Heuristic Concentration and Tabu Search: A Nose to Nose Comparison

K.E. Rosing
Applied Economics and
Tinbergen Institute
Erasmus University Rotterdam
Post Box 1738
NL-3000 DR Rotterdam
The Netherlands

and

C.S. ReVelle
Department of Geography and Environmental Engineering
Johns Hopkins University
514 Ames Hall
Baltimore MD 21218-2686
USA

29 April 1997

Keywords: Combinatorial optimization; Heuristics; Integer programming; Location; Heuristic concentration; Tabu search

ABSTRACT

In 1997 two papers applying the metaheuristics Tabu Search (TS) and Heuristic Concentration (HC) to the *p*-median problem were published in consecutive volumes of the *European Journal of Operational Research*. Here we apply the method of HC some of the data sets which were used for computational experience in the paper on TS and briefly set out the results.

ACKNOWLEDGEMENTS
The authors would like to thank Erik Rolland, David Schilling and John Current for kindly allowing us to use their data.

INTRODUCTION

In volume 96 of the *European Journal of Operational Research* Rolland *et al*. (1997, hereafter referred to as RSC) published a study of the efficiency of an implementation of the metaheuristic Tabu Search (Hansen, 1986; Glover, 1989, 1990, 1993) comparing it to the heuristics developed by Densham and Rushton (1992a, 1992b) and by Goodchild and Noronha (1983). The three methods were used to solve a series of 100 differently sized *p*-median problems. In volume 97 Rosing and ReVelle (1997, hereafter referred to as RR) published a description of and computational experience with a proposed new metaheuristic, Heuristic Concentration (HC), using 90 differently sized *p*-median problems. RR used as a base heuristic in their study the Teitz and Bart (1968) vertex substitution heuristic (T&B) and applied HC to the result.

The *p*-median problem is, on a network, to choose, amongst the intersections or termini (here termed "demand nodes" and symbolised as "n") some number of centres (here termed "facilities" and symbolised by "*p*") which minimize the summation of the weighted distance. Hakimi (1964, 1965) proved that an optimal solution existed consisting of a selection of the n nodes. The work of Hakimi makes this a combinatorial problem. ReVelle and Swain (1970) present an integer linear programme (ILP) for it.

All of the five heuristics referred to above are basically interchange heuristics (Pirlot, 1992, 1996) with the exception of HC which is a *selection* heuristic.

THE TWO METAHEURISTICS

Both TS and HC are what are called metaheuristics in that they constitute a series of ideas about how to approach a problem. These ideas are a guide to the development of a specific algorithm for a specific problem.

Tabu Search

A difficulty with any interchange heuristic is that its likelihood of terminating before reaching optimality appears to increase as a function of problem size. For the T&B (1968) RR report that optimal termination correlates with increasing combinatorial space at -0.75. We are unaware of similar studies for other interchange heuristics but it seems reasonable to expect similar results. The combinatorial space, the number of possible solutions, is given by $\binom{n}{p}$ Some of these possible solutions will be sub-optimal; those that satisfy the stopping criteria of the interchange heuristic are local optima (one or perhaps more will also be the global optimum).

TS perturbs an interchange heuristic, attempting to "bounce" the algorithm out of a local optima, and then continue on towards the global optimal. It does this by employing a memory (with differing grades of sophistication) of where it has been. This memory makes specific, already investigated, interchanges illegal in the hope that a possible short-term degradation of the objective function will lead to an uninvestigated region of the solution space and hence to further improvement of the objective function. This has been termed "steepest ascent, mildest descent" ([in a maximization problem] Hansen, 1986). Full details of the metaheuristic can be found in Glover (1986, 1989, 1990, 1993) and details of its implementation in the *p*-median context can be found in RSC.

Heuristic Concentration

The development of HC is a result of the observation that different random trials of an interchange heuristic generally give solutions that are highly similar in the specific demand nodes selected to be facilities. Viewed differently, the vast majority of demand nodes are never selected to be facilities. This allows the development of a concentration set (CS) as the union of the sets of facilities (each consisting of *p* nodes) found in different sub-optimal solutions. The best set of facilities is then extracted from the CS by means of an ILP. Another observation is that a number of demand nodes are frequently selected as facilities in **all** the different sub-optimal solutions. This allows the partitioning of the CS into two sets -- the CS free (CS_f) and the CS open (CS_o). The CS_o contains those nodes which appear in all solutions. It is assumed that they really are components of the optimal solution and they are fixed open. The remaining nodes in the CS are available to be chosen or not chosen; they are free. Thus, this set is termed the CS_f. Two ILPs can be written, one operating on the CS (ILP-1) and one operating on the $\{CS_o, CS_f\}$ (ILP-2). ILP-2 is much smaller and thus much faster. It includes however an assumption (certain nodes are facilities) which makes it slightly less likely to terminate optimally. Finally if no nodes are in all sub-optimal solutions then the ILP-2 does not exist. In this comparison the ILP-2 has been used in all cases except those (generally those with small values of *p*) where it does not exist. More details and explanation can be found in RR.

