Available online at www.sciencedirect.com
ScienceDirect
European Journal of Operational Research 179 (2007) 968–977
EUROPEAN JOURNAL OF OPERATIONAL RESEARCH
www.elsevier.com/locate/ejor

Using clustering analysis in a capacitated location-routing problem

Sérgio Barreto [a,e,*], Carlos Ferreira [b,e], José Paixão [c,e], Beatriz Sousa Santos [d]

[a] *Higher Institute of Accounting and Administration/ISCA, University of Aveiro, Aveiro, Portugal*
[b] *Department of Economics, Management and Industrial Engineering, University of Aveiro, Aveiro, Portugal*
[c] *Department of Statistics and Operations Research, University of Lisbon, Lisbon, Portugal*
[d] *Department of Electronics and Telecommunications/IEETA, University of Aveiro, Aveiro, Portugal*
[e] *Operational Research Centre, University of Lisbon, Lisbon, Portugal*

Received 15 December 2003; accepted 15 June 2005
Available online 3 May 2006

Abstract

The location routing problem (LRP) appears as a combination of two difficult problems: the facility location problem (FLP) and the vehicle routing problem (VRP). In this work, we consider a discrete LRP with two levels: a set of potential capacitated distribution centres (DC) and a set of ordered customers. In our problem we intend to determine the set of installed DCs as well as the distribution routes (starting and ending at the DC). The problem is also constrained with capacities on the vehicles. Moreover, there is a homogeneous fleet of vehicles, carrying a single product and each customer is visited just once. As an objective we intend to minimize the routing and location costs.

Several authors have integrated cluster analysis procedures in heuristics for LRPs. As a contribution to this direction, in this work several hierarchical and non-hierarchical clustering techniques (with several proximity functions) are integrated in a sequential heuristic algorithm for the above mentioned LRP model. All the versions obtained using different grouping procedures were tested on a large number of instances (adapted from data in the literature) and the results were compared so as to obtain some guidelines concerning the choice of a suitable clustering technique.

Keywords: Combinatorial optimization; Distribution; Heuristics; Cluster analysis; Location-routing

1. Introduction

In an ever more demanding society, having customers less and less willing to wait for the products they want to acquire, decisions concerning the location of distribution centres (DC) and tracing of distribution routes are a central problem, having implications on the complete supply chain (Bramel and Simchi-Levi, 1997). Nowadays, even small and medium enterprises should be aware that their future success may depend on the location–distribution decisions and recognise the need for flexible and efficient, as well as reliable, decision methods (White Paper, 2001).

* Corresponding author. Tel.: +351 234 380 110; fax: +351 234 380 111.
E-mail addresses: sergio.barreto@isca.ua.pt (S. Barreto), carlosf@egi.ua.pt (C. Ferreira), jpaixao@fc.ul.pt (J. Paixão), bss@det.ua.pt (B.S. Santos).

0377-2217/$ - see front matter © 2006 Published by Elsevier B.V.
doi:10.1016/j.ejor.2005.06.074

In the last four decades (Maranzana, 1963), the investigation on modelling and resolution of location-routing problems (LRP) has advanced and produced a large body of literature (Barreto et al., 2003b), allowing a complete and detailed view of this problem and its characteristics. However, a deficit in theoretical investigation at the simplest LRP level still exists. This deficit hinders a better understanding of its properties and the development of new approaches; furthermore, it holds back the creation of the solid foundations needed to support complex applications. As an attempt to contribute to this investigation, we present in this paper a study on a location routing problem, common in many organizations, which has two levels (customers and distribution centres) and vehicles with a limited capacity; we will call it a capacitated location-routing problem (CLRP).

2. A capacitated location-routing problem (CLRP)

Let a set of customers and potential distribution centres (DCs) be represented by points on the plane. Each customer has a certain demand (units of load); the location (installation) cost of each DC is known, as well as the unitary cost of distribution (function of covered distance). The vehicles (routes) and the potential DCs have a certain capacity (units of load). The purpose of this CLRP is, then, to choose the DCs that must be opened (installed) and to draw the routes from these DCs to the customers, having as an objective the minimization of the total cost (location and distribution costs).

Solving exactly a CLRP is a difficult task, since this type of problem is NP-complete (Srivastava, 1986). Thus, one way to "solve efficiently" large problems is to look for heuristics, which have some advantages such as: (i) getting "good solutions" in acceptable time; (ii) producing several "good solutions" allowing the user to choose the most suitable according to the scenario; (iii) being easy to understand, modify and implement, they allow to deal with larger problems.

3. A cluster analysis approach

Cluster analysis (Anderberg, 1973) studies the division of entities (as objects or individuals) in groups based in one or several of their characteristics. An important issue is the notion of group; according to Jain and Dubes (1988): "*Cluster may be described as connected regions of a multi-dimensional space containing a relatively high density of points, separated from other such regions by a region containing a relatively low density of points*".

