers can be stored together and loaded in the same vehicle for both the 1st and the 2nd-level vehicles. Moreover, the vehicles belonging to the same level have the same capacity. The satellites are capacitated and each satellite is supposed to have its own capacity, usually expressed in terms of maximum number of 2nd-level routes starting from the satellite or freight volume. Each satellite receives its freight from one or more 1st level vehicles.

We define as *1st-level route* a route made by a 1st-level vehicle which starts from the depot, serves one or more satellites and ends at the depot. A *2nd-level route* is a route made by a 2nd-level vehicle which starts from a satellite, serves one or more customers and ends at the same satellite.

The problem is easily seen to be NP-Hard via a reduction to VRP, which is a special case of 2E-CVRP arising when just one satellite is considered.

4.1 A Flow-based Model for 2E-CVRP

According to the definition of 2E-CVRP, if the assignments between customers and satellites are determined, the problem reduces to $1 + n_s$ VRP (1 for the 1st-level and n_s for the 2nd-level).

The main question when modeling 2E-CVRP is how to connect the two levels and manage the dependence of the 2nd-level from the 1st one.

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

In our model we will not consider the fixed costs of the vehicles, since we suppose they are available in fixed number. We consider the travel costs c_{ij}, which are of two types:

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

Another cost that can be used is the cost of loading and unloading operations at the satellites. Supposing that the number of workers in each satellite v_{s_k} is fixed, we consider only the cost incurred by the management of the freight and we define S_k as the unit cost of freight handling at the satellite v_{s_k}.

The formulation we present derives from the multi-commodity network design and uses the flow of the freight on each arc as main decision variables.

We define five sets of variables, that can be divided in three groups:

- The first group represents the arc usage variables. We define two sets of such variables, one for each level. The variable x_{ij} is an integer variable of the 1st-level routing and is equal to the number of 1st-level vehicles using arc (i, j). The variable y_{ij}^k is a binary variable representing the 2nd-level routing. It is equal to 1 if a 2nd-level vehicle makes a route starting from satellite k and goes from node i to node j, 0 otherwise.
- The second group of variables represents the assignment of each customer to one satellite and are used to link the two transportation levels. More precisely, we define z_{kj} as a binary variable that is equal to 1 if the freight to be delivered to customer j is consolidated in satellite k and 0 otherwise.
- The third group of variables, split into two subsets, one for each level, represents the freight flow passing through each arc. We define the freight flow as a variable Q^1_{ij} for the 1st-level and Q^2_{ijk} for the 2nd level, where k represents the satellite where the freight is passing through. Both variables are continuous.

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

$$D_k = \sum_{j \in V_c} d_j z_{kj}, \forall k \in V_s, \quad (1)$$

which represents the freight passing through each satellite k.

The model to minimize the total cost of the system may be formulated as

$V_0 = \{v_0\}$	Depot
$V_s = \{v_{s_1}, v_{s_2}, ..., v_{s_{n_s}}\}$	Set of satellites
$V_c = \{v_{c_1}, v_{c_2}, ..., v_{c_{n_c}}\}$	Set of customers
n_s	number of satellites
n_c	number of customers
m_1	number of the 1st-level vehicles
m_2	number of the 2nd-level vehicles
m_{s_k}	maximum number of 2nd-level routes starting from satellite k
K^1	capacity of the vehicles for the 1st level
K^2	capacity of the vehicles for the 2nd level
d_i	demand required by customer i
c_{ij}	cost of the arc (i, j)
S_k	cost for loading/unloading operations of a unit
	of freight in satellite k
Q^1_{ij}	flow passing through the 1st-level arc (i, j)
Q^2_{ijk}	flow passing through the 2st-level arc (i, j) and coming from satellite k
x_{ij}	number of 1st-level vehicles using the 1st-level arc (i, j)
y^k_{ij}	boolean variable equal to 1 if the 2st-level arc (i, j) is used by
	the 2nd-level routing starting from satellite k
z_{kj}	variable set to 1 if the customer c_i is served by the satellite k

Table 1: Definitions and notations

follows:

$$\min \sum_{i,j \in V_0 \cup V_s, i \neq j} c_{ij} x_{ij} + \sum_{k \in V_s} \sum_{i,j \in V_s \cup V_c, i \neq j} c_{ij} y_{ij}^k + \sum_{k \in V_s} S_k D_k \tag{2}$$

Subject to

$$\sum_{i \in V_s} x_{0i} \leq m_1 \tag{3}$$

$$\sum_{j \in V_s \cup V_0, j \neq k} x_{jk} = \sum_{i \in V_s \cup V_0, i \neq k} x_{ki} \quad \forall k \in V_s \cup V_0 \tag{4}$$

$$\sum_{k \in V_s} \sum_{j \in V_c} y_{kj}^k \leq m_2 \tag{5}$$

$$\sum_{j \in V_c} y_{kj}^k \leq m_{s_k} \quad \forall k \in V_s \tag{6}$$

$$\sum_{j \in V_c} y_{kj}^k = \sum_{j \in V_c} y_{jk}^k \quad \forall k \in V_s \tag{7}$$

$$\sum_{i \in V_s \cup v_0, i \neq j} Q_{ij}^1 - \sum_{i \in V_s \cup v_0, i \neq j} Q_{ji}^1 = \begin{cases} D_j & j \text{ is not the depot} \\ \sum_{i \in V_c} -d_i & \text{otherwise} \end{cases} \quad \forall j \in V_s \cup V_0 \tag{8}$$

$$Q_{ij}^1 \leq K^1 x_{ij} \quad \forall i, j \in V_s \cup V_0, i \neq j \tag{9}$$

