ROUTING DESIGN FOR LESS-THAN-TRUCKLOAD MOTOR CARRIERS USING ANT COLONY OPTIMIZATION

L. Barcos, V. Rodríguez, M. J. Álvarez and F. Robusté

Abstract

One of the most important challenges that confronts less-than-truckload carriers serving many-to-many distribution networks consists of determining how to consolidate flows of small shipments. The objective is to determine a route for each origin-destination pair that minimizes the cost while still guaranteeing a certain level of service. This research studies different aspects of the problem and provides a metaheuristic algorithm (based on Ant Colony Optimization techniques) capable of solving real-life problems in a reasonable computational time. The viability of the approach has been tested with a real case in Spain and encouraging results have been obtained.

Key words: Less-than-truckload operations optimization, freight transportation, Ant Colony Optimization, vehicle routing design.

1. Introduction

Many transport companies currently employ a wide variety of large-scale delivery systems, shipping goods across networks between many origin and destination pairs. This is the case of less-than-truckload (LTL) motor carriers and package carriers. A key to cost-effective shipping for these companies is to consolidate loads for different customers in order to travel with full or nearly full vehicles. In order to accomplish this, carriers establish networks consisting of regional consolidation centers, end-of-line (EOL) terminals, and break-bulk terminals (or hubs).

Normally, freight originating in a particular region is picked up on a daily basis during the afternoon by "small" trucks and is delivered to a regional EOL terminal (local transportation). Once the freight is in the origin EOL terminal, it is unloaded, sorted, consolidated and loaded onto a trailer for long-haul transportation. This transportation can be organized in a variety of different ways in order to move the freight up to the destination EOL terminal as economically as possible. When the load arrives at the destination EOL terminal it is unloaded, sorted and moved onto delivery trucks for final delivery. Long-haul movements occur mostly during the night so that the freight can be delivered in the morning. Figure 1 shows the flow of shipments through EOL terminals from the points of origin in a region to the final destination points.

Long-haul transportation may be carried out directly between the origin and destination EOL terminals. However, although this is the fastest method, in many cases it is not the most cost-effective since there is often insufficient freight to fill or even nearly fill a truck. Even for specific origin-destination requirements greater than one or more truckload(s), it is likely that the remaining load that does not fit exactly into a complete number of trucks, will need to be handled with LTL techniques. For this reason, other routing alternatives are used for consolidating freight, such as passing it through break-bulk terminals or using peddling/collecting routes.

Break-bulk terminals act as intermediate transshipment points where freight from many EOL terminals is unloaded, sorted, consolidated and reloaded onto other long-haul trailers which will take the freight to another break-bulk terminal or to the destination EOL terminal (see Roy, 2001 for a detailed description of the flow of shipments through a break-bulk terminal for a LTL carrier). Normally, these points act as terminals with high expectations of freight consolidation and play a very important role in this type of logistic network. Figure 2 shows how the freight can be consolidated if a break-bulk terminal is used for long-haul transportation. In many countries like Canada or Spain, break-bulk terminals perform as both break-bulk and EOL terminals. Sometimes, especially when shipments are small and the distance between the origin and the destination EOL terminals is long, shipments may pass through more than one break-bulk terminal.

The other strategy adopted by carriers for consolidating freight is to make multiple stops for collecting or delivering freight (collecting/peddling routes), as shown in Figure 3.

Cost-effective shipping, however, is not the only challenge for carriers since they have to ensure a certain level of service in terms of delivery time and frequency of service. Carriers working in Spain and Portugal usually offer a delivery time of 24 hours for most of their services (or 48 hours when this is not possible). Freight is, thus, normally collected in the afternoon and delivered the next morning.

Motor carriers face a complex logistical problem in managing their operations for long-haul transportation. It involves deciding if the freight for each origin-destination pair (origin EOL terminal – destination EOL terminal) should be transported directly, using a peddling/collecting route, or if consolidation should be carried out at one or several hubs. These decisions must also be taken bearing in mind the service conditions offered by the carrier. Which alternative is best? If the transportation has to be carried out using a peddling/collecting route, which stops should be made? If deliveries have to be made via break-bulk terminals, how many terminals should be used and which of them should be selected? Although each option has its own associated costs and delays, the choice does not depend solely on the locations of the origin and destination and the volume moved between the two points, but also on the demand throughout the entire network and on how the logistic costs of the system are configured. These decisions have network-wide impact and are complexly interconnected (Crainic and Roy, 1988); its optimization requires an integrated approach.

Furthermore, a characteristic feature of any freight transportation system is the need to move empty vehicles. This follows from the imbalances that exist in demand in certain regions of the area of activity: there are zones which generate more freight than that which is received, and vice versa. To correct these differences, empty vehicles must be sent from the areas in which an excess of empty vehicles has been created to those places which need them, in order to be able to perform the following day’s activities. As such, the design of routes must also take this aspect into account, and the most cost effective way of moving these vehicles must be sought.

This logistical problem, which acquires a prime role in the carrier’s planning process, is known as the *Load Planning Problem.* It consists of specifying how freight should be routed over the network, including the specification of movements of empty vehicles. The plan corresponds to the stable part of the demand for transportation, representing regular operations (Crainic, 2002). In order to adjust each day´s particular conditions, part of the load can be rerouted through other services and some departures can be added or cancelled.

The paper approaches this tactical problem by trying to find a set of routes for the long-haul transportation of loads that causes a minimal total system cost while ensuring a certain level of service. The methodology used for the research is based on Ant Colony Optimization techniques.

The following sections of this paper are structured as follows: Section 2 refers to several studies published in the literature related to the problem tackled here; Section 3 briefly describes what Ant Colony Optimization (ACO) consists of and some of its applications; Sections 4 and 5 include the detailed definition of the problem being tackled and its resolution by means of Ant Colony Optimization; Section 6 demonstrates how an ACO algorithm has been applied to a real situation, and Section 7 analyses the results obtained. Finally, some concluding remarks are presented in Section 8.

2. Routing design for LTL motor carriers: state of the art

The type of problem considered here is known as a *many-to-many* problem (i.e. one which involves shipping from several origins to several destinations in which each terminal acts simultaneously as origin and destination), unlike other related problems such as the Vehicle Routing Problem (VRP) or the Traveling Salesman Problem (TSP), which are considered to be *one-to-many* problems (Daganzo, 1996).

The classical mathematical programming approach for solving this type of problem exactly is not usually feasible for solving real-life problems. This fact is reflected in papers by, for example, Powell (1986), Powell and Sheffi (1983 and 1989) and Leung et al. (1990). Barcos (2002) makes a mixed-integer formulation of the problem, but its use for solving real life problems involves the use of a large number of variables and restrictions, and a prohibitive computation time. Consequently, heuristic algorithms which can provide solutions within a reasonable computation time are generally used.

There are studies which approach this type of problem with a high level of aggregation. Daganzo (1996) analyzes the problem with Continuous Approximations, working with the lowest level of detail possible in the data and trying to provide solutions in terms of design guidelines. In Hall (1987 and 1989) several shipment strategies via hubs are also analyzed (shipping through the hub closest to the origin or closest to the destination, shipping through the hub that offers the minimum travel distance, shipping first through the hub closest to origin and second through the hub closest to the destination, and other hybrid strategies). An attempt is made to identify different scenarios in which each of these strategies may be the most beneficial.

Powell (1986) and Powell and Sheffi (1983 and 1989) address the Load Planning Problem, which is defined as the specification of how freight should be routed (and consolidated) over the network, given a set of direct services between terminals. The authors implemented a heuristic procedure based on the hierarchical decomposition of the problem into a “master problem” and several subproblems. The “master problem” is a network design problem in which direct services offered by the carrier are established, with a minimum service frequency imposed. The total system cost is computed for each given configuration of selected services. Each time a modification is made in the network (by adding or dropping links), a routing subproblem and another empty balancing subproblem must be solved. To solve the first subproblem, shortest-path type procedures are used, while the second subproblem is solved using a minimum cost transshipment formulation with adjusted supply and demand.

In Leung et al. (1990) a problem-solving method also based on the decomposition of the problem into two inter-related subproblems is developed. The first subproblem considers the assignment of a first and a last break-bulk terminal on the route for every origin-destination pair. A pre-selection can be made within all the possible assignments, eliminating those assignments which violate the service-time restrictions (delivery within a certain time). The second subproblem seeks a minimum cost routing of the aggregated flow of goods among the break-bulk terminals (the routing problem, here, being restricted to the hub network, which is significantly smaller than the overall network). This is a procedure which iterates between both subproblems, where the routing subproblem constitutes an evaluation mechanism for a particular assignment. Lagrangian relaxation and shortest-path procedures have been used to approach these problems.

Crainic and Rousseau (1986) proposed a network optimization modeling framework for the medium-term planning problem of a multi-mode, multi-commodity freight transportation system. This general modeling framework was adapted for rail applications (Crainic, 1998) and also for LTL trucking (Roy and Delorme, 1989). In the latter application to LTL problems, the authors develop the model NETPLAN, which is intended to assist motor carriers in making decisions about the design of a service network, the routing of freight, and the balancing of empty vehicles. They use a non-linear mixed-integer programming formulation, where the main decision variables are the service frequencies and the volume of freight moving on each route through the network. NETPLAN trades off operating costs against speed and reliability of service at the objective function level, and provides the required service at minimum cost.

The present paper proposes a solution methodology based on Ant Colony Optimization. Until now there have been no known applications of this methodology for the resolution of this type of problem. However, encouraging results obtained from ACO when applied to other transport-related problems, such as the TSP or VRP, have motivated the present authors to test the feasibility of Ant Colony Optimization techniques in tackling routing design for LTL motor carriers.