This definition of group is an excellent reason to use cluster analysis in the resolution of LRPs. Recognizing groups of customers can be a good start to obtain good LRP solutions. Let us consider, as an example, Fig. 1 showing the location of the 150 biggest European cities (Daskin, 1995), where the squares represent potential DCs. Observing this figure, it seems natural that a CLRP related to these cities will tend to construct routes encompassing high density regions (population agglomerations). The identification by visual inspection of some groups of cities that can form good bases for the distribution routes (for example, the Canaries Islands, the Iberian Peninsula, the British Islands, the north of the Europe or Italy and Greece) is not difficult. However, what is relatively easy to visualize is more difficult to carry out. Nevertheless, cluster analysis provides a vast set of grouping methodologies that can be used in heuristic approaches to the CLRP.

The potential of cluster analysis for the resolution of LRPs (or problems directly related, as the facility location problem (FLP) and vehicle routing problem (VRP)) has also been recognized by other authors. Dantzig and Ramser (1959) were of the first ones to mention the identification of groups of points in the multiple travelling salesman problem (MTSP), stating that "*One would look for "clusters of points" and determine by trial and error the*

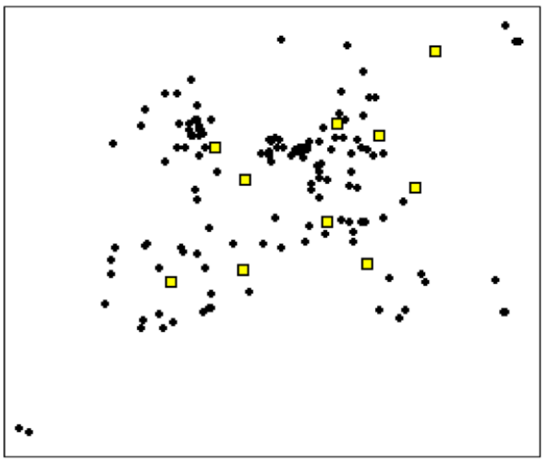

Fig. 1. Location of the 150 biggest European cities, where the squares represent potential DCs.

order in which they should be traversed, taking care that no loop crosses itself".

Concerning location problems, Kuehn and Hamburger (1963), also recognized the benefits of grouping (nearby customers) when they say that "...*in many cases a priori judgements can be made that customers in certain geographical regions will not be serviced from potential warehouses in other regions* (...) *customers can frequently be aggregated into concentrations of demand* (*for example, metropolitan chain grocery and wholesaler warehouses*) *because of geographical proximity*". This is probably one of the first explicit references on the interest of grouping customers. Fukunaga and Short (1978) used grouping algorithms in location problems as well.

The well known Clarke–Wright algorithm (Clark and Wright, 1964) uses grouping, with excellent results, to implement an original saving procedure to determine the proximity matrix. Bodin (1975) mentions the construction of groups as a technique to approach Routing and Scheduling Vehicles Problems. According to Madsen (1983), also Bednar and Strohmeier solve VRPs supported by grouping procedures. Cullen et al. (1981) make use of simple grouping procedures in the development of a heuristic for VRP. Branco and Coelho (1990) also use grouping processes in algorithms for a specific LRP (Hamiltonian p-median problem).

Min (1989) explicitly includes cluster analysis in algorithms for VRP with distribution and collection. He uses the group average proximity measure which, according to Romesburg (1984), has advantages over the single linkage or complete linkage measures. In another paper, Min et al. (1992) use the ward measure of proximity due to its tendency to form groups of equal dimension and the good results reported in some studies.

Srivastava (1993) also considers a LRP heuristic based on grouping procedures, using the minimum spanning tree to determine a set of groups, which he refines later by means of a 1-optimal method.

Bruns and Klose (1995) and Klose (1995) make explicit reference to the integration of the cluster analysis in a heuristic procedure for the LRP. They employ hierarchical grouping techniques with single and complete linkage, group average and ward proximity measures.

Min (1996) considers as well a sequential method for a LRP with capacity that starts by grouping customers through a hierarchical method and uses the ward proximity measure.

Actually, although one can find in the literature a significant number of attempts to integrate grouping techniques in algorithms for the LRP, the same is not true when looking for comparative studies among the several grouping techniques in order to evaluate their real capacities.

4. Proximity measures among groups

Cluster analysis considers different grouping methods accordingly to the type of variables (qualitative, quantitative, binary or mixed). In this paper we have quantitative variables (co-ordinates of customers and DCs on the plane), and the classification will use hierarchical and non-hierarchical methods (or partition methods).

Several measures have been proposed to determine the proximity between points on the plane (Anderberg, 1973; Gower, 1985), however the most common for quantitative data is the Euclidean metric that determines the proximity between the points $I=(x_i,y_i)$ and $J=(x_j,y_j)$ as $d(I,J)=\sqrt{(x_i-x_j)^2+(y_i-y_j)^2}$.

Based on this concept of proximity between two elements, some measures of proximity among groups have been proposed: single linkage (nearest neighbour), complete linkage (farthest neighbour), group average, centroid, ward and saving. With the exception of the saving measure, proposed in this study, these are some of the most popular in cluster analysis literature (Anderberg, 1973; Jain and Dubes, 1988; Kaufman and Rousseeum, 1990; Everitt, 1993). Fig. 2 shows a graphical and analytical representation of these proximity measures.