$$\sum_{i \in V_c \cup k, i \neq j} Q_{ijk}^2 - \sum_{i \in V_c \cup k, i \neq j} Q_{jik}^2 = \begin{cases} z_{kj} d_j & j \text{ is not a satellite} \\ -D_j & \text{otherwise} \end{cases} \quad \forall j \in V_c \cup V_s, \forall k \in V_s \tag{10}$$

$$Q_{ijk}^2 \leq K^2 y_{ij}^k \quad \forall i, j \in V_s \cup V_c, i \neq j, \forall k \in V_s \tag{11}$$

$$\sum_{i \in V_s} Q_{iv_0}^1 = 0 \tag{12}$$

$$\sum_{j \in V_c} Q_{jkk}^2 = 0 \quad \forall k \in V_s \tag{13}$$

$$y_{ij}^k \leq z_{kj} \quad \forall i \in V_s \cup V_c, \forall j \in V_c, \forall k \in V_s \tag{14}$$

$$y_{ji}^k \leq z_{kj} \quad \forall i \in V_s, \forall j \in V_c, \forall k \in V_s \tag{15}$$

$$\sum_{i \in V_s \cup V_c} y_{ij}^k = z_{kj} \quad \forall k \in V_s, \forall j \in V_c \tag{16}$$

$$\sum_{i \in V_s} y_{ji}^k = z_{kj} \quad \forall k \in V_s, \forall j \in V_c \tag{17}$$

$$\sum_{i \in V_s} z_{ij} = 1 \quad \forall j \in V_c \tag{18}$$

$$y_{kj}^k \leq \sum_{l \in V_s \cup V_0} x_{kl} \quad \forall k \in V_s, \forall j \in V_c \tag{19}$$

$$y_{ij}^k \in \{0,1\}, \quad \forall k \in V_s \cup V_0, \forall i, j \in V_c \tag{20}$$

$$z_{kj} \in \{0,1\}, \quad \forall k \in V_s \cup V_0, \forall j \in V_c \tag{21}$$

$$x_{kj} \in \mathbb{Z}^+, \quad \forall k, j \in V_s \cup V_0 \tag{22}$$

$$Q_{ij}^1 \geq 0, \forall i, j, \in V_s \cup V_0, \; Q_{ijk}^2 \geq 0, \; \forall i, j \in V_s \cup V_c, \forall k \in V_s. \tag{23}$$

The objective function minimizes the sum of the traveling and handling operations costs. Constraints (4) show, for $k = v_0$, that each 1st-level route begins and ends at the depot, while when k is a satellite, impose the balance of vehicles entering and leaving that satellite. The limit on the satellite capacity is satisfied by constraints (6). They limit the maximum number of 2nd-level routes starting from every satellite (notice that the constraints also limit at the same time the freight capacity of the satellites). Constraints (7) force each 2nd-level route to begin and end to one satellite and the balance of vehicles entering and leaving each customer. The number of the routes in each level must not exceed the number of vehicles for that level, as imposed by constraints (3) and (5).

Constraints (8) and (10) indicate that the flows balance on each node is equal to the demand of this node, except for the depot, where the exit flow is equal to the total demand of the customers, and for the satellites at the 2nd-level, where the flow is equal to the demand (unknown) assigned to the satellites. Moreover, constraints (8) and (10) forbid the presence of subtours not containing the depot or a satellite, respectively. In fact, each node receives an amount of flow equal to its demand, preventing the presence of subtours. Consider, for example, that a subtour is present between the nodes i, j and k at the 1st level. It is easy to check that, in such a case, does not exist any value for the variables Q^1_{ij}, Q^1_{jk} and Q^1_{ki} satisfying the constraints (8) and (10). The capacity constraints are formulated in (9) and (11), for the 1st-level and the 2nd-level, respectively. Constraints (12) and (13) do not allow residual flows in the routes, making the returning flow of each route to the depot (1st-level) and to each satellite (2nd-level) equal to 0.

Constraints (14) and (15) indicate that a customer j is served by a satellite k ($z_{kj} = 1$) only if it receives freight from that satellite ($y^k_{ij} = 1$). Constraint (18) assigns each customer to one and only one satellite, while constraints (16) and (17) indicate that there is only one 2nd-level route passing through each customer. At the same time, they impose the condition that a 2nd-level route departs from a satellite k to deliver freight to a customer if and only the customer's freight is assigned to the satellite itself. Constraints (19) allow a 2nd-level route to start from a satellite k only if a 1st-level route has served it.

5 Valid inequalities for 2E-CVRP

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

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

$$\sum_{i,j \in S_c} y_{ij}^k \leq |S_c| - 1, \forall S_c \subset V_c,\ 2 \leq |S_c| \leq |V_c| - 2 \tag{24}$$

These inequalities explicitly forbid the presence in the solution of subtours not containing the depot, already forbidden by Constraints (10).

These inequalities can be strengthened by considering that, given a subset of second-level edges y_{ij}^k belonging to the same satellite, the cardinality of the subset of customers appearing in (24) is equal to the sub of a given subset of variables z_{kj}. More precisely, the inequality (24) can be rewritten as follows:

$$\sum_{i,j \in S_c} y_{ij}^k \leq \sum_{j \in S_c} z_{kj}, \forall S_c \subset V_c,\ 2 \leq |S_c| \leq |V_c| - 2 \tag{25}$$

In the following, we will refer to inequalities (25) as *edge cuts*. The number of potential valid inequalities (24) and (25) is exponential, so we should need a separation algorithm to add them. As we will show in Section 7, in practice the inequalities involving sets S_c with cardinality more than 3 are unuseful and the separation algorithm can be substituted by a direct inspection of the constraints up to cardinality equal to 3.