3. Ant Colony Optimization

Ant Colony Optimization algorithms are models inspired by the behavior of ant colonies. Studies have been made to show how ants, which are almost blind, are capable of following the shortest route paths from their nests to feeding sources and back. This behavior is due to the ants' capacity for transmitting information between themselves, through a pheromone trail along the chosen path. In this way, while an isolated ant moves essentially at random, the "agents" of an ant colony detect the pheromone trail left by other ants and tend to follow that trail. These ants then deposit their own pheromone along the path, thus reinforcing it and making it more attractive. It can thus be said that the process is characterized by a positive feedback loop, in which the probability of an ant choosing a path increases with the number of ants which have previously used the same path.

The first algorithm based on Ant Colony Optimization (Ant System algorithm) was applied to the TSP (Dorigo et al, 1996), from which quite promising results were obtained. Improved versions of this algorithm have been developed which have been applied not only to the TSP but also to other Combinatorial Optimization problems such as the VRP and the Quadratic Assignment Problem (QAP). References include Dorigo et al. (1999), Gambardella et al. (1999), Gambardella and Dorigo (1997), Bullnheimer et al. (1999), Stützle and Dorigo (1999a), Stützle (1997) and Merkle et al. (2002). The present research paper develops a metaheuristic approach based on one of these improved algorithms (the Max Min Ant System, MMAS), which was first applied to the TSP (see Stützle and Hoos, 1997).

In ACO algorithms, artificial ants act as computational agents which transmit information in some way. The ACO algorithms are iterative processes. In each iteration, each one of the artificial ants which make up the colony constructs a solution to the problem. These agents construct the solutions in a probabilistic manner, being guided by an artificial pheromone trail and by data which has been heuristically calculated a priori; i.e., the virtual ants are not totally blind; instead, they are capable of including heuristic information in the construction of solutions. Therefore, when these algorithms are applied to the resolution of a problem, the pheromone trail which the ants will deposit must be determined, as must the heuristic information which will be worked with. The probabilistic rule followed by the ants to construct solutions must be defined taking these two elements into account.

When an ant has constructed a solution to the problem, this solution may be improved by applying a local search algorithm. For example, short runs of Tabu Search were used for the QAP and a *3-opt local search* for the TSP, as shown in Stützle and Dorigo (1999a and 1999b).

When all the ants of the colony have constructed a solution, the pheromone trails deposited by the virtual ants have to be updated, for which a pheromone updating rule needs to be defined. This updating rule takes into account the evaporation which the pheromone undergoes over time.

After the pheromone has been updated, the process is iteratively repeated until a termination condition is given (e.g. a maximum number of iterations or a given CPU run time). The solution provided by the algorithm will be the best solution found in the whole iterative process.

In general, all the ACO algorithms follow a specific algorithmic scheme:

Step 1: Set parameters and initialize pheromone trails.
Step 2: Construct a solution for each virtual ant.
Step 3: Improve each solution applying a local search.
Step 4: Update pheromone trails.
Step 5: If continuation is allowed, go to *Step 2*; otherwise stop.

4. Definition of the problem

The problem being tackled considers a network over which a carrier operates; it is composed of N EOL terminals which potentially act concurrently as origins and destinations of freight. Some of these EOL terminals operate as *break-bulk terminals* as well. Determining the number of break-bulk terminals in the network and their location constitutes an important strategic decision and has thus received a lot of attention (Ballou and Masters, 1993; Miller, 1993 and Rodriguez et al., 2007). In this paper, the number and location of the *break-bulk terminals* in the network are considered to be predetermined.

For each origin-destination pair (i, j) within the system, there is a load flow Q_{ij} (volume of freight per day) that must be transported regularly from the origin to the destination. The load must be routed and consolidated across the network, which involves choosing from among several possible routing strategies; therefore, for each origin-destination pair, there will be a variable which indicates the routing option chosen for that pair. The selection of these routing strategies should produce the least total system cost (objective function of the problem) while ensuring a certain level of service (restriction of the problem).

It is assumed that the LTL carrier guarantees the delivery of the freight within 24 hours, or within 48 hours at the latest. An unlimited fleet with homogeneous capacity C is considered. All freight to be transported between each origin-destination pair follows the same route, except in those cases where the load exceeds vehicle capacity C. In those cases, all vehicles which are full are sent directly to their destination (V_{ij}^F representing the number of full vehicles from origin *i* to destination *j*) while an attempt is made to consolidate the remaining load, q_{ij}. Thus, it can be said, without loss of generality, that $q_{ij} < C$. This is usual practice for less-than-truckload motor carriers (Leung et al, 1990) including the Spanish company being used in this study.

Searching for routes in order to consolidate loads q_{ij}, whilst trying to minimize the total cost of the system and respecting certain service requirements, constitutes a sizeable and complex problem, the principle aspects of which are explained below.

4.1. Variables of the problem and routing strategies

Given a load q_{ij}, determined by $q_{ij} = Q_{ij} - V_{ij}^F C$, it is necessary to choose a routing option for it. The variables of the model assign one of the following shipping alternatives to each origin-destination pair:

- Shipping directly from origin *i* to destination *j*.
- Shipping via the hub that, of the two closest to the origin, generates the least costly route: *Hub1(i, j)*.
- Shipping via the hub that, of the two closest to the destination, generates the least costly route: *Hub2(i, j)*.
- Shipping through two hubs: first through *Hub1(i, j)* and then through *Hub2(i, j)*.
- Shipping through the hub which generates the least costly route: *Hmin(i,j)*.
- Using a collecting or peddling route, with no predetermined limit on the number of stops. The possibility of mixing deliveries and collections in the same route is not considered since this would imply the need to reorganize the load within the vehicle, which in many cases would be costly and difficult (see Daganzo, 1996).

It should be noted that the possibilities of routing via one or several hubs have been restricted in order to reduce the space of solutions (as the number of hubs in the system increases, the number of routing possibilities for each origin-destination pair increases more than proportionally). These routing strategies are based on efficient design guidelines and appear in various papers on the subject, e.g. Daganzo (1996), Hall (1987), Hall (1989) and Robusté et al. (1996). In addition, the size of geographical areas covered by some real-life networks (as is the case in this study) and the necessity of guaranteeing the delivery of the freight within 24 hours, or within 48 hours at the latest, have dictated that only routes passing through no more than two hubs have been considered: three or more transshipments of freight would have taken up too much time and would have made it difficult to provide the demanded level of service.

A notable feature of this model is that peddling/collecting routes are explicitly dealt with. Other studies, such as those conducted by Powell and Sheffi (1983 and

1989), Powell (1986), Leung et al. (1990) and Roy and Delorme (1988) do not analyze these types of routes.

4.2. Objective function: the total system cost

A feasible solution for a given problem is constructed by the selection of a routing alternative for each origin-destination pair. This selection configures a scenario of freight flows (and, as such, vehicles) traveling along the links of the network, of freight that must be processed in the hubs and of empty vehicles traveling in order to balance the network. This scenario generates three types of costs:

- Costs associated with the distance covered by each vehicle. These are evaluated by considering a cost *R* per unit of distance traveled.
- Costs associated with stops made by vehicles. These costs are calculated assuming a fixed cost *P* per vehicle-stop made.
- Handling costs in hubs. Freight that is transshipped at a hub generates a handling cost which is calculated assuming a cost c_r per unit of freight handled at hub *r*. Unit cost c_r is subject to economies of scale or of technology. This endows the hubs with a hierarchy, such that, for the same routing distance it may be preferable to make transshipments in hubs where load processing is cheaper.

The model tries to find a solution which minimizes the sum of all these costs. With this cost structure, if a vehicle travels directly from the origin to the destination or the vehicle follows a peddling/collecting route, the cost caused by this vehicle can be calculated as: $Rd_{total} + n_s P$, where d_{total} is the total distance covered by the vehicle and n_s is the number of stops made to collect or deliver freight (including the stops made at the origin and at the destination points). The same formula can be used if the vehicle moves empty.

On the other hand, if the vehicle travels from/to a hub *r*, the cost involved can be computed as: $Rd_{total} + 2P$, where d_{total} is the total distance covered by the vehicle. In addition, a handling cost $c_r q_{ij}$ has to be added for any load q_{ij} which is in the vehicle and is transshipped at the hub involved (hub *r*).

4.3. Restriction in the level of service

The objective function must be minimized subject to the restriction of the level of service. This restriction demands the existence of a minimum percentage of freight (or dispatches) served in 24 hours, and the rest of the freight served within 48 hours. We define the transit time, *Time(i, j),* for each origin-destination pair (*i, j*), as the time this origin-destination pair is served. For a given solution, this transit time must be calculated for all the origin-destination pairs, taking into account the following aspects concerning delays:

- Driving time, which depends on the distance to be covered and on the speed estimated over a given link of the network.
- The stopping time at EOL terminals for collecting or delivering the freight. A fixed time θ_{sf} and another variable time θ_{sv} proportional to the load that is collected or delivered are assumed for any EOL terminal in the system.
- Reorganization time at hubs; i.e., the time needed to receive an inbound truck for transshipment, unload the freight, sort it and reload it in an outbound truck. Hub *r* is considered capable performing these operations and having an outbound vehicle ready to leave in time θ^{r}_{hub}. This time is measured from the

arrival of the last truck which contains freight that must travel in the outbound truck.

- Waiting time in the hubs for all the loads that must travel together to arrive at the hub.

Once the duration of the service (transit time) for each origin-destination pair has been computed, the percentage of freight served in 24 hours and in 48 hours can be calculated in order to evaluate if the given solution is feasible in terms of level of service. To this end, it is important to take into account that, normally, a vehicle can't leave the origin EOL terminal before a certain leaving time TL_i (e.g. 9 p.m.). Before this time, the carrier must collect shipments in local areas and deliver them to the origin terminal. There, documentation is prepared and the freight is unloaded and sorted according to its immediate destination (Roy, 2001). Furthermore, vehicles need to arrive at the destination EOL terminal before a time TA_j (e.g. 8 a.m.) so that the load can be delivered to the final destinations in the morning.