According to Jain and Dubes (1988), Jardine and Sibson considered that the single linkage is the only measure satisfying all the mathematical criteria they had defined. However, they concluded that the use of this measure usually produces worse results, when compared to the other measures.

An important element on the centroid and ward proximity measures is the centroid (gravity centre) of each group. In our case for two groups, A and B, we will have m_{A} and m_{B} defined as

$$m_{\mathrm{A}}=\left(\frac{\sum_{I\in A}x_i}{|A|},\frac{\sum_{I\in A}y_i}{|A|}\right),\quad m_{\mathrm{B}}=\left(\frac{\sum_{J\in B}x_j}{|B|},\frac{\sum_{J\in B}y_j}{|B|}\right).$$

The ward proximity measure (Kaufman and Rousseeum, 1990) also includes the sum square error of a group quantified as the sum of the square

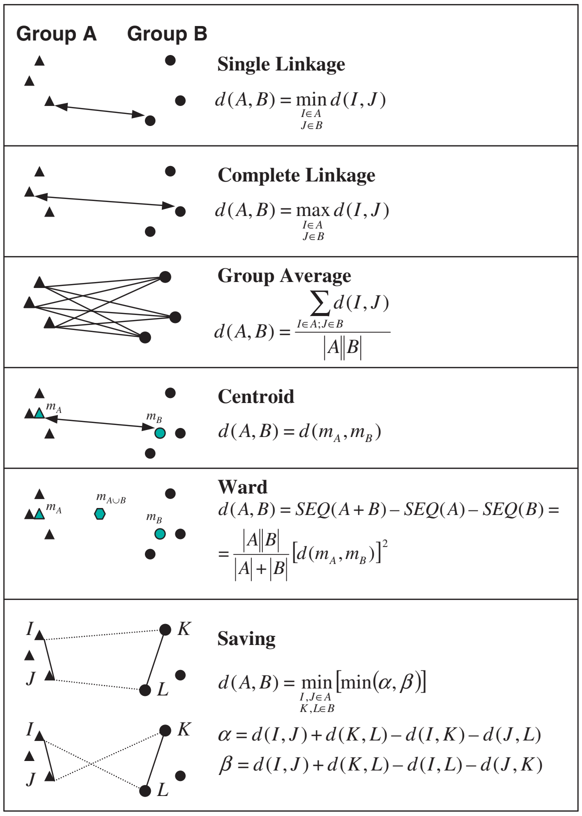

Fig. 2. Graphical and analytical representation of some proximity measures.

deviation of the elements of the group to the centroid, $\mathrm{SEQ}(A) = \sum_{I \in A}[d(I, m_A)]^2$. While this measure is widely used, the mean square error may prove inadequate to investigate the data structure, due to its tendency to form groups of equal size (Milligan and Schilling, 1985; Min, 1996). On the other hand, this can be an advantage to solve the LRP.

Golden and Meehl (1980) claim that group average, complete linkage and ward produce better results than single linkage and centroid for a specific data set. Bayne et al. (1980) affirm that ward and complete linkage are preferable to the centroid and group average measures.

Some authors have attempted to demonstrate the superiority of one or another proximity measure. Despite these efforts, we still do not have a clear idea on the general or specific potentialities of each one. The lack of agreement among the several studies reinforces the idea that there is not a measure which is adequate for all the applications and cases. In each case, the measure must be carefully chosen and, probably, only after tests with several measures, definitive conclusions can be drawn (Edelbrock and McLaughlin, 1980).

5. A cluster analysis based heuristic

Heuristic methods for resolution of LRPs can be classified as sequential and iterative. The former solve sequentially the location and the vehicle routing problems. The latter solve the same problems iteratively, while it is possible to improve the solution.

According to Min (1996) and Balakrishnan et al. (1987), for limited capacity of vehicles and significant fixed cost of distribution centres, the sequential methods are preferable from a computational point of view. Moreover, Srivastava and Benton (1990) conclude that the error associated to the obtained solutions is perfectly acceptable. As a consequence, we propose for the CLRP a sequential heuristic of the type distribution-first, location-second, which we present below.

A sequential heuristic for the CLRP

Input: Co-ordinates of $N = \{1, 2, \ldots, n\}$ customers on the plane with demand $e_i : i \in N$.
Co-ordinates of $P = \{n+1, n+2, \ldots, n+p\}$ potential DCs with capacity u_k and location cost $f_k : k \in P$.
w = vehicle capacity.
Output: Vehicles routes based in the DCs.
Step 1. Construct groups of customers with a capacity limit.
Step 2. Determine the distribution route in each customer group.
Step 3. Improve the routes.
Step 4. Locate the DCs and assign the routes to them.

Fig. 3 shows the results of each step of the heuristic applied to an instance of a CLRP with 50 customers and 5 potential DC (adapted from Christofides and Eilon (1969)).