The aim of flow cuts is to reduce the splitting of the values of the binary variables when the continuous relaxation is performed, strengthening the BigM constraints (11). The idea is to reduce the constant K^2 by considering that each customer reduces the flow by an amount equal to its demand d_i. Thus the following inequalities are valid:

$$\begin{cases} Q^2_{ijk} \leq (K^2 - d_i) y^k_{ij}, \forall i, j \in V_c \;\; \forall k \in V_s \\ Q^2_{ijk} - \sum_{l \in V_s} Q^2_{jlk} \leq (K^2 - d_i) y^k_{ij} \;\; \forall i, j \in V_c, \forall k \in V_s. \end{cases} \tag{26}$$

Constraints (26) are of the same order of magnitude of (11), so they can be directly introduced into the model.

6 Math-based Heuristics for 2E-CVRP

In this section we introduce heuristic for the 2E-CVRP based on the information that can be obtained by solving the linear relaxation of the model presented in the previous section. Algorithms of this type are often called math-heuristics. If we consider the model of 2E-CVRP presented in Section 4, we can notice that, given feasible values to the variables dealing with the customer-satellite assignment, the z_{kj} variables, the problem is simply partitioned in at most $k+1$ CVRP instances, one for the 1st-level and one for each satellite with at least a customer assigned. Thus, given the values of z_{kj}, the associated solution can be computed by means of any heuristic or exact method developed for the CVRP. Thus, our idea is to focus our search on z_{kj}, using the model (2)-(23) to guide the search process. According to these guidelines, we develop two math-based heuristic methods to find feasible solutions of the customer-satellites based of the usage of simplified versions of 2E-CVRP model.

The first heuristic considers a continuous model derived from (2)-(23). Given the optimal solution of the continuous model, the heuristic apply a diving procedure on z_{kj}. Differently to similar procedures, in our case we privilege the fixing of the variables to 0, letting the model to adapt the values if the remaining variables. Moreover, in order to recover possible infeasibilities due to the fixing, a restarting procedure is incorporated. More precisely, the procedure works as follows (see Algorithm 1 for the pseudocode):

- the set of compulsory forced variables $forcedVars$ is emptied;
- while an integer solution of z_{kj} variables is not found or a maximum number of trials is not reached proceed with the diving;
  - set to 1 the z_{kj} in $forcedVars$;
  - solve the continuous model (2)-(23);
  - if the solution is integer in the z_{kj} variables and the corresponding assignment of 2nd-level vehicles satisfy the capacity constraints on the satellites, solve the corresponding CVRP instances;
  - otherwise
    - get the $p < P$ z_{kj} variables with a value near to zero and with the largest residual cost and force them to zero;
    - if $p = 0$, get the q z_{kj} with a value greater or equal to 0.5 and force them to 1.
    - optimize the continuous model continuous model (2)-(23);
    - if the model is infeasible, take the last fixed variable and add to

$forcedVars$, unfix the other fixed variables, increase the number of trials and restart the process.

In the second heuristic method we consider that the number of variables z_{kj} in model (2)-(23) is quite small and a MIP solver can find a near-optimal solution with a limited computational effort of 2E-CVRP model with variables y_{ij}^k and x_{ij} considered as continuous. Thus, we consider a simplified version of model (2)-(23) where (20) and (22) are ignored. Moreover, we add to the simplified model the integer variables v_k, representing the vehicles used by satellite k, and the following constraints:

$$\sum_{j \in V_c} z_{k,j} d_j <= K^2 v_k \ \forall k \in V_s \tag{27}$$

$$\sum_{k \in V_s} v_k <= m_2 \tag{28}$$

$$\sum_{k \in V_s} v_k <= m_{s_k} \tag{29}$$

Constraints are (27)-(29) are used to ensure the satisfaction of the capacity constraints of the satellites even when the y_{ij}^k are not integral. In the following, we will refer to this simplified model as *semi-continuous 2E-CVRP* model.

Thus, the semi-continuous heuristic works as follows:

- solve the semi-continuous 2E-CVRP model by meas of a MIP solver with a time-limit of 60 seconds;
- put in a list the S best integer solutions found while solving the semi-continuous 2E-CVRP model;
- for every solution in the list, consider the assignments satellite-customer given by z_{kj} and solve the corresponding CVRP instances with a time

Algorithm 1 Diving-based Heuristic

$numTrials = 0$

$forcedVars = \{\oslash\}$

while $numTrials < MaxTrials$ **or** $solutionFound =$ **false do**

set the variables in $forcedVars$ to 1

Solve the continuous model

while one or more variables z_{kj} are not integer **do**

if All the z_{kj} have integral value **then**

Solve $m + 1$ CVRP instances

$solutionFound =$ **true**

else

Get the $p \leq P$ variables with value $z_{kj} \leq 0.1$ and having the largest residual cost

if $p \neq 0$ **then**

Set the p variables to 0

else

Get the $q \leq Q$ variables with value $z_{kj} \geq 0.5$

Set the q variables to 1

Solve the continuous model

if Model is infeasible **then**

$numTrials = numTrials + 1$

Get the last variable $z_{\overline{kj}}$ rounded

$forcedVars = forcedVars \cup \{z_{\overline{kj}}\}$

limit of 5 seconds.

The threshold S is used to explore more integer solutions of the semi-continuous 2E-CVRP model, ensuring at the same time an upper limit to the computational effort.

In both the heuristics, an exact or a heuristic method to solve the CVRP problems is needed. A comparison of the results obtained by means of both exact and heuristic methods for CVRP is given in Section 7.1.