THE DATA SETS

RSC tested their TS procedure on 20 different of data sets of sizes ranging from 13 to 500 demand nodes. A number of different values of *p* were utilized with each data set yielding 100 problems. Since: 1). HC is designed for larger data sets where optimality is less likely to be achieved by an interchange heuristic (Rosing, 1997) and 2). RSC's TS heuristic seemed less effective on the larger (value of n) data sets[1] we requested the opportunity to re-examine their data where n ≥ 100 with HC. RSC responded most helpfully e-mailing the n = 100, 200, 300, 400, and 500 node data sets immediately.

It was now that problems began to develop. The weights for the networks with n = 300, 400, 500 were damaged in the e-mail transmission. The computational results reported here had been obtained before new sets of weights arrived. Difficulties in reconstructing data matrices from "forward star" format were compounded by the fear that these large networks would be unsolvable because of excessive branch and bound we had encountered. Accordingly we limit ourselves, for this comparison, to the n = 100 and 200 data sets. In order to confirm that these were the correct data sets (and that we had interpreted them correctly) we attempted to solve the 21 available instances optimally using an earlier and reliable ILP formulation/matrix generator (Rosing *et al.*, 1979). For n = 100 our optimal solutions were identical to the optimal which RSC report in their Table 1 with the exception of 100a, p = 5. RSC report 59962 as the ILP optimal. This is the first feasible we encounter in branch and bound but the global optimal is 57708. Finding the same optimal solutions confirmed we were using the correct data. For n = 200 we were unable to find optimal solutions for comparison purposes. After 4.5 days of trying to resolve fractions in the n = 200, p = 20 problem (3000 nodes resolved) we were forced to terminate the programme. T&B was run 50 times with each different value of *p*. Several of the local optima found by

---

1 In their study 90.7% of instances with n ≤ 50 were optimal, 78.6% of instances with n ≤ 70, but only 46.7% of instances with n ≥ 100 reach the optimal (or best found for data sets ≥ 200).

T&B were identical to entries in RSC's Table 2 (n = 200). In addition in one case (n = 200, *p* = 15) HC found an identical best known. We take this as sufficient confirmation that our interpretation of this data set is also correct.

One of the nice things about the *p*-median is its integer friendliness (ReVelle, 1993). We are aware of only one formal study showing the strong integer characteristics (Morris, 1978) of the *p*-median; but references in the literature to the paucity of fractional solutions and their ease of resolution are too frequent to enumerate. This data is different however. The LP for all 15 of the n = 100 problems terminated with fractions that had to be resolved by branch and bound. Details are shown in Table 1. The tableau of the fully specified 100 node problem is always 10001 rows and 10000 columns. On the left of Table 1 is the number of branch and bound nodes that had to be resolved to prove optimality together with the time in branch and bound and the total time to solve the problem. On the right are similar statistics for HC (we shall return to this portion of the Table presently).

Concerning the data RSC tell us: "… where each node served as both a demand point and a potential facility site. The graphs were randomly generated in a 100 x 100 square with demand at each node being a randomly generated integer distributed in the range of 0 - 100." From this and from inspection of the data it would appear that the distances are random numbers, in the range 1 - 100 assigned to i, j pairs. There are no zero distances but there are zero weights. This means that in the weighted distance matrix there are whole rows which involve no cost. These demand nodes can assign anywhere at no cost. It must be randomness and the lack of structure to the data which is responsible for all the complexly fractional solutions and the difficulty in resolving the fractions.

THE EXPERIMENT

For the 100 node data sets (optimal known) the T&B heuristic was run 50 times. The total time for 50 runs, in seconds, is given in Table 2, column "Heur." The runs were then sorted into ascending order by functional value and the lists inspected. Since we are trying to judge the effectiveness of HC and not T&B in those cases (mostly small values of *p*) where the optimal solution had been found by T&B the optimal solutions were eliminated from the list of solutions. HC will never terminate at less than the best solution in the list. In order to make our comparison of TS and HC fair the we must insure that the information available to the ILP consists only of sub-optimal solutions. The first (best) 15 different sub-optimal solutions were then used to build the CS for ILP-2. The ILP-2 model fixes open ($X_{jj} = 1$) any facilities which were in all 15 solutions (the CS_o). The nodes which were in some but not all solutions constitute the CS_f and may be chosen to be facilities or not chosen. In some cases with smaller values of *p* the CS_o was empty. Then, automatically, model ILP-1 (CS model) is used. Which model was used is shown in Table 1, the in column headed "ILP." The time, in seconds, to create the MPS standard input is given in Table 2, the column headed "HCcon." The model (ILP-1 or ILP-2) was then solved using the Cplex (1995) 4.0. The solution time is given in Table 2, column "Cplex" and the total time for the three steps is given in the column "Total." A Sun Sparcserver 20 (60 Mhz) was used for all calculations.