Thus, it can be said that an origin-destination pair (i, j) is served in 24 hours if the delivery is realized in a margin of time equal to $M_{ij}^{24} = TA_j - TL_i$. That is, the pair (i, j) will have a service of 24 hours if $Time(i, j) \le M_{ij}^{24}$. A margin of time, M_{ij}^{48}, for the service in 48 hours can be established in a similar way. In the formulation proposed here we assume that TL_i and TA_j are both the same for every terminal in the network. More detailed explanation concerning the calculation of the service times and the restriction in the level of service can be found in Barcos (2002).

Another feature that characterises the present work is the way in which the level of service has been considered. In this study, the restriction in the level of service is defined as a minimum percentage of load that must be delivered in 24 hours. The remaining load must be delivered in 48 hours. The optimization algorithm looks for solutions that fulfil this pre-established level of service and for each origin-destination pair determines which route to follow and the type of service (24/48 hours) that corresponds to it. However, in Roy and Delorme (1988) a standard service time is predetermined for each origin-destination pair, and the percentage of pairs satisfying the service targets is an output of the model (controlled through a penalty cost appearing in the objective function). In the model introduced by Powell and Sheffi (Powell and Sheffi, 1988) a minimum frequency is used for each load planning link as a heuristic approach for ensuring an acceptable level of service in most traffic lanes. No explicit constraints on origin-destination transit times are considered.

5. ACO algorithm

This section describes the principal aspects of the ACO algorithm, which has been implemented in C language for the purpose of resolving the problem presented in Section 4.

5.1. General structure of the algorithm

A general scheme of the structure of the algorithm is shown in Figure 4. Firstly, the information regarding the freight transportation system must be entered as inputs (see appendix 1 for a detailed list of the input data).

The next step in the process is to determine the number of full vehicles which will travel directly from origin to destination and those remaining loads that will work for consolidation, i.e., the calculation of V_{ij}^{F} and q_{ij} for every origin-destination pair (i, j) in the system.

In its next phase the algorithm tries to consolidate all the loads q_{ij}, choosing from among the routing alternatives described in section 4. Given the expectation of significant loading consolidation promised by the use of hubs and also for greater simplicity, it was decided to break down the general routing problem into two subproblems: the D-H subproblem, and the D-H-P/C subproblem. Both are solved using ant colony optimization techniques. The **D-H** subproblem consists of finding the optimal solution to the general problem while ignoring the possibility of making peddling/collecting routes (only **D**irect routes or routes via **H**ubs are considered, hence the name of the subproblem). The point of departure for the **D-H-P/C** subproblem is the solution found for the D-H subproblem (ψ^0). This second phase attempts to refine and improve the solution by introducing **P**eddling/**C**ollecting routes (hence the name of the subproblem).

When the solution is found for this second subproblem (ψ*), the algorithm solves the problem of empty balancing. Empty vehicle movements are determined using integer linear programming formulation of a classic transportation problem in which empty vehicles travel from the place where an "excess" of vehicles is created to where there is a "lack" of vehicles in order to satisfy the demand for the following period. The variables in this formulation, V_{ij}^E, are the number of empty vehicles travelling from terminal *i* to terminal *j*. The problem can be written as:

$$Min \sum_{i \in ST} \sum_{j \in DT} \left(2P + Rd_{ij}\right) V_{ij}^E \qquad (1)$$

Subject to

$$\sum_{j \in DT} V_{ij}^E = vs_i \qquad i \in ST \qquad (2)$$

$$\sum_{i \in ST} V_{ij}^E = vd_j \qquad j \in DT \qquad (3)$$

$$V_{ij}^{E} \geq 0 \text{ and integer} \qquad i \in ST, j \in DT \qquad (4)$$

Where ST and DT correspond, respectively, to the sets of terminals where an excess of vehicles (supply) and a lack of them (demand) is generated. The parameter vs_i is the number of vehicles supplied by terminal $i \in ST$, while vd_j is the number of vehicles demanded by terminal $j \in DT$.

Once the empty balancing problem has been solved, the final solution to the general problem can be obtained. This solution is composed of: the full vehicles sent directly at the beginning of the process (V_{ij}^{F}), all the consolidation routes included in ψ^* and the empty balancing movements, V_{ij}^{E} (see Figure 4).

5.2. Resolution of the D-H subproblem using Ant Colony Optimization techniques

The methodology used for the resolution of this subproblem is based on the Max Min Ant System (MMAS), described by Stützle and Dorigo (1999a, 1999b).

As described in section 3, ACO algorithms are iterative processes. In each iteration of the resolution process a colony of virtual ants constructs solutions to the given problem in a probabilistic manner. Each ant of the colony constructs its own individual solution, so there are as many solutions as ants in the colony. In the D-H subproblem, ants construct solutions choosing a routing strategy for each origin-destination pair within the system. This selection is chosen from the first five routing strategies described in section 4.1 (i.e. peddling and collecting routes are excluded). In this probability-based selection there are two elements that play a significant role: pheromone trails and the heuristic information parameter.

Once the ants have constructed a solution to the problem, they leave a pheromone trail on every element of it (i.e., they assign a pheromone trail to the chosen option for each origin-destination pair). The better the constructed solution is (i.e., the less costly it is), the stronger the trail will be, so that ants of later iterations are more likely to choose the route options that have provided good solutions in the past.

In addition, the ants don't only depend on the experience gained at the time the routes are constructed. They can also be guided by a heuristic information parameter calculated in advance. The idea behind this parameter is that ants are most likely to choose those route options that, given the specific structure of the problem and following a heuristic logic, present certain expectations of being a "good" selection.

Therefore, before beginning the iterative process, a preliminary stage must be carried out in which the possible routes to choose and the parameter of heuristic information are determined for each origin-destination pair. In this preliminary phase all the pheromone trails are also initialized. Following this preliminary stage, the iterative process (whose flow diagram can be seen in Figure 5) begins.

In each iteration, each ant of the colony constructs a solution to the problem. $\psi_1(t,k)$ refers to the solution constructed by ant *k* in iteration *t* of the process. The sub index 1 means that this solution corresponds to the first subproblem D-H. When all the ants of the colony have constructed a solution to the subproblem, the most cost-effective one of them is selected. This will be the *iteration-best solution* $\psi_1^{ib}(t)$. Nevertheless, it is worth mentioning that only solutions that fulfil the service level required (e.g. a minimum percentage of freight served in 24 hr) are considered in this selection. If, for a given iteration, none of the solutions constructed by the ants fulfils the service restrictions (which is not improbable at the early stages of the iteration process if the service level is very restrictive), then $\psi_1^{ib}(t)$ will be chosen by default. Shipping all the loads directly, for example, could be appropriate as a default solution when the service level is very restrictive. After computing $\psi_1^{ib}(t)$, the *global-best solution* found up to iteration *t*, $\psi_1^{gb}(t)$, must be updated. This means that if the cost of $\psi_1^{ib}(t)$ is less than the cost of $\psi_1^{gb}(t-1)$, then $\psi_1^{ib}(t)$ will be registered as the *global-best* solution found up to that moment, i.e. as $\psi_1^{gb}(t)$.

Prior to starting a new iteration in the process, the pheromone trails must be updated. Pheromone trails are subject to an evaporation (debilitation) effect, which enables the process to "forget" bad decisions made by the ants in previous iterations. Updating also involves the ants reinforcing some of the trails by depositing their own pheromone in order to make those trails more attractive in later iterations. The update process is described in detail in section 5.2.2.

The iterative process continues until a termination condition is given. The termination condition in this approach is a predefined number of iterations, NI_1. The final solution to this subproblem corresponds to the *global-best-solution* obtained after completion of these iterations, i.e. $\psi_1^{gb}(NI_1)$.

5.2.1. The probabilistic rule, the pheromone trails and the heuristic information parameter

We let *Z* be the set of five routing strategies considered for each origin-destination pair. Thus, $Z = \{\text{direct, Hmin, Hub 1, Hub 2, Hub 1\&2}\}$. The first option corresponds to a direct shipment; the second corresponds to the routing strategy associated with shipping through the hub that generates the least total cost, *Hmin*; the third and forth options correspond to shipments via one hub, *Hub1* and *Hub2* respectively (as described in section 4), and Hub1&2 indicates that the load is shipped via two hubs. Given an origin-destination pair (i, j), we define:

$\tau_{ij}^{z}(t)$ = *pheromone trail for the pair (i, j) and routing option z in iteration t*

η_{ij}^{z} = *heuristic information parameter for pair (i, j) and routing option z*

The probability with which an ant from the colony chooses the routing strategy *z* for the origin-destination pair (i, j) in iteration *t* is:

$$P_{ij}^{z}(t) = \frac{\left[\tau_{ij}^{z}(t)\right]^{\alpha_1}\left[\eta_{ij}^{z}\right]^{\beta_1}}{\sum_{z \in Z}\left[\tau_{ij}^{z}(t)\right]^{\alpha_1}\left[\eta_{ij}^{z}\right]^{\beta_1}} \qquad (5)$$

where α_1 and β_1 are two algorithm parameters which determine the relative influence of the pheromone trail and the heuristic information for the first subproblem D-H (hence the indicative “1”). This decision rule has been adapted from that used in Stützle and Dorigo (1999b) and in other ACO related articles. If $\alpha_1 = 0$, ants will choose a routing strategy exclusively guided by a stochastic heuristic rule (i.e., ants will not take into account their previous experience). If $\beta_1 = 0$, ants will choose a routing strategy, considering only the information given by the pheromone, which could lead, in some cases, to the rapid emergence of a stagnation situation. Therefore, a trade-off between the influence of the pheromone strength and that of the heuristic information parameter appears to be necessary.