In step 2, whenever the group has 40 customers or less, the TSP routes are determined by an exact algorithm which solves the relaxation of the subcycles constraints with more than three customers. These constraints are introduced later if they are violated in this step. If the route integrates more than 40 customers, a two stages heuristic procedure is used. In the first stage, a feasible solution is obtained using a choice criterion of the farthest type

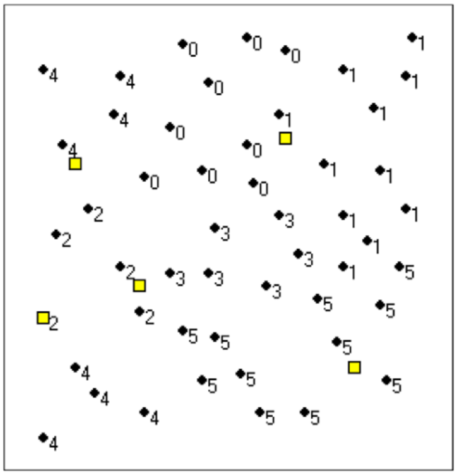

Fig. 3a. Step 1. Capacity limited group construction.

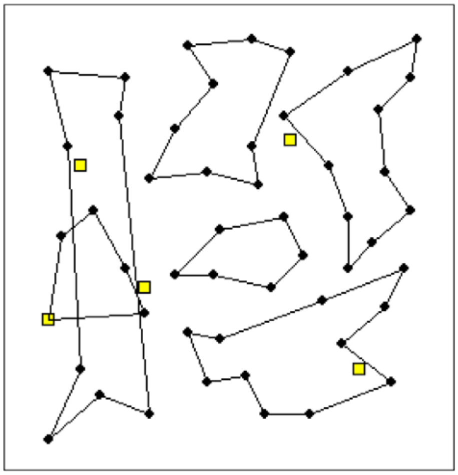

Fig. 3b. Step 2. Route design in each group. Cost = 526.

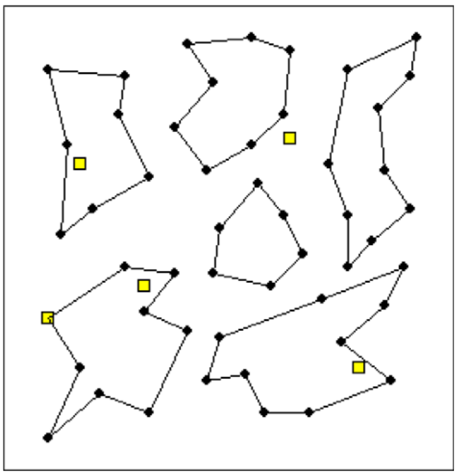

Fig. 3c. Step 3. Route improvement. Cost = 463.

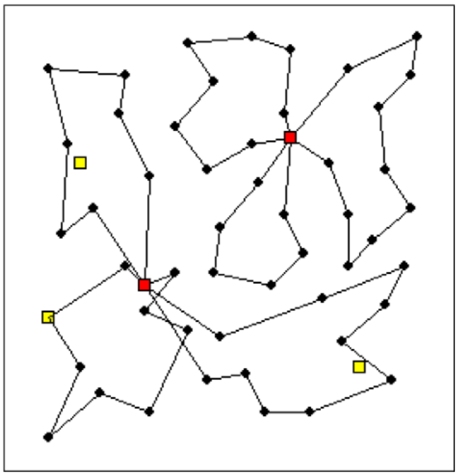

Fig. 3d. Step 4. DC location and route assignment to the open DCs. Cost = 614.

and an insertion criterion of the saving type. In the second stage, the solution is improved through a 3-optimal local search procedure.

In the improvement of the routes (step 3) a 3-optimal local search procedure is used, as the one proposed by Branco and Coelho (1990) for the Hamiltonian p-median problem.

After step 3, each route collapses into one customer with a saving type DC assignment cost. Then, the single source capacitated location problem is solved, leading to a feasible solution of the CLRP.

To implement step 1, four grouping methods are considered (two hierarchical and two non-hierarchical):

1. One-phase hierarchical method.
2. Two-phase hierarchical method.
3. Direct-assignment, non-hierarchical method.
4. Sequential-assignment, non-hierarchical method.

For each of these methods, the six proximity measures presented in Fig. 2 are evaluated. The integration of these methods in step 1 leads to four versions of the proposed heuristic, V1, V2, V3 and V4 (Fig. 4).

5.1. One phase hierarchical method (V1)

This is an agglomerative hierarchical method that performs an iterative merging of the nearest groups. There are several ways to implement agglomerative hierarchical methods, the most common ones being the spanning tree and the Johnson

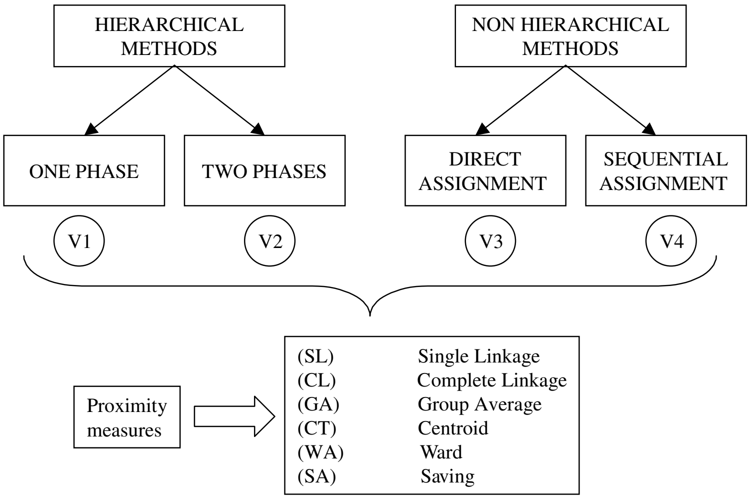

Fig. 4. Heuristic versions and proximity measures.