In the case of the 200 node data sets a slightly different procedure had to be followed since the optimal solutions are not known. There are three different values of *p*. In the problem with $p = 10$ T&B found a better solution than the "best known" shown in RSC's Table 2. The best 15 (including the new "best known") were used to construct the CS and the model solved. The solution was identical to the new "best known" from the T&B. The new "best known" was then eliminated from the 50 solutions (analogous to removing optimal solutions from the solution list in the n = 100 problems). A new CS was created without this knowledge and the problem was solved again. The same new "best known" solution was again found by HC. In the other two cases the best result from the T&B were inferior to the "best known" reported by RSC. One case (p = 20) the solution to the ILP was better than RSC's "best known" and in the other (p = 15) equal.

Comparison of the "B&B seconds" of Table 1, right-hand half and the "Cplex" of Table 2 indicates that the resolution of fractions accounts for the excessive time required when, particularly, model ILP-1 was employed.

THE RESULTS

Table 1, the right, shows in the column "ILP" which model ILP-1 or ILP-2 was used. ILP-1 is used when no single node was selected in the solution of all of the best 15 of the sub-optimal solutions from the T&B. The extreme difference in the size of the matrices with ILP-1 and ILP-2 is shown in the columns giving the size of the matrix and labelled "Rows" and "Cols." The number of branch and bound nodes resolved and the number of seconds the step took complete are also show in Table 1. Comparison of the columns "B&B nodes" and "B&B seconds" on the left and the right half of the table also indicates that the use of HC concentrates not only the solution but also the work required to resolve fractions.

Table 2 is completed by showing the time taken by TS as reported in Tables 1 and 2 of RSC. Table 3 give the optimal (or "best known") functional values and the gap, defined as (heuristic_solution minus optimal_solution[2]) divided by optimal_solution[2], for first TS and

---

2 or best_known in the case of the 200 node data sets.

second HC for the 21 problems. The new optimal or best known values are reported in **bold** type as are all 0.0% gaps.

REFERENCES

Cplex (1995) *Using the Cplex Callable Library* Incline Village, NV: Cplex Optimization.

Densham, P.J. and Rushton G. (1992a) "Strategies for Solving Large Location-Allocation Problems by Heuristic Methods" *Environment and Planning, Series A* **24** 289-304.

Densham, P.J. and Rushton G. (1992b) "A More Efficient Heuristic for Solving Large *P*-Median Problems *Papers in Regional Science* **71** 307-329.

Glover, F. (1986) "Future Paths for Integer Programming and Links to Artificial Intelligence" *Computers and Operations Research* **5** 533-549.

Glover, F. (1989) "Tabu search - Part I" *ORSA Journal of Computing* **1/3**.

Glover, F. (1990) "Tabu search - Part II" *ORSA Journal of Computing* **2/1**.

Glover, F. and Laguna, M.(1993) "Tabu search" in: C. Reeves (ed.) *Modern Heuristic Techniques for Combinatorial Problems* (London: Blackwell) 70-150.

Goodchild, M.F. and Noronha, V. (1983) *Location-Allocation for Small Computers* Monograph No. 8, Iowa City, Iowa: Department of Geography, University of Iowa.

Hansen, P. (1986) "The Steepest Ascent Mildest Descent Heuristic for Combinatorial Programming" Paper presented at: *Congress on Numerical Methods in Combinatorial Optimization* Capri, Italy.

Hakimi, S.L. (1964) "Optimum Location of Switching Centers and the Absolute Centers and Medians of a Graph *Operations Research* **12** 450-459.

Hakimi, S.L. (1965) "Optimal Distribution of Switching Centers in a Communication Network and Some Related Graph Theoretic Problems" *Operations Research* **13** 462-475.

Pirlot, M. (1992) "General Local Search Heuristics in Combinatorial Optimization: A Tutorial" *JORBEL - Belgian Journal of Operations Research, Statistics and Computer Science* **32** 7-67.