The heuristic information parameter associated with each origin-destination pair (i, j) and for each routing option *z* is defined as:

$$\eta_{ij}^{z} = \frac{1}{Cost_{min}^{z}(q_{ij})} \qquad (6)$$

When z corresponds to the strategy of direct shipping, $Cost_{min}^{direct}(q_{ij})$ is calculated as the total cost of the direct route. If z corresponds to a routing strategy which passes through hubs (bearing in mind that four possibilities of shipping via hubs were considered), then $Cost_{min}^{z}(q_{ij})$ corresponds to the cost which can be proportionally imputed to q_{ij} when all the vehicles covering the route travel are full. In this sense, $Cost_{\min}^{z}(q_{ij})$ is the **minimum cost** which can be imputed to the load when it is transported according to the routing strategy z. For example, if the load q_{ij} is moved through the hub *Hub1(i,j)*, this cost is calculated as:

$$Cost_{min}^{Hub1}(i,j) = \frac{q_{ij}}{C}(4P + Rd_{i,Hub1} + Rd_{Hub1,j}) + q_{ij}c_{Hub1} \tag{7}$$

With this definition of the heuristic information parameter and using formula (5), there is a greater likelihood that ants will choose those routing strategies that have a lower corresponding minimum cost.

5.2.2. Pheromone trails update

After all the ants have constructed a solution, the following formulas, which have been adapted from those proposed by Stützle and Dorigo (1999b), are used to update the pheromone trails:

$$\tau_{ij}^{z}(t+1) = (1-\rho_1)\tau_{ij}^{z}(t) + \Delta\tau_{ij}^{best1}(t) \qquad \forall (i,j), \quad \forall z \tag{8}$$

$$\Delta\tau_{ij}^{best1}(t) = \begin{cases} \dfrac{1}{Cost\ (\psi_1^{best}(t))} & \text{If z is chosen for (i, j) in solution } \psi_1^{best}(t) \\ 0 & \text{Otherwise} \end{cases} \tag{9}$$

where $0 \le \rho_1 \le 1$ corresponds to the pheromone evaporation rate. To exploit the best solution found, only one ant is allowed to add pheromone after each iteration. This ant will be the one which provides the solution $\psi_1^{best}(t)$, where *best* can correspond to either the *iteration-best solution* $\psi_1^{ib}(t)$ or the *global-best solution* $\psi_1^{gb}(t)$. $\Delta\tau_{ij}^{best1}(t)$ is the amount of pheromone added by this ant and $Cost(\psi_1^{best}(t))$ means the cost of the solution $\psi_1^{best}(t)$. Note that the better the solution $\psi_1^{best}(t)$ is, the more pheromone will be deposited by the ant. Experience with the MMAS approach shows that it is expedient to use the *iteration-best solution* and the *global-best solution* alternately in the pheromone update; in general, best results are obtained by gradually increasing the frequency that the *global-best solution* is chosen for the trail update (Stützle and Dorigo, 1999b). By shifting the emphasis in this way from the *iteration-best* to the *global-best solution*, we progress from a vigorous exploration of the search space to an exploitation of the best solution found so far.

In order to avoid search stagnation[1], the MMAS allows the pheromone values to move only between a maximum and a minimum level of pheromone during the iteration process. We define $[\tau_{min1}, \tau_{max1}]$ as the permitted range of the pheromone trail values for the first subproblem. In the preliminary phase of the iterative process, all the pheromone trails are initialized with the maximum value, i.e. $\tau_{ij}^{z}(t=0)=\tau_{max1}$. If after the pheromone update of a given iteration *t* we have $\tau_{ij}^{z}(t)>\tau_{max1}$, we set $\tau_{ij}^{z}(t)=\tau_{max1}$; analogously, if $\tau_{ij}^{z}(t)<\tau_{min1}$, we set $\tau_{ij}^{z}(t)=\tau_{min1}$. In addition, when almost all the pheromone trails not associated with the *global- best solution* are very close to τ_{min1}, a process of re-initialization restores their values to τ_{max1} again. Premature search stagnation can be avoided by this re-initialization of the pheromone and the exploration of solutions improved.

5.3. Resolution of the D-H-P/C subproblem using Ant Colony Optimization techniques

The point of departure for the solution of the D-H-P/C subproblem is the solution found for the D-H subproblem, i.e. $\psi^0 = \psi_1^{gb}(NI_1)$. In this second subproblem, we try to improve the initial solution by introducing peddling/collecting routes. As such, there will be loads which, in the initial solution are transported directly or through hubs and which, in this second phase, will change to being transported by peddling/collecting routes.

The construction of peddling/collecting routes is complicated as there are an excessive number of possible combinations. To reduce the space of possible solutions, a *Set of Candidates for Peddling (* S_{ij}^P *)* and a *Set of Candidates for Collecting (* S_{ij}^C) are assigned to each *(i, j)* pair. Thus, if a collecting/peddling route such as that described in Figure 6(a)/6(b) has to be constructed, the algorithm should choose a load q_{lj} / q_{il} from those pertaining to the set S_{ij}^C / S_{ij}^P. These sets of candidates are determined on the expectation that the above peddling/collecting routes will improve the initial solution. The configuration of these sets is explained in section 5.3.1.

Once the sets of candidates have been determined, a new iterative ACO process begins. This process tries to modify the initial solution ψ^0, sending some of the loads that were initially transported directly or via hubs on peddling/collecting routes. The pseudo code for this second phase of the algorithm can be written as:

Determination of sets of candidates

---

[1] When stagnation occurs, the values of few pheromone trails go so high that the ants would always select, for a given origin-destination pair, the corresponding routing options again and again, making further solution improvements impossible.

Initialization of the second ACO process

FOR t=1 to NI_2 *DO*

FOR k=1 to NA_2 *DO*

Construct a D-H-P/C solution $\psi_2(t,k)$

END FOR

Selection of $\psi_2^{ib}(t)$

Local improvement of $\psi_2^{ib}(t)$ *obtaining* $\psi_2^{ib*}(t)$

Update $\psi_2^{gb}(t)$

Pheromone update

END FOR

NI_2 and NA_2 are, respectively, the number of iterations and the number of ants in the colony for this second ACO process (hence the indicative "2"). $\psi_2(t,k)$, $\psi_2^{ib}(t)$ and $\psi_2^{gb}(t)$ have the same meaning as $\psi_1(t,k)$, $\psi_1^{ib}(t)$ and $\psi_1^{gb}(t)$ for the first subproblem, respectively. Nevertheless, it should be noted that, in this second phase of the approach, a local improvement of the *iteration-best* solution is performed for each iteration, resulting in $\psi_2^{ib*}(t)$ (see section 5.3.4 for more details about the local improvement process).

5.3.1. Construction of the Sets of Candidates for Collecting/Peddling

Consider the case where we wish to construct the set S_{ij}^C. This set will be composed of those loads q_{lj} that are going to be considered as candidates when collecting routes that follow the sequence of stops *i-l-j* are constructed (Figure 6a). Given a load q_{lj}, this will belong to the set S_{ij}^{C} if the following conditions are met:

- The loads involved in the collecting route *i-l-j* do not exceed the capacity of the vehicle, i.e., $q_{ij} + q_{lj} \leq C$. This condition guarantees the feasibility of the collecting route in terms of capacity.
- The cost of this collecting route does not exceed the sum of the **minimum costs** that can be attributed to q_{ij} and q_{lj} (for this concept see section 5.2.1) when transported according to the initial solution. Completion of this condition generates certain expectations that sending q_{ij} and q_{lj} via collecting route *i-l-j* (instead of being sent in the manner indicated by ψ^{0}) will reduce the cost of the initial solution.
- The delivery time (24 or 48 hours) of q_{ij} and q_{lj} when transported on this collecting route should be the same as or better than the delivery time of both loads in the initial solution. This condition ensures that the level of service associated with the initial solution is maintained or improved when introducing collecting routes.

Similar conditions are imposed on loads q_{il} pertaining to set S_{ij}^{P} (see Figure 6b).

5.3.2. Initialization of the second ACO process. Pheromone trails and heuristic information

A trail of pheromone and a heuristic information parameter are defined for each element of sets S_{ij}^{C} and S_{ij}^{P}:

Pheromone trails $\qquad \tau_{ilj}^{C}(t) \quad \forall q_{lj} \in S_{ij}^{C}; \qquad \tau_{ilj}^{P}(t) \quad \forall q_{il} \in S_{ij}^{P} \qquad (10)$

Heuristic information u_{ilj} $\forall q_{lj} \in S_{ij}^{C}$, $\forall q_{il} \in S_{ij}^{P}$ (11)

As occurred in the D-H problem, a range is established for the pheromone trails $[\tau_{min2}, \tau_{max2}]$. All the trails must be initialized with the value τ_{max2} at the start of the iterative process.

The heuristic information chosen for this subproblem is given by the parameters u_{ilj}. These parameters are calculated using formula (12), which confers higher parameter values on those peddling/collecting routes *i-l-j* which deviate less from the main direction of travel, *i-j* (see Figure 7). In this formula, d_{il} and d_{lj} correspond to the distance between terminals *i* and *l*, and *l* and *j*, respectively. These parameters must be calculated at the start of the second ACO process.

$$u_{ilj} = \frac{1}{d_{il} + d_{lj}} \quad \forall q_{lj} \in S_{ij}^{C}, \forall q_{il} \in S_{ij}^{P} \tag{12}$$

5.3.3. Construction of a solution for the D-H-P/C subproblem

Artificial ants construct solutions for the D-H-P/C subproblem, departing from the solution found for the D-H subproblem and improving on it by adding peddling/collecting routes. The heuristic process which the ants employ in order to construct peddling/collecting routes along which to send loads that in the initial solution were sent directly or via hubs is explained below. The flow chart corresponding to this process is shown in Figure 8.