7 Computational tests

In this section, we analyze the behavior of the model and the heuristics in term of solution quality and computational efficiency. Being 2E-CVRP introduced for the first time in this paper, in Subsection 7.1 we define some benchmark instances, extending the instance sets from the VRP literature. All the tests have been performed on a 3 GhZ Pentium PC wth 1 Gb of Ram. The models and the routines have been implemented in Mosel language and tested by means of XPress 2006 solver (Dash Associates, 2006). Section 7.2 is devoted to present the computational results on a wide set of benchmark instances and the impact of the valid inequalities of Section 5 on the computational results, while Section 7.3 presents the computational results of the model and the valid inequalities on the overall sets of instances.

7.1 Instance sets

In this section we introduce different instance sets for 2E-CVRP. The instances cover up to 51 nodes (1 depot and 50 customers) and are grouped in three sets. The first three sets have been built from the existing instances for VRP by Christofides and Eilon denoted as E-n13-k4, E-n22-k4, E-n33-k4 and E-n51-k5 (Christofides and Eilon, 1969). All the instance sets can be downloaded from the web site of OR-Library (Beasley, 1990).

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

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

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

A summary of the main features of the different sets are reported in Table 2. The first column reports the set of instances, while the number of instances in shown in Column 2. Columns 3 and 4 contain the number of satellites and customers, respectively. The number of vehicles for the 1st and the 2nd level can be read in Columns 5 and 6, while Columns 7 and 8 give the capacity of the vehicles of the two levels. In the remaining columns the rule used to localize the satellites and the customers are specified. More in detail, for the satellites the value *All pairs* indicates that all the possible pairs have been computed, while

Set	Instances	ns	nc	m1	m2	C1	C2	Satellite distribution	Customer distribution
1	66	2	12	3	4	15000	6000	Pair	Chistofides and Eilon E-n13-k4 instance
2	6	2	21	3	4	15000	6000	Random	Chistofides and Eilon E-n22-k4 instance
2	6	2	32	3	4	20000	8000	Random	Chistofides and Eilon E-n33-k4 instance
2	6	2	50	3	5	400	160	Random	Chistofides and Eilon E-n51-k5 instance
2	3	4	50	4	5	400	160	Random	Chistofides and Eilon E-n51-k5 instance
3	6	2	21	3	4	15000	6000	Border Random	Chistofides and Eilon E-n22-k4 instance
3	6	2	32	3	4	20000	8000	Border Random	Chistofides and Eilon E-n33-k4 instance
3	6	2	50	3	5	400	160	Border Random	Chistofides and Eilon E-n51-k5 instance

Table 2: Summary of the benchmark tests

Random that the satellites are randomly selected. About the customers, the name reported in the column is the name of the instance by Christofides and Eilon, 1969.

7.2 Valid inequalities computational results

In this section we present the computational results of the 2E-CVRP model on instances belonging to Set1 and Set2 using the valid inequalities introduced in Setion 5 within a computation time limit of 10000 seconds.

With respect to the edge cuts, a series of tests was carried out using a simple procedure testing all the subtours up to cardinality 5. The procedure, coded in Mosel, iteratively solves the continuous problem and checks the violated cuts up to 10 iterations. According to the results, the subtours of cardinality greater than 3 are ineffective for the quality of both lower bounds and final solution.

In table 3 the results of the 66 instances corresponding to the problem with 12 customers and 2 satellites are given. The optimum is reported in the second column, while columns 3 and 4 contain the time in seconds needed to solve the instances without and with the valid inequalities introduced in Section 5, while column 5 reports the number of edge cuts added. Finally, the last column presents the percentage of decreasing/increasing of computational time due to the usage of the valid inequalities. We do not present the lower bounds at the root node with and without cits, being the difference more than 2%. This behaviour, as we will show with the results of Set 2, is mainly due to the small size of the instances themselves.

According to the results most instances are solved in less than one minute, and only 10 of them need more than 2 minutes to be solved. There are however seven instances for which the computational times are greater than 10 minutes. This gap is mostly related to the satellite location. In fact, the greatest computational times are related to the situation where choosing which satellite to use has little or no effect on the final solution. In this situation, the model finds an optimal solution quickly, but spends a lot of time closing the nodes of the decision tree due to the poor quality of the lower bound obtained by the continuous relaxation of the model. Better behavior is obtained with the valid inequalities. As a counter effect, on some instances, the computational time still increases, but this is mainly due to the fact that the management of the additional inequalities. Moreover, the number of added cuts in quite limited, with a mean of 40 cuts added to the original formulation.

The results on Set 2 instances are presented in Table 4. Table 4 presents the behavior of the lower bound computed with a continuous relaxation of the model found without and with the valid inequalities. More precisely, columns 1 and 2

Satellites	OPT	Time		Cuts	% Time
		Without cuts	With cuts		
1,2	280	1312.34	1032.34	34	-21.34%
1,3	286	861.94	298.03	42	-65.42%
1,4	284	1445.05	306.11	56	-78.82%
1,5	218	2.06	1.98	51	-3.83%
1,6	218	7.92	2.27	46	-71.40%
1,7	230	19.95	5.95	72	-70.16%
1,8	224	2.50	2.75	60	10.00%
1,9	236	13.34	6.44	73	-51.75%
1,10	244	14.27	6.08	63	-57.40%
1,11	268	28.70	8.73	34	-69.57%
1,12	276	45.05	31.33	4	-30.45%
2,3	290	849.17	393.84	64	-53.62%
2,4	288	895.19	658.84	60	-26.40%
2,5	228	4.48	2.92	46	-34.83%
2,6	228	4.20	11.89	52	182.89%
2,7	238	7.09	3.53	80	-50.20%
2,8	234	6.00	4.17	50	-30.47%
2,9	246	11.69	7.08	50	-39.44%
2,10	254	26.25	8.67	113	-66.96%
2,11	276	37.27	9.70	50	-73.96%
2,12	286	226.48	156.67	8	-30.82%
3,4	312	1704.41	873.74	24	-48.74%
3,5	242	4.61	2.78	72	-39.66%
3,6	242	13.13	2.00	24	-84.76%
3,7	252	17.05	1.92	20	-88.73%
3,8	248	7.08	2.23	52	-68.44%
3,9	260	6.17	3.97	48	-35.69%
3,10	268	33.27	10.30	54	-69.05%
3,11	290	17.50	6.33	62	-63.84%
3,12	300	13.39	8.03	72	-40.02%
4,5	246	6.39	3.28	30	-48.66%
4,6	246	10.17	3.61	28	-64.51%
4,7	258	12.16	5.81	28	-52.19%