methods (Johnson, 1967). The former considers the formation of groups from the minimum spanning tree of a graph (Gower and Ross, 1969) having the disadvantage of supporting only the single linkage and complete linkage proximity measures. The later needs to store the triangular matrix of proximity between the groups; however it allows using all proximity measures showed in Fig. 2 and a similar algorithmic construction for the hierarchical and non-hierarchical cases.

Beginning with groups consisting of only one customer, the hierarchical algorithm leads to the formation of one single group. In CLRPs the capacity limit avoids the consecutive joining of groups, acting as a natural stopping criterion on their final number. However, the construction of groups with limited capacity leads to a difficulty to avoid drawback related to the final part of the grouping process: the groups reaching their capacity limits prevent the merging of the near groups since this will exceed their capacity. For this reason the merging of far away groups often occurs (as it can be observed in Figs. 3a and 3b) producing the *biased* effect (Klose, 1995).

5.2. Two phase hierarchical method (V2)

The two phase hierarchical method starts by applying the hierarchical method without capacity constraints, freely constructing a number of groups, thus preventing the undesirable biased effect. Knowing the demand of the customers (e_i) and the maximum vehicle capacity (w), the minimum number of vehicles (r) is determined as

$$r = \left\lceil \frac{\sum_{i \in N} e_i}{w} \right\rceil.$$

Next, the hierarchical grouping method is applied until there are only r groups. The lack of capacity constraints probably leads to the formation of groups that violate these constraints. In this case, the second phase is performed using a procedure that allows transferring customers from the groups exceeding capacity limits to other groups that can receive them without exceeding their capacity. Customer transfer is based on a relative proximity measure defined as follows.

5.2.1. Relative proximity measure

Let a set $G = \{G_1, G_2, \ldots, G_r\}$ include r groups on the plane. Let $i \in G_l \in G$ be a customer and $\mathrm{Prox}(i, G_j)$ be the proximity between customer i and group $G_j \in G$. The proximity coefficient of customer i is defined as

$$\rho_{\mathrm{prox}}(i) = \frac{\min_{G_j \in G \setminus G_l} \mathrm{Prox}(i, G_j)}{\mathrm{Prox}(i, G_l)}.$$

This proximity coefficient is a relative measure, independent of the customer space distribution, which represents the degree of proximity to the next group. The numerator of the expression represents the external proximity and the denominator the internal proximity.

In this second phase, customers having the least proximity coefficient are transferred, as long as the receiving group does not exceed its capacity. If, however, no customers can be transferred (and the violation of capacity still exists), the number of vehicles (r) is increased by one unit and the procedure restarts.

5.3. Direct assignment non-hierarchical method (V3)

While the hierarchical methods begin with a set of groups with one element and, through a nested process, converge to one group, the non-hierarchical methods are devised to construct r groups, where r is known a priori or is determined as part of the method.

In the direct assignment non-hierarchical method the minimum number of groups (r) is determined as in the two phase hierarchical method. Then, r customer sources (who will serve of first customer in each group) are established. To prevent the biased effect, the customer sources must be located on the boundary; therefore, they are chosen using a farthest neighbour proximity measure. The remaining customers are directly assigned to the group whose vertex source is next.

5.4. Sequential assignment non-hierarchical method (V4)

In the previous method, capturing new customers to the group depends entirely on the customer source; the remaining elements of the group do not contribute for this gravitational action. In the sequential assignment non-hierarchical algorithm, the responsibility of capturing not yet assigned customers is shared by the entire group. That is why it is necessary to calculate, in all iterations, the proximity between the free customers and the modified groups. Customer sources are determined as in the previous method.

6. Computational tests

To evaluate the four versions of the heuristic (V1, V2, V3, V4) and the six proximity measures (SL, CL, GA, CT, WA, SA) computational tests were carried out on 19 CLRP instances (see Table 1) obtained from the literature (Or, 1976; Perl, 1983) or adapted from data related with VRP (Gaskell, 1967; Christofides and Eilon, 1969; Min et al., 1992; Daskin, 1995). Data relative to the used instances are available in Barreto (2003) and the complete tables of results are available in Barreto (2004).