Firstly, the algorithm constructs a list (L) with all the origin-destination pairs within the system and arranges them in decreasing order according to direct distances between origin and destination. The first pair on the list is taken and the availability of candidate loads for peddling or collecting is checked. If there are candidates, then the ant chooses from among them in a probabilistic manner according to the formulas:

$$P_{ilj}^{C}(t) = \frac{\left[\tau_{ilj}^{C}(t)\right]^{\alpha_2}\left[u_{ilj}\right]^{\beta_2}}{\sum_{f|q_{fj}\in S_{ij}^{C}}\left[\tau_{ifj}^{C}(t)\right]^{\alpha_2}\left[u\right]^{\beta_2} + \sum_{f|q_{if}\in S_{ij}^{P}}\left[\tau_{ifj}^{P}(t)\right]^{\alpha_2}\left[u_{ifj}\right]^{\beta_2}} \tag{13}$$

$$P_{ilj}^{P}(t) = \frac{\left[\tau_{ilj}^{P}(t)\right]^{\alpha_2}\left[u_{ilj}\right]^{\beta_2}}{\sum_{f|q_{fj}\in S_{ij}^{C}}\left[\tau_{ifj}^{C}(t)\right]^{\alpha_2}\left[u_{ifj}\right]^{\beta_2} + \sum_{f|q_{if}\in S_{ij}^{P}}\left[\tau_{ifj}^{P}(t)\right]^{\alpha_2}\left[u_{ifj}\right]^{\beta_2}} \tag{14}$$

Given the pair (*i*,*j*), $P_{ilj}^{C}(t)$ refers to the probability with which the ant chooses the load $q_{lj} \in S_{ij}^{C}$ to make a collecting route according to the *i-l-j* sequence in iteration *t* (see Figure 6a); $P_{ilj}^{P}(t)$ corresponds to the probability with which the ant chooses the load $q_{il} \in S_{ij}^{P}$ to make a peddling route according to the *i-l-j* sequence (Figure 6b). Given this probabilistic rule, the ants are more likely to choose those candidate loads with a higher value of u_{ilj} (and thus produce routes which deviate less from the main direction of travel) and that, in previous iterations have led to better solutions (the information coming from the pheromone trails).

After the choice is made, the load chosen must be eliminated from the list L and from the sets S_{ij}^{C} and S_{ij}^{P}. Then, an attempt is made to introduce more stops into the peddling/collecting route which is being constructed, until one of the following two conditions is produced, in which case the route is said to be *saturated:*

- There is insufficient space in the vehicle to introduce another load (taken from those candidate loads which are still available)
- The route becomes so long that it affects the delivery times of the loads involved.

To introduce new stops in a collecting route, loads from S_{lj}^{C} are chosen, with q_{lj} being the last load introduced into the route. In the case of a peddling route, a load is chosen from those in S_{il}^{P}, with q_{il} being the last load introduced into the route (note that in S_{lj}^{C} and S_{il}^{P} there will only be loads which have not been used up to that moment). This choice will also be made in a probabilistic manner, using very similar formulas to (13) and (14).

When a route is saturated, a new route begins to be constructed with the next origin-destination pair available from list L. When no more pairs are available in L, the construction process for the solution is complete. The solution will consist of:

- the peddling/collecting routes constructed by the ant,
- the loads not involved in these peddling/collecting routes being transported according to the initial solution (the solution obtained for the D-H subproblem). Thus, some direct routes and routes via hubs will still remain, though some of these might have fewer loads to transport than in the initial solution. This will be because some of the loads are now transported on peddling/collecting routes.

5.3.4. Local improvement and pheromone update

Given an iteration *t*, when all the ants of the colony have constructed a solution, the algorithm tries to improve the *iteration-best* solution obtained. A local improvement for all the ants of the colony would consume an excessive running time. This process detects vehicles which are not sufficiently full and tries to send the loads in each via other routing alternatives, taking advantage of space in other partially full vehicles. In this way, the cost of the solution decreases since the number of nearly empty vehicles is minimized (for more details of the improvement process see Barcos, 2002).

After this local improvement, resulting in $\psi_2^{ib*}(t)$, the *global-best* solution is recorded and the pheromone trails are updated. Similar formulas to those already used in the resolution of the D-H subproblem are employed in this updating process. Additionally, pheromone values are reinitialized when needed, following a similar strategy to that employed for the D-H subproblem.

Once all the iterations of the process have been concluded, the final solution to the D-H-P/C subproblem corresponds to $\psi^* = \psi_2^{gb}(NI_2)$

6. Application of the algorithm to a real case

One of the main objectives behind this paper was to test the feasibility of the algorithm in the resolution of real-life problems by implementing the algorithm in C language. In order to do that, the algorithm was applied to the real case of an LTL carrier operating in Spain and Portugal.

The 49 main EOL terminals (2352 origin-destination pairs) with which the company was working in Spain and Portugal were considered. Six of these terminals acted simultaneously as break-bulk terminals. The service requirements introduced into the algorithm corresponded to the levels provided by the company at the time when the study was conducted, 87% of the dispatches being delivered in 24 hours, the rest were delivered within 48 hours. We worked with a load flow matrix obtained from statistical studies of historical data provided by the company. These load flows reflected a scenario of loads (Q_{ij}) to be transported on a regular daily basis. The distances between each origin-destination pair was determined by means of a road map.

Cost parameters were estimated from data provided by the company. Estimations could also be made for the time required for loading and unloading of freight for peddling/collecting and time spent in the reorganization of the freight in the hubs which was assumed to be the same for all the hubs in the system. A mean travel speed was also assumed for all the routes and the capacity of the vehicles was considered equivalent to that of a trailer, taking into account a percentage of wasted storage space arising from the organization of the load in the vehicles. This percentage was estimated using historical data.

In addition, it was also necessary to set the value of the parameters that influence the two ACO processes used in this paper. Several experiments were carried out on the real problem to determine the value of the parameters which lead, in general, to better solutions (see Barcos, 2002). These values are summarized in Table 1.

Initially, the values used for the pheromone limits were calculated from the following formulas:

$$\tau_{max} = \frac{1}{\rho} \cdot \frac{1}{Cost(\psi^{gb})} \qquad \tau_{min} = \frac{\tau_{max}}{2n} \qquad (15)$$

where n is the number of origin-destination pairs in the system and ψ^{gb} the global best solution. These formulas were adapted from those used in Stützle and Dorigo (1999a and 1999b). However, experience showed that the performance of the algorithm was better when values similar to those which appear in the table were used for τ_{max1} and

τ_{min1}. In the case of τ_{max2} and τ_{min2}, the formulas (15) have been shown to be appropriate.

The optimization process began with the calculation of the load flows q_{ij} by means of dispatching full vehicles directly (see Figure 4). Afterwards, the D-H subproblem was solved five times in order to find different solutions which could serve as initial solutions for the D-H-P/C subproblem (ψ^0). Previous experiments revealed that the probability of finding a better final solution depended considerably on the quality of ψ^0. With the best solution found for the D-H subproblem, the D-H-P/C subproblem was then solved five times. A significant reduction of the objective function was thus obtained due to the introduction of the peddling/collecting routes. With the best solution found for this second subproblem (ψ^*) the empty balancing movements were calculated to provide the final solution for the general problem.

The cost of the final solution was compared to that of the regular routes that the company were working with at that moment and which had been established by them over the years through experience[2]. The comparison was made in terms of the cost per unit of freight transported. The best solution found by the algorithm for an estimated average speed of 80 km/hr indicated cost savings of 10.8 %, with 87% of the shipments being delivered in 24 hours.

In addition, the computation time required by the algorithm using an Intel Core2 Duo CPU T7100 with 1,8GHZ and 2046MB was less than 38 CPU minutes (32.7 min. for one run of the D-H subproblem, 5.2 min. for one run of the D-H-P/C subproblem),

---

[2] This actual cost provided by the company was also used as an initial estimation of $Cost(\psi^{gb})$ in order to calculate formula (15) at the beginning of the application process. Later, the cost of the global-best solutions provided by the algorithm in subsequent experiments were used.

which represents a very reasonable time when planning at a tactical level, as in this case.

Figure 9 shows a convergence diagram of the solution provided by the algorithm for each of its subproblems. It can be observed how the introduction of peddling/collecting routes for solving the D-H-P/C subproblem represents a significant improvement in the result and how, subsequently, there is a refinement of the solutions as the number of iterations increases.

7. Discussion of results

During the experiments it was observed that the cost of the solution provided by the algorithm was very sensitive to the average vehicle speed used, a slight decrease in velocity provoking a considerable increase in cost.

In fact, when the speed diminishes, in order to maintain the level of service demanded of the system, loads which were delivered via hubs are now transported directly or by using peddling/collecting routes, thereby augmenting the cost of the solution considerably. This is clearly reflected in Table 2.

Other studies, such as Crainic and Roy (1988) and Roy and Delorme (1989) also show that transportation and handling costs increase with the service performance of solutions, this increase being more pronounced for higher levels of service.

In Table 2, it can be observed that the solutions provided by the algorithm showed savings in cost for speeds higher than 77 km/hr, maintaining a level of service of 87%. Although the mean travel speed estimated for the problem was 80 km/hr, an analysis was considered necessary to evaluate the shortcomings of the algorithm in a conservative scenario. In the analysis the two following aspects were taken into account:

(a) The minimum speed that can provide a level of service of 87% is 75 km/hr. for the problem at hand. For speeds less than this, not even sending all the loads directly to their destinations can provide the level of service required above.

(b) As far as the present study is concerned, in order to complete the 24-hour service, the transit time of the loads must be less than $M_{ij}^{24} = 10.5\, hr$. In practice, however, these service restrictions could be relaxed for some origin-destination pairs. In fact, according to the departure and arrival timetables of the vehicles programmed by the company for their routes, 17 % of the shipments had a transit time of 11 hours or more, yet were considered as having been served in 24 hours.