Satellites	OPT	Time		Cuts	% Time
		Without cuts	With cuts		
4,8	252	5.25	0.64	24	-87.79%
4,9	264	6.56	3.94	12	-40.01%
4,10	272	15.28	4.34	32	-71.57%
4,11	296	11.11	5.89	46	-46.97%
4,12	304	13.91	6.22	46	-55.28%
5,6	248	3.28	3.31	22	0.98%
5,7	254	1.97	2.56	16	30.18%
5,8	256	9.34	4.77	40	-48.99%
5,9	262	6.59	4.45	36	-32.47%
5,10	262	2.08	2.70	26	30.08%
5,11	262	1.73	1.70	12	-1.79%
5,12	262	1.41	1.67	12	18.92%
6,7	280	17.70	18.77	60	6.00%
6,8	274	7.64	6.50	8	-14.92%
6,9	280	15.22	12.27	26	-19.40%
6,10	280	7.73	8.38	8	8.29%
6,11	280	7.11	5.74	8	-19.33%
6,12	280	14.88	4.45	8	-70.06%
7,8	292	4.42	2.06	12	-53.35%
7,9	300	8.97	8.67	30	-3.30%
7,10	304	12.63	13.52	30	7.06%
7,11	310	23.88	8.03	16	-66.36%
7,12	310	19.94	10.16	16	-49.06%
8,9	326	40.81	18.25	80	-55.28%
8,10	326	17.86	15.33	70	-14.17%
8,11	326	11.55	5.94	8	-48.58%
8,12	326	6.84	5.44	8	-20.53%
9,10	338	24.27	17.47	54	-28.01%
9,11	350	17.52	17.58	78	0.35%
9,12	350	16.25	13.41	8	-17.50%
10,11	358	40.98	53.23	16	29.89%
10,12	358	23.19	22.59	16	-2.56%
11,12	400	40.45	37.95	31	-6.18%

Table 3: 12 customers and 2 satellites instances: valid inequalities improvements

contain, respectively, the number of customers in the original Christofides and Eilon's instances and the position of the satellites given as customer number. The values and the gap with the best solution of the first lower bound (calculated at the root node) without and with the valid inequalities are reported in columns 3-6, while the same data on the final lower bound (calculated at the end of the optimization process), increased by letting the solver to apply lift-and-project cuts during the optimization, and its gap are presented in columns 7-10. The number of cuts added at the root node in shown in column 11, while the best solution after 5000 seconds and 10000 seconds are reported in columns 12 and 13, respectively(bold values mean optimal values).

From these results it can be seen that the use of the cuts helps the model to reduce the gap by up to 9%. The behavior is confirmed by considering the values of the feasible solutions found by the model without and with the valid

		First Bound				Final Bound						
		Without cuts		With cuts		Without cuts		With cuts				
CVRP Instance	Satellites	Bound	Gap	Bound	Gap	Best Bound	Gap	Best Bound	Gap	Cuts	Sol 5000 ss	Best Sol.
E-n22-k4	7,18	399.20	4.28%	407.35	2.33%	417.07	0.00%	417.07	0.00%	189	417.07	**417.07**
	9,15	358.24	6.94%	368.06	4.39%	384.96	0.00%	384.96	0.00%	70	384.96	**384.96**
	10,20	423.48	10.01%	434.55	7.66%	457.07	2.88%	470.60	0.00%	134	470.60	**470.60**
	11,15	348.22	6.27%	357.85	3.67%	371.50	0.00%	371.50	0.00%	82	371.50	**371.50**
	12,13	374.11	12.43%	390.62	8.57%	417.61	2.25%	427.22	0.00%	156	427.22	**427.22**
	13,17	349.70	10.97%	363.18	7.54%	372.66	5.12%	392.78	0.00%	122	392.78	**392.78**
E-n33-k4	2,10	626.48	14.20%	696.45	4.62%	688.05	5.77%	725.50	0.64%	186	731.21	730.16
	3,14	610.21	14.61%	675.52	5.47%	656.75	8.10%	701.04	1.90%	148	714.63	714.63
	4,18	611.44	13.59%	651.90	7.87%	641.59	9.33%	683.42	3.42%	204	708.01	707.62
	5,6	636.93	19.10%	694.93	11.73%	711.73	9.60%	764.80	2.86%	106	787.29	787.29
	8,26	648.41	15.41%	716.30	6.55%	707.48	7.70%	739.24	3.56%	236	766.49	766.49
	15,23	662.62	14.96%	747.91	4.01%	741.57	4.83%	764.38	1.90%	102	779.19	779.19
E-n51-k5	3,18	536.23	10.51%	541.31	9.66%	528.80	11.75%	576.97	3.71%	432	599.20	599.20
	5,47	502.85	10.49%	508.45	9.50%	499.64	11.07%	513.09	8.67%	401	561.80	561.80
	7,13	505.31	14.89%	508.44	14.36%	496.98	16.29%	526.91	11.25%	340	593.71	593.71
	12,20	544.35	15.82%	550.39	14.89%	542.77	16.07%	550.99	14.79%	340	646.66	646.66
	28,48	499.29	7.23%	502.90	6.56%	489.82	8.99%	524.00	2.64%	384	538.22	538.22
	33,38	513.01	7.34%	516.76	6.66%	503.56	9.05%	540.14	2.44%	456	576.54	553.64
E-n51-k5	3,5,18,47	465.35	33.03%	496.83	28.50%	479.91	30.93%	502.82	27.63%	975	724.09	694.83
	7,13,33,38	462.99	19.03%	494.15	13.58%	480.45	15.97%	509.35	10.92%	1124	685.45	571.80
	12,20,28,48	476.98	34.13%	497.91	31.24%	482.01	33.43%	506.99	29.98%	966	915.43	724.09