To evaluate the performance of each version of the heuristic, it is essential to define "good performance" and decide which measures should be used. Naturally, a good heuristic produces good results (in this case low cost CLRP solutions). However, a heuristic may provide a very good solution for a

Table 1
Instances lower and upper bounds

CLRP instance		Vehicles capacity	LB	UB	Gap
1	Gaskell67—21 × 5	6000	424.9*	435.9	2.59
2	Gaskell67—22 × 5	4500	585.1*	591.5	1.09
3	Gaskell67—29 × 5	4500	512.1*	512.1*	0.00
4	Gaskell67—32 × 5	8000	556.5	571.7	2.73
5	Gaskell67—32 × 5	11,000	504.3*	511.4	1.41
6	Gaskell67—36 × 5	250	460.4*	470.7	2.24
7	Christofides69—50 × 5	160	549.4	582.7	6.06
8	Christofides69—75 × 10	140	744.7	886.3	19.01
9	Christofides69—100 × 10	200	788.6	889.4	12.78
10	Perl83—12 × 2	140	204.0*	204.0*	0.00
11	Perl83—55 × 15	120	1074.8	1136.2	5.71
12	Perl83—85 × 7	160	1568.1	1656.9	5.66
13	Perl83—318 × 4	25,000	–	580,680.2	–
14	Perl83—318 × 4	8000	–	747,619.0	–
15	Min92—27 × 5	2500	3062.0*	3062.0*	0.00
16	Min92—134 × 8	850	–	6238.0	–
17	Daskin95—88 × 8	9,000,000	356.4	384.9	8.00
18	Daskin95—150 × 10	8,000,000	43,938.6	46,642.7	6.15
19	Or76—117 × 14	150	12,048.4	12,474.2	3.53
			Average gap		4.81
			Median		3.13

certain instance and have a poor performance in others. In this case we cannot say that it's a "good" heuristic. Thus, the evaluation of a heuristic must have into account its capacity to generate frequently "good" solutions. In short, a good heuristic must generate, for most of the instances, "good" solutions; still, it may not be able to find the best solution.

Three types of success rates were considered: (i) number of generated solutions equal to the best known solutions (BKS); (ii) number of solutions within a 2% tolerance from the BKS; (iii) number of solutions within a 5% tolerance from the BKS. Fig. 5 shows the success rates for all the versions using all proximity measures.

V2 heuristic, using the complete linkage proximity measure, yielded the best success rates. With a tolerance of 2%, the best results were obtained also by V2, but now for the group average proximity measure. V1 and V3 produced more solutions within a 5% tolerance. V3 version using SL, CL, GA, CT and SA proximity measures produced equivalent success rates.

A standardization of the data (customers and DC) for the square $[0, 500]^2$ was performed in order to get comparable results for all the instances used as well as to allow the use of some statistical measures. Fig. 6 shows the average results for all the versions and proximity measures: CAR represents the average cost after a routing process (after step 2), CAI the average cost after the improvement step (step 3) and CAL the average cost after the location procedure (after step 4) or the average CLRP cost. V1 and V3 produced the best average results. V2 and V4 produced the worst results. The average standard deviation in each version is 108.2, 217.0,

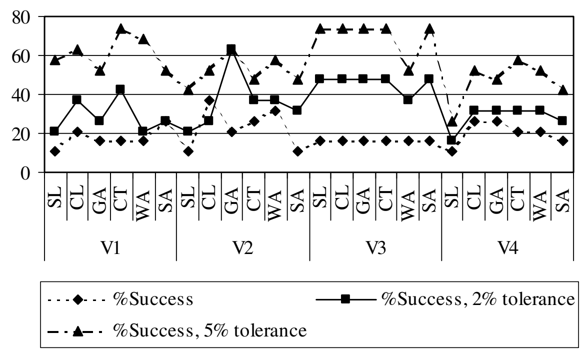

Fig. 5. Success rates obtained for the four versions of the heuristic (V1, V2, V3, V4) and the six proximity measures (SL, CL, GA, CT, WA, SA).

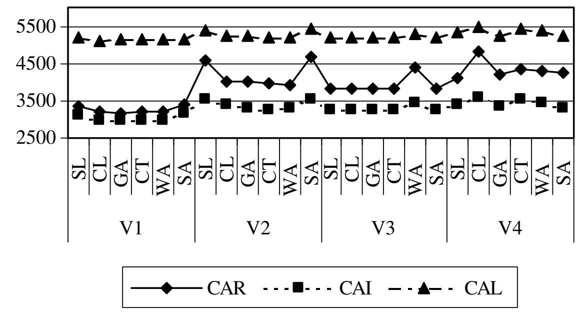

Fig. 6. Average results after steps 2, 3 and 4.

142.6 and 221.7, respectively, confirming the good performance of V1.

The route improvement procedure allowed a decrease of 16% in the route costs and, as shown in Fig. 6, had a significant impact in the cases of (V2,SL), (V2,SA), (V3,WA) and (V4,CL), attenuating the consequences of a poor initialization. Moreover, in Barreto et al. (2003a) we have shown that CLRP solutions obtained without the route improvement step are 25% worst. Furthermore, this step has an important role in the elimination of the biased effect as shown in Figs. 3a, 3b and 3c.

Fig. 7 shows the CLRP average results and confirms the good performance of V1 version. The group average (GA) proximity measure produces the most balanced results.