Taking these two aspects into account, the analysis was carried out using a speed of 75km/hr with service restrictions relaxed for 17% of the shipments (i.e., 17% of the shipments were allowed to have a transit time of between 10.5 and 11 hours, yet were considered as having been served in 24 hours).

As can be seen in Table 3 (in the column corresponding to 75* Km/hr), the results obtained indicated cost savings of 8.3% as compared to the existing route system of the company. Furthermore, the table reflects a significant reduction in costs relative to those obtained when no relaxed service restrictions were considered (column corresponding to 75 km/hr). The reason for this cost reduction is that some loads that previously could only be transported directly or by using peddling/collecting routes (in order to not exceed the time limits) are now sent via break-bulk terminals, taking advantage of the consolidation afforded by those hubs. In Table 3, this is reflected by the reduction in the costs due to direct and peddling/collecting routes and the increase in the costs due to routes via hubs.

In addition, it was also decided to evaluate the sensitivity of cost to increases in load reorganisation times at the hubs, θ^{r}_{hub}. In this experiment a speed of 75 km/hr was used. Again, a level of service of 87% was demanded and service restrictions were relaxed for 17% of the shipments. The results reflect the fact that the algorithm continues to supply less costly solutions than the company's existing system of routes, even when the load reorganisation times at the hubs increases by up to 20% of that used initially (see Table 4).

8. Concluding remarks

The algorithm developed carries out the routing design for less-than-truckload motor carriers, determining, for each origin-destination pair , the route to follow and the corresponding type of service (24/48 hours). It does this by looking for solutions which minimize the total cost of the system and that respect a predetermined minimum level of service. The algorithm considers direct routes, routes via hubs and peddling/collecting routes.

The results obtained in this paper are encouraging and show great consistency with reality. These results imply a significant reduction in cost compared to the cost of the actual routes used by the company and ensure an equal or higher level of service. At the same time, the computation time used is very reasonable, given that the problem being resolved involves planning at a tactical level. Thus, ACO optimization applied to the design of transport routes in *many-to-many* real size systems has demonstrated to be a viable and promising methodology.

In addition, the algorithm proposed in this paper can be a very useful tool when analysing the cost/service trade-offs and the impact that certain changes in the system (such as changes in the mean travel speed, reorganization time at the hubs or changes in the time limit for arrival of the vehicles at the destination terminals) can have on the result. The algorithm can also be exploited to evaluate the consequences of changes to the network of terminals, such as the opening or closure of certain hubs.

However, since this was a first attempt to find a solution to this real-life problem using optimization methodologies, the application has been carried out in a simplified manner. This means that the model may be improved by including other aspects of the real-life problem more in line with reality. Accordingly, for example, the possibility of using a fleet of non-homogenous vehicles was not considered in this paper (container trucks are usually used for long distances, but on some routes it may be appropriate to use vehicles with less capacity). In addition, the model does not include capacity restrictions at the hubs, so there is the possibility that some hubs may be saturated and others underused. An effort must be made in future research to improve these and other aspects of the model. Results obtained justify continuing the research with ACO approach.

References

Ballou, R., Masters, J.M., 1993. Commercial software for locating warehouses and other facilities. Journal of Business Logistics 4 (2), 71-107.

Barcos, L., 2002. Routing optimization for many-to-many freight transportation systems using Ant Colony Optimization. PhD thesis, Tecnun, University of Navarra (in Spanish).

Bullnheimer, B., Hartl, R.F., Strauss, C., 1999. Applying the Ant System to the Vehicle Routing Problem. In: Voss, S., Martello, S., Osman I.H., Roucairol, C. (Eds.), Meta-Heuristics: Advances and Trends in Local Search Paradigms for Optimization, Kluwer, Boston, 285-296.

Crainic, T.G., Rousseau, J.M., 1986. Multicommodity, Multimode Freight Transportation: A General Modeling and Algorithmic Framework for the Service Network Design Problem. Transportation Research B 20 (3), 225-242.

Crainic, T.G., 1988. Rail Tactical Planning: Issues, Models and Tools. In: Bianco, L., Bella, A.L., (Eds.), Freight Transport Planning and Logistics, Springer-Verlag, Berlin, 463-509.

Crainic, T.G., 2002. A Survey of Optimization Models for Long-Haul Freight Transportation. CRT-2002-05, Centre de recherché sur les transports, Université de Montréal.

Crainic, T.G., Roy, J., 1988. O.R. Tools for Tactical Freight Transportation Planning. European Journal of Operational Research 33, 290-297.

Daganzo, C.F., 1996. Logistics Systems Analysis. Springer-Verlag, Berlin.

Dorigo, M., Di Caro, G., Gambardella, L.M., 1999. Ant Algorithms for Discrete Optimization. Artificial Life 5 (3), 137-172.

Dorigo, M., Maniezzo, V., Colorni, A., 1996. The Ant System: Optimization by a Colony of Cooperating Agents. IEEE Transactions on Systems, Man and Cybernetics 26B (1), 29-42.

Gambardella, L.M., Dorigo, M., 1997. HAS-SOP: Hybrid Ant System for the

Sequential Ordering Problem. Technical Report IDSIA-11/97, IDSIA, Lugano, Switzerland.

Gambardella, L.M., Taillard, É., Agazzi, G., 1999. MACS-VRPTW: A Multiple Ant Colony System for Vehicle Routing Problems with Time Windows. In: Corne, D., Dorigo, M., Glover, F. (Eds.), New Ideas in Optimization, McGraw-Hill.

Hall, R.W., 1987. Comparison of strategies for routing shipments through transportation terminals. Transportation Research A 21 (6), 421-429.

Hall, R.W., 1989. Configuration of an overnight package air network. Transportation Research A 23 (2),139-149.

Leung, J.M.Y., Magnanti, T.L., Singhal, V., 1.990. Routing in point-to-point delivery systems: Formulations and solution heuristics. Transportation Science 24 (4), 245-260.

Merkle, D., Middendorf, M., Schmeck, H., 2002. Ant Colony Optimization for Resource-Constrained Project Scheduling. IEEE Transactions On Evolutionary Computation 6 (4), 333-346.

Miller, T., 1993. Learning about facility location models. Distribution 22 (5), 47-50.

Powell, W.B., 1986. A local improvement heuristic for design of less-than-truckload motor carrier networks. Transportation Science 20 (4), 246-257.

Powell, W.B., Sheffi, Y., 1983. The load planning problem of motor carriers: problem description and a proposed solution approach. Transportation Research A 17 (6), 471-480.

Powell, W.B. and Sheffi, Y., 1989. Design and implementation of an interactive optimization system for network design in the motor carrier industry. Operations Research 37 (1), 12-29.

Robusté, F., Almogera, J.M., Gargallo, X., Ardanuy, A., 1996. Sistema de ayuda a la decisión para la reestructuración de la red logística de una empresa de transporte urgente. In: Actas del II Symposium de Ingeniería de los Transportes, Vol II, pp. 269-277, Madrid, Colegio de Ingenieros de Caminos, Canales y Puertos.

Rodriguez, M.V, Álvarez M.J., Barcos L., 2007. Hub Location under capacity constraints. Transportation Research E 43, 495–505.

Roy, J., 2001. Recent trends in logistics and the need for real-time decision tools in the trucking industry. Proceedings of the 34th Hawaii International Conference on System Sciences.

Roy, J., Delorme, L., 1989. NETPLAN: A Network Optimization Model for Tactical Planning in the Less-than-Truckload Motor-Carrier Industry. INFOR 27 (1), 22-35.

Stützle, T., 1997. An Ant Approach to the Flow Shop Problem. In: Proceedings of the 6th European Congress on Intelligent Techniques & Soft Computing (EUFIT'98) 3, 1560-1564, Verlag Mainz, Aachen.

Stützle, T. and Dorigo, M., 1999a. ACO Algorithms for the Quadratic Assignment Problem. In: Corne, D., Dorigo, M., Glover, F. (Eds.), New Ideas in Optimization, McGraw-Hill.

Stützle, T. and Dorigo, M., 1999b. ACO Algorithms for the Traveling Salesman Problem. In: Miettinen, K., Makela, M., Neittaanmaki, P., Periaux, J.(Eds), Evolutionary Algorithms in Engineering and Computer Science, Wiley.

Stützle, T., Hoos, H., 1997. The MAX-MIN Ant System and Local Search for the Traveling Salesman Problem. In: Baeck, T., Michalewicz, Z., Yao, X. (Eds.), Proceedings of the IEEE International Conference on Evolutionary Computation, ICEC'97, 309-314.

Appendix 1. Input data used by the algorithm

The information regarding the freight transportation system which must be entered as inputs is summarized below:

- Network configuration: EOL terminals, set of break-bulk terminals, and distances between terminals.
- Load flows matrix: the matrix Q represents the load flow (volume of freight per day) that must be regularly transported between each origin-destination pair. The elements of this matrix are the Q_{ij} flows mentioned in previous sections.
- Fleet of vehicles: an unlimited fleet with homogenous capacity C is assumed.
- Cost structure: R, P and handling costs in each hub, c_r (see section 4)
- Time delay parameters: speed of service v_{ij}; stopping times at the EOL terminals for delivering or collecting freight (θ_{sf} and θ_{sv}); reorganization time for each hub r, θ^r_{hub}; earliest departure time at the origins (TL_i for each origin i), latest arrival time at the destinations (TA_j for each destination j) and level of service demanded, $L_{service}$.

TABLES AND FIGURES

D-H	D-H-P/C
$\alpha_1 = 2$	$\alpha_2 = 1$
$\beta_1 = 23$	$\beta_2 = 1$
$\rho_1 = 0.6$	$\rho_2 = 0.1$
$\tau_{max1} = 0.01$	$\tau_{max2} = 1/[\rho_2 * Cost(\psi_2^{gb})]$
$\tau_{min1} = 0.0$	$\tau_{min2} = \tau_{max2} / 2n$
NA_1: 150	NA_2: 100
NI_1 : 12000	NI_2: 1250

NA: number of ants in the colony; NI: number of iterations;

Table 1. Parameter settings.