Table 4: Results on instances of Set 2

inequalities. According to these results, for up to 32 customers the model is able to find good quality solutions in 5000 seconds at most. When the number of customers increases to 50, more than 5000 seconds are required to find a good solution. Moreover, the use of the cuts increases the average model quality in terms of the initial solutions and the lower bounds. The gaps between the best solutions and the best bounds are quite small for instances involving up to 32 customers, but increase for 50-customer instances, with a gap up to 29% for the 4 satellite instances.

7.3 Overall computational results

In this section we present the results of the tests in Set 2 and 3. All the results have been obtained using the model with the valid inequalities activated. Due to the computational experience on the edge cuts, we limited the generation of the cuts to cycles of length 3. The results are related to Set 2 and 3, being the sets with the largest size on terms of customers and satellites. The results of each set are summarized in Tables 5a and 5b. Each table contains the instance name and the number of satellites in Columns 1 and 2. Columns 3 and 4 contain the best solution and the lower bound computed by continuous relaxation of the model. Finally, the percentage gap of the best solution compared with the lower bound is presented in Column 5.

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

The instances generated from the classical CVRP instances present a distribution of the customers which is quite different from the distribution in realistic applications in urban and regional delivery. Moreover, the model is able to find solutions with an average gap less than 9% in Set 2 and around 14% in Set 3. This is quite large, but understandable considering that the lower bounds come from the simple continuous relaxation of the model with cuts.

Given the complexity of the model, and the number of integer variables and constraints involved in particular, it is not surprising that the solver requires more than 3 hours to obtain a reasonable solution. On the other hand, heuristic methods can help to close the gap with the lower bound with a limited computational effort. Tables 6a and 6b presents the results of the math-based heuristics derived from the complete 2E-CVRP model. Each table contains the instance name and the number of satellites in Columns 1 and 2. Column 3 report the best

Instance	Satellites	Final Solution	Best Bound	Gap
E-n22-k4-s6-17	2	**417.07**	**417.07**	0.00%
E-n22-k4-s8-14	2	**384.96**	**384.96**	0.00%
E-n22-k4-s9-19	2	**470.60**	**470.60**	0.00%
E-n22-k4-s10-14	2	**371.50**	**371.50**	0.00%
E-n22-k4-s11-12	2	**427.22**	**427.22**	0.00%
E-n22-k4-s12-16	2	**392.78**	**392.78**	0.00%
E-n33-k4-s1-9	2	730.16	725.50	0.64%
E-n33-k4-s2-13	2	714.63	701.04	1.94%
E-n33-k4-s3-17	2	707.62	683.42	3.54%
E-n33-k4-s4-5	2	787.29	764.80	2.94%
E-n33-k4-s7-25	2	766.49	739.24	3.69%
E-n33-k4-s14-22	2	779.19	764.38	1.94%
E-n51-k5-s2-17	2	599.20	576.97	3.85%
E-n51-k5-s4-46	2	561.80	513.09	9.49%
E-n51-k5-s6-12	2	593.71	526.91	12.68%
E-n51-k5-s11-19	2	646.66	550.99	17.36%
E-n51-k5-s27-47	2	538.22	524.00	2.71%
E-n51-k5-s32-37	2	553.64	540.14	2.50%
E-n51-k5-s2-4-17-46	4	694.83	502.82	38.19%
E-n51-k5-s6-12-32-37	4	571.80	509.35	12.26%
E-n51-k5-s11-19-27-47	4	724.09	506.99	42.82%
Mean				8.70%

(a) - Set 2

Instance	Satellites	Final Solution	Best Bound	Gap
E-n22-k4-s13-14	2	**526.10**	**526.10**	0.00%
E-n22-k4-s13-16	2	**521.04**	**521.04**	0.00%
E-n22-k4-s13-17	2	**496.34**	**496.34**	0.00%
E-n22-k4-s14-19	2	498.80	479.95	3.93%
E-n22-k4-s17-19	2	512.81	512.75	0.01%
E-n22-k4-s19-21	2	520.42	509.96	2.05%
E-n33-k4-s16-22	2	691.26	626.78	10.29%
E-n33-k4-s16-24	2	675.01	617.30	9.35%
E-n33-k4-s19-26	2	687.49	640.06	7.41%
E-n33-k4-s22-26	2	698.98	642.35	8.82%
E-n33-k4-s24-28	2	690.54	626.50	10.22%
E-n33-k4-s25-28	2	658.04	612.61	7.42%
E-n51-k5-s12-18	2	858.63	648.72	32.36%
E-n51-k5-s12-41	2	823.05	629.01	30.85%
E-n51-k5-s12-43	2	835.16	683.72	22.15%
E-n51-k5-s39-41	2	960.64	676.62	41.97%
E-n51-k5-s40-41	2	827.77	663.06	24.84%
E-n51-k5-s40-43	2	1001.89	702.84	42.55%
Mean				14.12%