Generally, the analysis of the success rates and average results allows ordering the versions by decreasing performance: V1, V3, V2 and V4. Concerning the proximity measures, the best performance was obtained by GA followed by CT, SA, CL, SL and WA measures. Table 1 shows, for the 19 instances of the LRPC, a lower bound (LB) and an upper bound (UB) of the optimal solution cost. The running time was less than one second, except for instances 11 (202 seconds maximum), 13 and 14 (115 seconds). The CLRP instance column contains information about the author, the

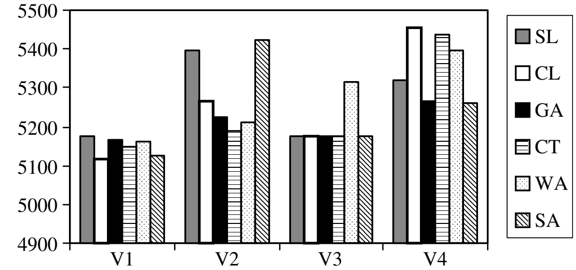

Fig. 7. CLRP average results (after step 4).

publication year and the number of customers and potential DCs. The lower bound was obtained with a relaxed 2-index integer linear programming formulation (Barreto, 2004) and the upper bound is the best known solution (BKS) obtained using the heuristic. The values followed by an asterisk correspond to the optimal solution. It was not possible to get LBs for instances 13, 14 and 16. For all the instances, the Gap falls between a minimum of 0% and a maximum of 19.01% with an average of 4.81% and a median of 3.13%.

7. Conclusions

Due to the complexity of the capacitated location-routing problem (CLRP), the heuristic approach is a promising way to find good solutions for medium and large problems. In this paper, a cluster analysis based sequential heuristic that uses simple procedures was presented. Moreover, four grouping techniques (hierarchical and non-hierarchical) and six proximity measures (single linkage, complete linkage, group average, centroid, ward and saving) were used to obtain several versions of the heuristic. Computational tests on 19 CLRP instances adapted from the literature were performed in order to compare their performance. The obtained results seem to indicate that version V1 had a slightly better performance than the other versions; version V3 was in second place followed by version V2, while version V4 had the worst performance.

Concerning proximity measures, the group average measure has produced the most balanced results, followed by the measures centroid, saving, complete linkage, single linkage and ward. However, all the measures obtained good results for some instances; thus, it seems an advisable strategy to use several versions of the proposed heuristic and several proximity measures and then choose the best solution.

In absolute terms, the average Gap was 4.81% with a median of 3.13% and, in some instances, optimality was reached.

Several authors used clustering techniques in algorithms to solve the CLRP but never justified its use nor evaluated its performance. In this work the good performance of the clustering procedures is demonstrated; the average gap is narrow (less than 5%) and the result improves when the median (≈3%) is considered, attenuating the previous influence of the two outliers (19.01 and 12.78) which are due to the LB solution and not to (UB) heuristic solution.

Despite these encouraging results, there are yet many opportunities ahead. The work is in progress, and eventually improvements are expected. Not all the clustering methods neither the proximity measures were tested and this is an open opportunity. In this process the construction of groups (cluster) with limited capacity is a crucial moment and it is necessary to improve the methods to perform it.

For instance, further improvement could be obtained in the sequential heuristic for the CLRP using the following additional steps:

Step 5. Determine the solution of the TSP in each route (including the DC).

Step 6. Apply a 1-optimal procedure to the customers, subjected to the capacity constraints. Return to step 5 if any route was changed.

Moreover, in the non-hierarchical case, another possibility to obtain better solutions is to run the same version several times using different customer sources.

All the above ideas attempt to improve the solution concerning cost; nevertheless, the choice of a final solution does not solely depend on its cost. Other objectives and aspirations of managers can also be relevant. In this context, the possibility to generate several "good" solutions in an acceptable time can be very important in order to support decisions.

This study confirms the potentialities of using clustering techniques in the CLRP approach. The grouping procedures are very fast and allow obtaining good alternative solutions corresponding to different configurations. This is a good example of an investigation opportunity, with promising results, using two distinct scientific areas, cluster analysis and operational research.

Acknowledgement

This research was partially financially supported by the European Social Fund.

UNIÃO EUROPEIA
Fundo Social Europeu

References

Anderberg, M.R., 1973. Cluster Analysis for Applications. Academic Press, New York.

Balakrishnan, A., Ward, J.E., Wong, R.T., 1987. Integrated facility location and vehicle routing models: Recent work and future prospects. American Journal of Mathematical Sciences 7 (1&2), 35–61.

Barreto, S.S., 2003. Available from: <http://sweet.ua.pt/~iscf143/>.

Barreto, S.S., 2004. Análise e modelização de problemas de localização-distribuição [Analysis and modelling of location-routing problems]. PhD Dissertation, University of Aveiro, Aveiro, Portugal (in Portuguese).

Barreto, S.S., Ferreira, C.M., Paixão, J.M., 2003a. Using clustering analysis in a capacitated location-routing problem. Communication presented at XIV Meeting of the European Working Group on Locational Analysis, September 11–13, Corfu, Greece.

Barreto, S.S., Ferreira, C.M., Paixão, J.M., 2003b. Problemas de localização-distribuição: uma revisão bibliográfica [Location-routing problems: a bibliographical review]. In: Proceedings of the VI Congreso Galego de Estatística e Investigación de Operacións, November 5–7, University of Vigo, Vigo, Spain, pp. 93–100 (in Portuguese).