	Mean travel speed (km/hr)					
	80	79	78	77	76	75
Total cost (in 1000 €)	286.2	294.9	315.0	328.9	371.3	398.6
Cost savings (%)	10.8	8.1	1.8	-2.5	-15.7	-24.2
Costs of the consolidation routes (in 1000 €)						
Cost due to direct routes	16.7	25.0	29.1	31.8	39.0	48.0
Cost due to routes via hubs	113.0	107.5	104.2	103.1	93.2	72.4
Cost due to peddling/collecting routes	127.2	133.5	152.3	164.4	210.2	250.3

Table 2. Variation of solution costs as estimated average speed decreases. Total costs include full vehicle costs, consolidation route costs and empty return costs.

	Mean travel speed (km/hr)	
	75	75*
Total cost (in 1000 €)	398.6	294.1
Cost savings(%)	-24.2	8.3
Costs of the consolidation routes (in 1000 €)		
Cost due to direct routes	48.0	25.4
Cost due to routes via hubs	72.4	111.9
Cost due to peddling/collecting routes	250.3	127.7

Table 3. Variation of solution costs when service restrictions are relaxed for 17% of the shipments

	Reorganization time in hubs θ^{r}_{hub}		
	2.5 (initial value)	2.75	3
Total cost (in 1000 €)	294.9	308.1	320.5
Cost savings (%)	8.1	4.0	0.1

Table 4. Variation of solution costs when the load reorganization time at hubs increases.

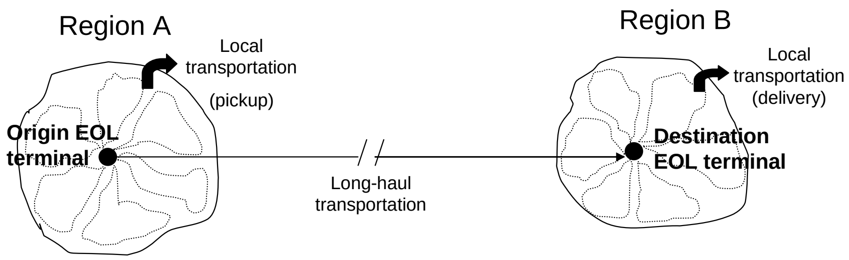

Figure 1. Local and long-haul transportation.

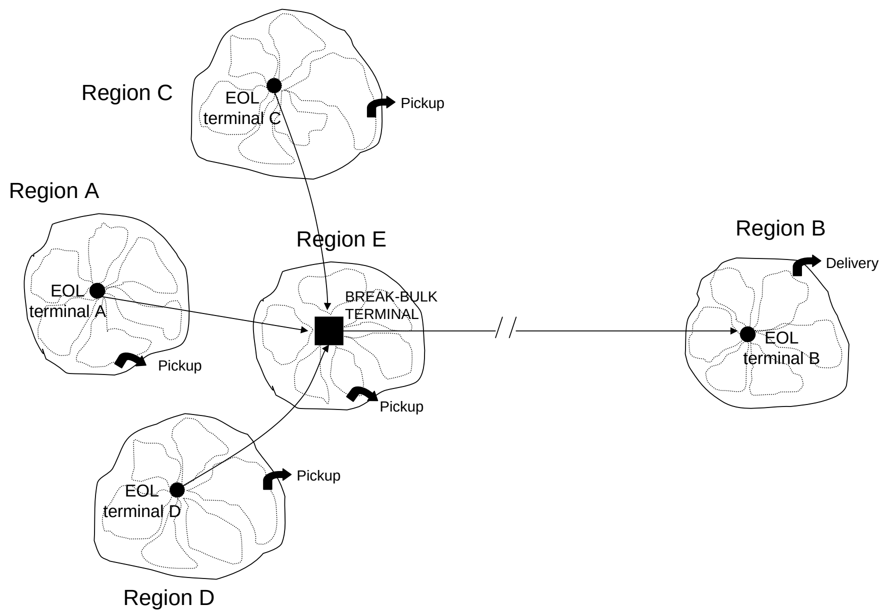

Figure 2. Long-haul transportation through a break-bulk terminal.

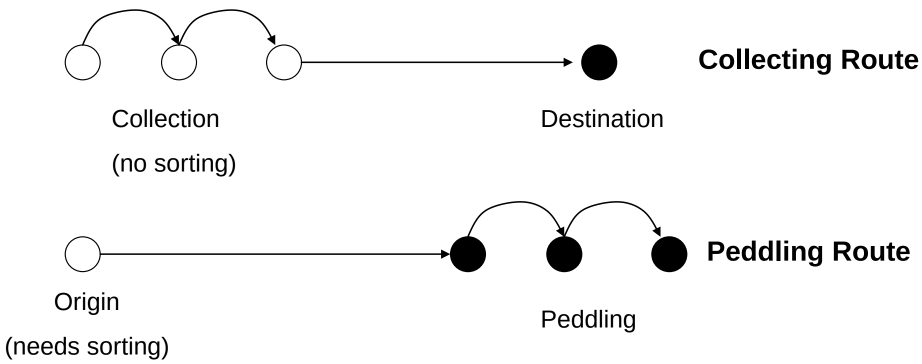

Figure 3. Collecting and peddling routes

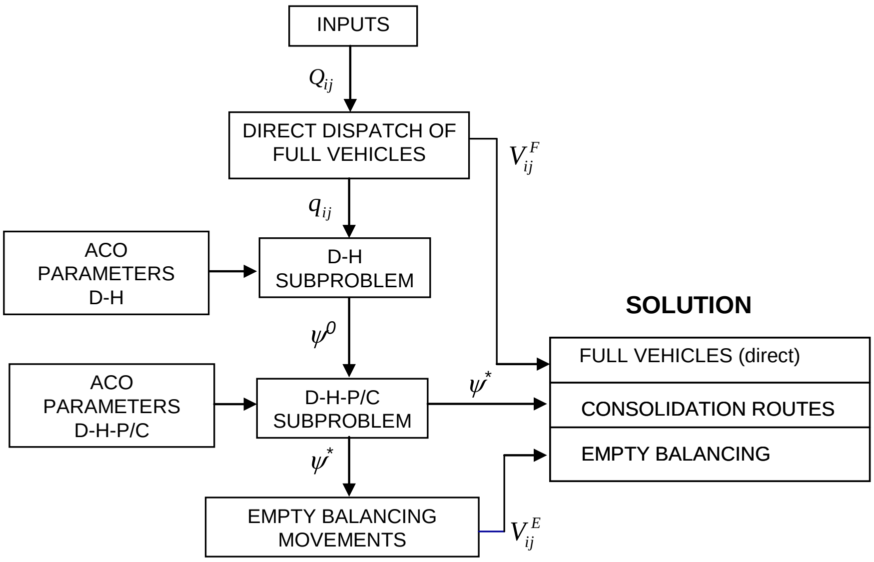

Figure 4. General scheme of the ACO algorithm.