(b) - Set 3

Table 5: Results of the MIP model on Set 2 and Set 3

solution obtained by the model. Columns 5, 6, 7 and 8 show the behavior of the diving and the semi-continuous heuristic, giving, for each heuristic, the value of the objective function and the computational time, while the best solution obtained combining the two heuristics is shown in column 9. Column 10 gives the value of the best lower bound known for each problem. Finally, columns 11 and 12 present the percentage gap of the best model solution and the best heuristic solution compared with the best lower bound, respectively. Both diving and semi-continuous heuristics have been tested solving the CVRP subproblems by means of the CVRP Branch and Cut algorithm by Ralphs et al., 2003 stopped after 5 seconds. In semi-continuous heuristic the parameter S, related to the maximum number of integer solutions of the semi-continuous model used by the heuristic, is set to 5.

According to the results, there is not a heuristic dominating the other. Moreover, the combination of diving and semi-continuous let us to reduce the mean gap from the lower bound. In particular this is true for Set 3, where the mean gap is reduced by 9%. This is more evident in 50 customers instance, where the mean gap is reduced from 32% of the MIP model to 10% of the heuristics. The benefits of the heuristics are also clear from the computational point of view, presenting a mean value of 36 seconds and worst case of 150 seconds in instance E-n51-k5-s40-43.

Obviously the results are affected by the method used to solve the CVRP subproblems. In fact the truncated Branch and Cut we used can have unreasonable computational times while the size of the CVRP instances increases. Thus, we substituted the Branch and Cut with the hybrid algorithm by Perboli et al., 2008 stopped after 3 seconds. We do not report the results, being exactly the same of the version, but with a mean computational effort of 24 seconds.

8 Conclusions

In this paper, we introduced a new family of VRP models, the Multi-Echelon VRP. In particular, we considered the 2-Echelon Capacitated VRP, giving a MIP formulation and valid inequalities for it. The model and the inequalities have been tested on new benchmarks derived from the CVRP instances of the literature, showing a good behavior of the model for small and medium sized instances.

Moreover, two different heuristics based on the MIP model have been presented. Both the heuristics present good performances both from the computa-

Instance	Satellites	Model	Diving	Time ss	SC	Time ss	Best Heur	Best LB	Gap Model	Gap Best Heur
E-n22-k4-s6-17	2	417.07	417.07	2.1	417.07	1.6	417.07	417.07	0.00%	0.00%
E-n22-k4-s8-14	2	384.96	441.41	2.2	408.14	0.9	408.14	384.96	0.00%	6.02%
E-n22-k4-s9-19	2	470.60	472.23	1.9	470.60	1.2	470.60	470.60	0.00%	0.00%
E-n22-k4-s10-14	2	371.50	435.92	2.6	440.85	0.0	435.92	371.50	0.00%	17.34%
E-n22-k4-s11-12	2	427.22	427.22	2.3	429.39	1.1	427.22	427.22	0.00%	0.00%
E-n22-k4-s12-16	2	392.78	425.65	2.3	425.65	1.0	425.65	392.78	0.00%	8.37%
E-n33-k4-s1-9	2	730.16	772.57	11.8	736.92	0.2	736.92	725.50	0.64%	1.57%
E-n33-k4-s2-13	2	714.63	749.94	15.8	736.37	0.7	736.37	701.04	1.94%	5.04%
E-n33-k4-s3-17	2	707.62	801.19	22.8	739.47	0.7	739.47	683.42	3.54%	8.20%
E-n33-k4-s4-5	2	787.29	838.31	7.0	816.59	1.5	816.59	764.80	2.94%	6.77%
E-n33-k4-s7-25	2	766.49	756.88	6.8	756.88	4.8	756.88	739.24	3.69%	2.39%
E-n33-k4-s14-22	2	779.19	779.06	3.0	779.06	0.5	779.06	764.38	1.94%	1.92%
E-n51-k5-s2-17	2	599.20	666.83	28.2	628.53	66.3	628.53	576.97	3.85%	8.94%
E-n51-k5-s4-46	2	561.80	543.24	29.4	543.20	30.3	543.20	513.09	9.49%	5.87%
E-n51-k5-s6-12	2	593.71	560.22	33.9	554.80	7.1	554.80	526.91	12.68%	5.29%
E-n51-k5-s11-19	2	646.66	584.09	76.3	592.06	28.8	584.09	550.99	17.36%	6.01%
E-n51-k5-s27-47	2	538.22	538.20	35.9	538.20	26.2	538.20	524.00	2.71%	2.71%
E-n51-k5-s32-37	2	553.64	584.59	34.4	587.12	64.7	584.59	540.14	2.50%	8.23%
E-n51-k5-s2-4-17-46	4	694.83	590.63	45.6	542.37	124.3	542.37	502.82	38.19%	7.86%
E-n51-k5-s6-12-32-37	4	571.80	571.80	38.2	584.88	109.2	571.80	509.35	12.26%	12.26%
E-n51-k5-s11-19-27-47	4	724.09	724.09	40.2	724.09	64.8	724.09	506.99	42.82%	42.82%
Mean									8.70%	8.42%