Bayne, C.K., Beauchamp, J.J., Begovich, C.L., Kane, V.E., 1980. Monte carlo comparisons of selected clustering procedures. Pattern Recognition 12, 51–62.

Bodin, L.D., 1975. A taxonomic structure for vehicle routing and scheduling problems. Computers and Urban Society 1, 11–29.

Bramel, J., Simchi-Levi, D., 1997. The Logic of Logistics. Springer-Verlag, New York.

Branco, I.M., Coelho, J.D., 1990. The hamiltonian *p*-median problem. European Journal of Operational Research 47, 86–95.

Bruns, A., Klose, A., 1995. An iterative heuristic for location-routing problems based on clustering. In: Proceedings of the Second International Workshop on Distribution Logistics, The Netherlands, pp. 1–6.

Christofides, N., Eilon, S., 1969. An algorithm for the vehicle-dispatching problem. Operational Research Quarterly 20 (3), 309–318.

Clark, G., Wright, J.W., 1964. Scheduling of vehicles from a central depot to a number of delivery points. Operations Research 14, 568–581.

Cullen, F.H., Jarvis, J.J., Ratliff, H.D., 1981. Set partitioning based heuristics for interactive routing. Networks 11, 125–143.

Dantzig, G.B., Ramser, J.H., 1959. The truck dispatching problem. Management Science 6 (1), 80–91.

Daskin, M.S., 1995. Network and Discrete Location: Models, Algorithms and Applications. John Wiley & Sons, Inc., New York.

Edelbrock, C., McLaughlin, B., 1980. Hierarchical cluster analysis of intraclass correlations: A mixture model study. Multivariate Behavioral Research 15, 299–318.

Everitt, B.S., 1993. Cluster Analysis. Arnold, London.

Fukunaga, K., Short, R.D., 1978. Generalized clustering for problem localization. IEEE Transactions on Computers C 27 (2), 176–181.

Gaskell, T.J., 1967. Bases for vehicle fleet scheduling. Operational Research Quarterly 18 (3), 281–295.

Golden, R.R., Meehl, P.E., 1980. Detection of biological sex – an empirical test of cluster methods. Multivariate Behavioural Research 15, 475–496.

Gower, J.C., 1985. Measures of similarity, dissimilarity and distance. Encyclopedia of Statistical Sciences 5, 397–405.

Gower, J.C., Ross, G.J.S., 1969. Minimum spanning trees and single linkage cluster analysis. Applied Statistics 18, 54–64.

Jain, A.K., Dubes, R.C., 1988. Algorithms for Clustering Data. Prentice Hall, Englewood Clifs, New Jersey.

Johnson, S.C., 1967. Hierarchical clustering schemes. Psychometrika 32 (3), 241–254.

Kaufman, L., Rousseeum, P.J., 1990. Finding Goups in Data: An Introduction to Cluster Analysis. John Wiley & Sons Inc., New York.

Klose, A., 1995. Using clustering methods in problems of combined location and routing. Operations Research Proceedings. Springer, pp. 411–416.

Kuehn, A.A., Hamburger, M.J., 1963. A heuristic program for locating warehouses. Management Science 9 (4), 643–666.

Madsen, O.B.G., 1983. Methods for solving combined two level location-routing problems of realistic dimensions. European Journal of Operational Research 12, 295–301.

Maranzana, F.E., 1963. On the location of supply points to minimize transportation costs. IBM Systems Journal 2, 129–135.

Milligan, G.W., Schilling, D.A., 1985. Asymptotic and finite-sample characteristics of four external criterion measures. Multivariate Behavioral Research 20, 97–109.

Min, H., 1989. The multiple vehicle routing problem with simultaneous delivering and pick-up points. Transportation Research A 23 (5), 377–386.

Min, H., 1996. Consolidation terminal location–allocation and consolidated routing problems. Journal of Business Logistics 17 (2), 235–263.

Min, H., Current, J., Schilling, D., 1992. The multiple depot vehicle routing problem with backhauling. Journal of Business Logistics 13 (1), 259–288.

Or, I., 1976. Traveling salesman–type combinatorial problems and their relation to the logistics of regional blood banking. PhD Dissertation, Northwestern University, Evanston, Illinois, USA.

Perl, J., 1983. A unified warehouse location-routing analysis. PhD Dissertation, Northwestern University, Evanston, Illinois, USA.

Romesburg, H.C., 1984. Cluster Analysis for Researchers. Lifetime Learning Publications, Belmont, Canada.

Srivastava, R., 1986. Algorithms for solving the location-routing problem. PhD Dissertation, The Ohio State University.

Srivastava, R., 1993. Alternate solution procedures for the location-routing problem. OMEGA The International Journal of Management Sciences 21 (4), 497–506.

Srivastava, R., Benton, W.C., 1990. The location-routing problem: Considerations in physical distribution system design. Computers and Operations Research 17 (5), 427–435.

White Paper, 2001. European transport policy for 2010: Time to decide. The European Commission, COM(2001)370.