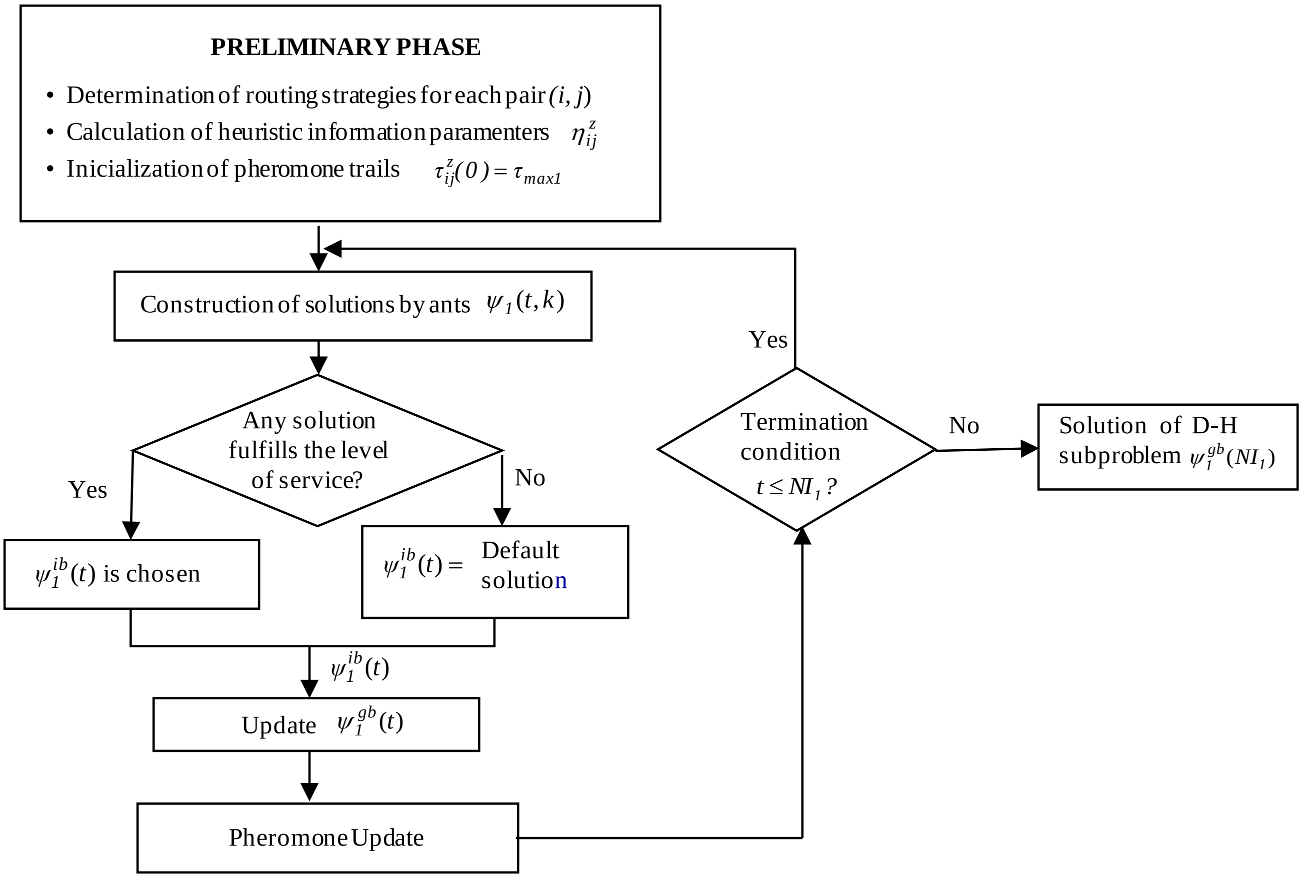

Figure 5. Iterative process for the D-H subproblem

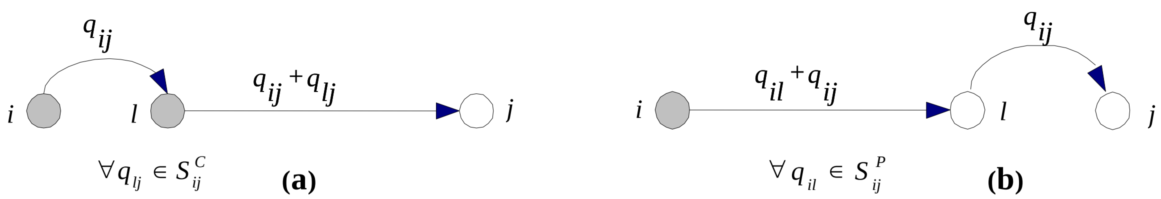

Figure 6. Sets of candidate loads for collecting/peddling

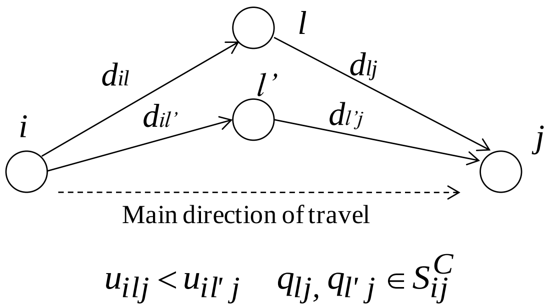

Figure 7. Comparison of the heuristic information parameter values for two loads pertaining to the set S_{ij}^{C}

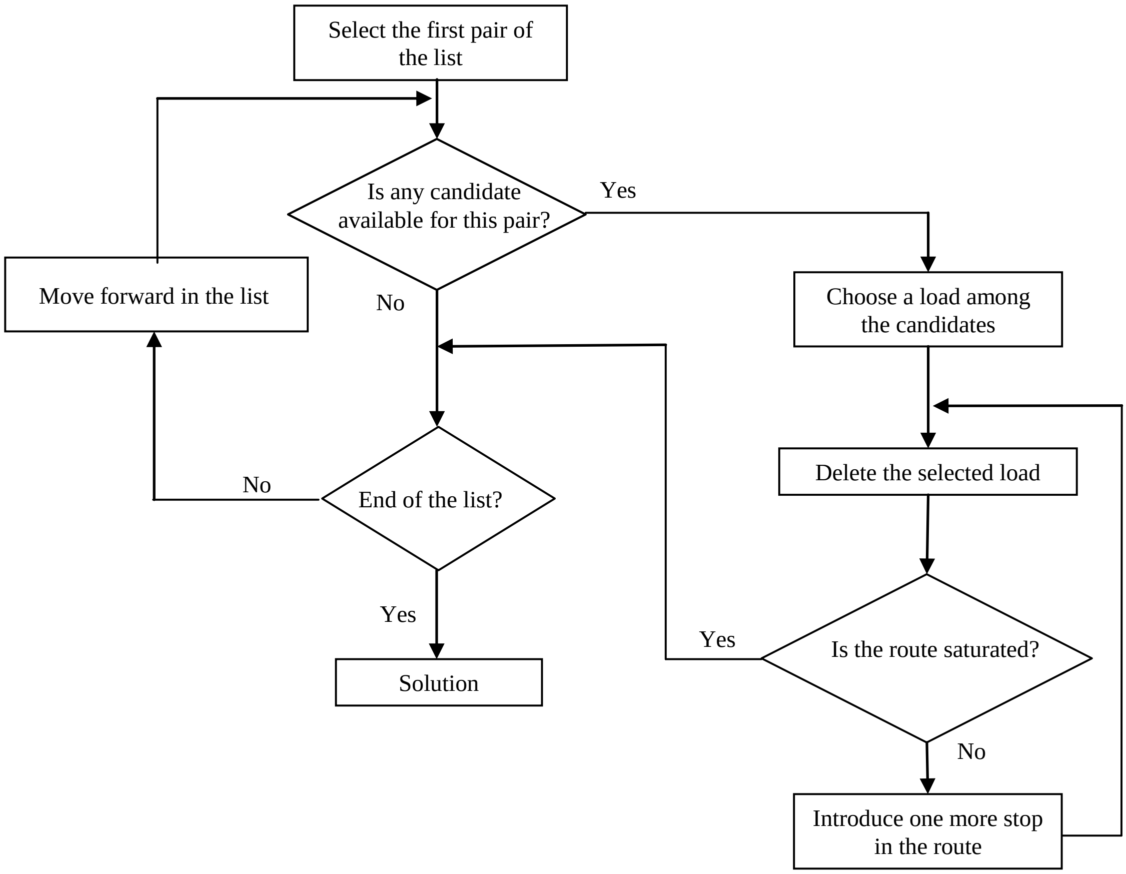

Figure 8. Construction of solutions for the D-H-P/C subproblem

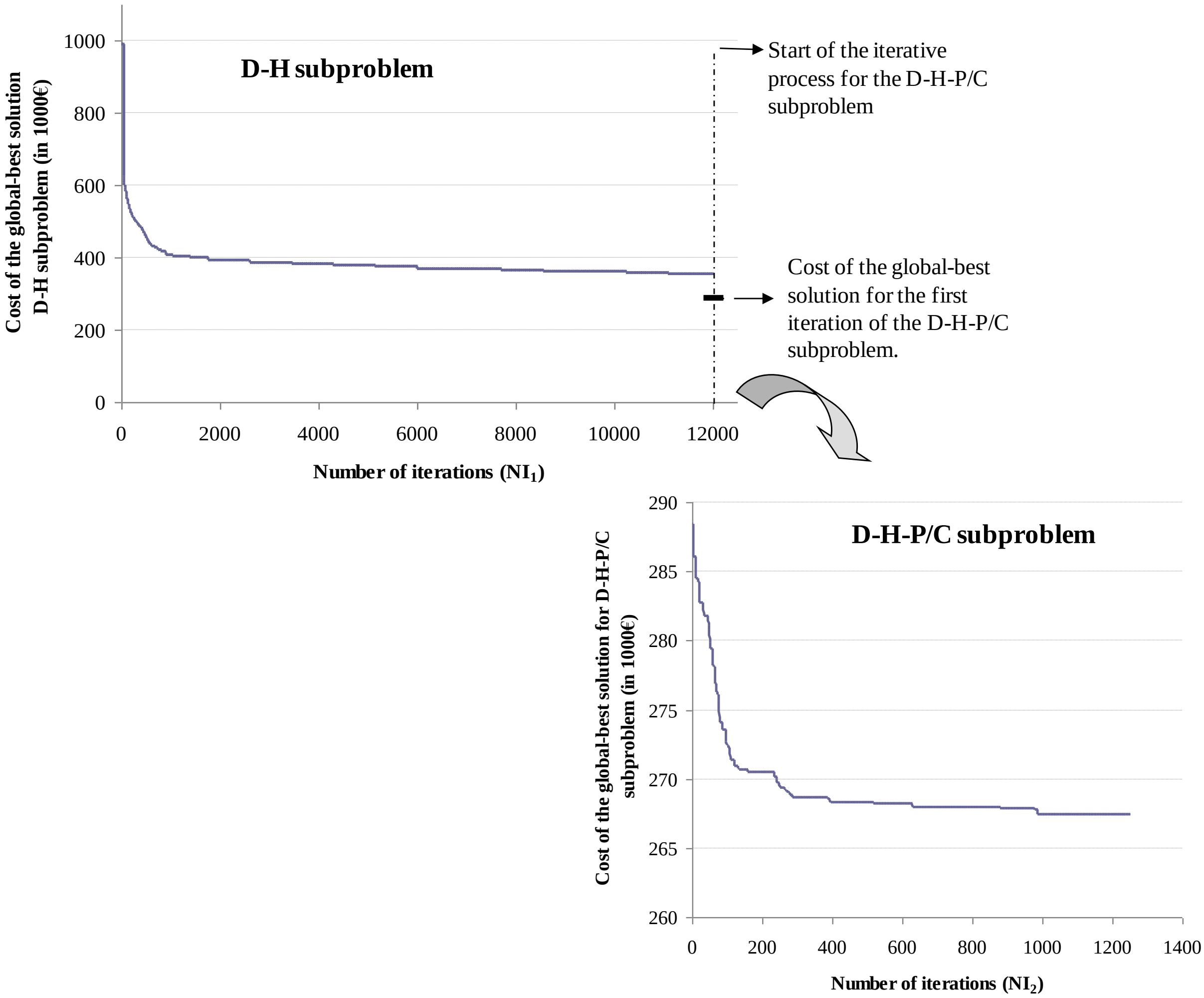

Figure 9. Convergence diagrams for a typical run of the problem addressed in the application. Cost of the global best solution in thousands of euros per working day.