(a) - Set 2

Instance	Satellites	Model	Diving	Time ss	SC	Time ss	Best Heur	Best LB	Gap Model	Gap Best Heur
E-n22-k4-s13-14	2	526.10	561.15	1.9	526.54	0.6	526.54	526.10	0.00%	0.08%
E-n22-k4-s13-16	2	521.04	521.04	0.3	521.10	0.2	521.04	521.04	0.00%	0.00%
E-n22-k4-s13-17	2	496.34	496.34	1.7	496.39	0.5	496.34	496.34	0.00%	0.00%
E-n22-k4-s14-19	2	498.80	551.95	3.9	523.61	1.0	523.61	479.95	3.93%	9.10%
E-n22-k4-s17-19	2	512.81	512.81	9.9	521.84	5.4	512.81	512.75	0.01%	0.01%
E-n22-k4-s19-21	2	520.42	527.57	2.4	527.57	2.0	527.57	509.96	2.05%	3.45%
E-n33-k4-s16-22	2	691.26	674.71	6.8	672.19	15.6	672.19	626.78	10.29%	7.25%
E-n33-k4-s16-24	2	675.01	668.82	6.7	674.69	0.9	668.82	617.30	9.35%	8.35%
E-n33-k4-s19-26	2	687.49	744.42	7.0	680.38	5.2	680.38	640.06	7.41%	6.30%
E-n33-k4-s22-26	2	698.98	735.25	6.2	680.38	4.9	680.38	642.35	8.82%	5.92%
E-n33-k4-s24-28	2	690.54	702.86	7.5	692.66	6.6	692.66	626.50	10.22%	10.56%
E-n33-k4-s25-28	2	658.04	682.42	6.7	650.55	12.0	650.55	612.61	7.42%	6.19%
E-n51-k5-s12-18	2	858.63	719.64	25.6	692.54	46.1	692.54	648.72	32.36%	6.75%
E-n51-k5-s12-41	2	823.05	743.91	33.2	708.29	21.4	708.29	629.01	30.85%	12.60%
E-n51-k5-s12-43	2	835.16	711.73	15.8	715.76	38.3	711.73	683.72	22.15%	4.10%
E-n51-k5-s39-41	2	960.64	742.07	32.1	10078.90	115.7	742.07	676.62	41.97%	9.67%
E-n51-k5-s40-41	2	827.77	733.60	29.2	737.88	33.2	733.60	663.06	24.84%	10.64%
E-n51-k5-s40-43	2	1001.89	803.24	18.7	846.25	131.4	803.24	702.84	42.55%	14.29%
Mean									16.95%	7.68%

(b) - Set 3

Table 6: Results of the math-heuristics on Set 2 and Set 3

tional and the solution quality point of view.

9 Acknowledgments

The authors are grateful to Jesús Gonzáles Feliu for his contribution to a previous version of the paper.

This project has been partially supported by the Ministero dell'Università e della Ricerca (MUR) (Italian Ministry of University and Research), under the Progetto di Ricerca di Interesse Nazionale (PRIN) 2007 "Optimization of Distribution Logistics".

References

Albareda-Sambola, M., J. Diaz, E. Fernandez. 2005. compact model and tight bounds for a combined location-routing problem. *Computers and Operations Research*, 32 407–428.

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

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

Beasley, J. E. 1990. Or-library: distributing test problems by electronic mail. *Journal of the Operational Research Society*, 41 1069–1072.

Christofides, N., S. Eilon. 1969. An algorithm for the vehicle dispatching problem. *Operational Research Quarterly*, 20 309–318.

Cordeau, J.-F., M. Gendreau, A. Hertz, G. Laporte, J.-S. Sormany. 2005. New heuristics for the vehicle routing problem. In A. Langevin, D. Riopel, editors, *Logistics Systems: Design and Optimization*, pages 279–297. Springer, New York.

Cordeau, J.-F., M. Gendreau, A. Mercier. 2001. A unified tabu search heuristic for vehicle routing problems with time windows. *Journal of the Operational Research Society*, 52 928–936.

Cordeau, J.-F., G. Laporte, M. Savelsbergh, D. Vigo. 2007. Vehicle Routing. In C. Barnhart, G. Laporte, editors, *Transportation*, Handbooks on Operations Research and Management Science, pages 367–428. North Holland, Amsterdam.

Crainic, T. G., M. Gendreau. forthcoming. Intelligent freight transportation systems: Assessment and the contribution of operations research. *Transportation Research part C*.

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

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

Dash Associates. 2006. *XPRESS-MP User guide and reference manual (Release 2006).* Dash Associates, Northants.

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

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

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

Nagy, G., S. Salhi. 2007. Location outing: Issues, models and methods. *European Journal of Operational Research*, 177 649–672.

Perboli, G., F. Pezzella, R. Tadei. 2008. Eve-opt: an hybrid algorithm for the capability vehicle routing problem. *Mathematical Methods of Operations Research*. doi:10.1007/s00186-008-0236-7.

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

Prins, C., C. Prodhon, A. Ruiz, P. Soriano, R. Wolfler Calvo. 2007. Solving the capacitated location-routing problem by a cooperative lagrangean relaxation - granular tabu search heuristic. *Transportation Science*, 41 470–483.

Ralphs, T., L. Kopman, W. Pulleyblank, L. T. Jr. 2003. On the capacitated vehicle routing problem. *Mathematical Programming Series B*, 94 343–359.

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

Shen, Z.-J. M., C. Coullard, M. S. Daskin. 2003. A joint location-inventory model. *Transportation Science*, 37 40–55.

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

Toth, P., D. Vigo. 2002. An overview of vehicle routing problem. In P. Toth, D. Vigo, editors, *The Vehicle Routing Problem*, volume 9 of *SIAM Monographs on Discrete Mathematics and Applications*, pages 1–51. SIAM, Philadelphia.

Toth, P., D. Vigo. 2003. The granular tabu search and its application to the vehicle routing problem. *Journal on Computing*, 15(1) 333–346.

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