Pirlot, M. (1996) "General Local Search Methods" *European Journal of Operational Research* **92** 493-511.

Morris, J.D. (1978) "On the Extent to Which Certain Fixed Charge Depot Location Problems can be Solved by LP" *Journal of the Operational Research Society* **29** 71-76.

Rolland, E., Schilling, D.A. and Current, J.R. (1997) "An Efficient Tabu Search Procedure for the *P*-Median Problem" *European Journal of Operational Research* **96**: 329-342.

ReVelle, C.S. (1993) "Facility Siting and Integer-Friendly Programming" *European Journal of Operational Research* **65** 147-158.

ReVelle, C.S. and Swain, R. (1970) "Central Facilities Location" *Geographical Analysis* **2** 30-42.

Rosing, K.E. (1997) "An Empirical Investigation of the Power of a Vertex Substitution Heuristic" *Environment and Planning, Series B* **24** 59-67.

Rosing K.E., ReVelle, C.S. and H. Rosing-Vogelaar (1979) "The *P*-Median and its Linear Programming Relaxation: An Approach to Large Problems" *Journal of the Operational Research Society* **30** 815-823.

Rosing, K.E. and ReVelle, C.S. (1997) "Heuristic Concentration: Two Stage Solution Construction" *European Journal of Operational Research* **97**: 75-86.

Teitz, M.B. and Bart, P. (1968) "Heuristic Methods for Estimating the Generalized Vertex Median of a Weighted Graph" *Operations Research* **16** 955-961.

Table 1, Problem Sizes and the Amount of Branch and Bound

	Full (10001 x 10000)			Heuristic Concentration				
	B&B	B&B	Total				B&B	B&B
p=	nodes	seconds	seconds	ILP	Rows	Cols.	nodes	seconds
			Data 100a					
5	216	3816.3	4021.5	1	3071	3000	88	179.2
7	446	5108.2	5298.7	1	3269	3200	217	320.8
8	240	2672.9	2830.6	2	1802	1737	66	55.5
10	272	2113.7	2253.7	2	1595	1526	52	24.5
13	75	463.8	567.3	2	1867	1803	37	13.7
15	26	78.6	153.2	2	1332	1271	11	2.7
17	18	68.3	143.2	2	1075	1017	6	0.9
20	18	46.7	89.7	2	1054	1003	12	0.9
			Data 100b					
5	60	1263.9	1487.5	1	2873	2800	88	168.8
7	164	1990.6	2156.5	1	3269	3200	103	185.8
10	120	956.3	1074.0	1	3863	3800	86	127.4
13	80	383.4	480.6	1	3962	3900	58	62.8
15	14	90.0	183.7	2	1131	1072	32	4.8
17	19	63.2	143.6	2	960	901	17	1.4
20	14	40.4	91.7	2	846	781	8	0.4
			Data 100c					
10	84	890.9	1021.9	1	3863	3800	59	124.2
15	331	1420.3	1418.5	1	4655	4600	128	204.2
20	112	338.9	405.9	2	1718	1667	119	26.5
			Data 200					
10				2	3923	3760	255	832.8
15				2	5770	5620	503	2796.1
20				2	3996	3857	1251	2490.8

Table 2, Times required.

p=	Heur.	HCcon	Cplex	Total	Tabu
		Data 100a			
5	6.0	2.1	208.0	216.1	42.9
7	6.9	2.2	342.6	351.7	45.1
8	7.8	1.5	65.4	74.7	46.7
10	9.5	1.4	30.9	41.8	48.4
13	11.5	1.7	19.9	33.1	51.5
15	13.0	1.3	6.8	21.1	53.4
17	13.9	1.3	3.8	19.0	52.6
20	15.9	1.2	3.2	20.3	58.2
		Data 100b			
5	5.8	2.1	194.7	202.6	42.7
7	7.7	2.2	210.1	220.0	45.4
10	10.8	2.7	147.3	160.8	48.8
13	12.1	3.0	62.8	77.9	51.6
15	14.1	1.2	7.1	22.4	53.3
17	15.3	1.1	3.1	19.5	56.2
20	16.7	1.1	1.7	19.5	54.5
		Data 100c			
10	10.1	2.5	149.0	149.0	54.3
15	14.3	3.2	204.2	221.7	58.1
20	16.7	1.4	30.9	49.0	69.9
		Data 200			
10	43.3	3.5	876.7	922.5	446.6
15	62.9	4.6	2903.4	2970.9	497.5
20	81.1	3.9	2521.2	2606.2	544.1

All times in seconds

Table 3, Optimal (Best Known) Functional Values and Gap.

p=	Optimal Value	Tabu Gap	HC Gap
		Data 100a	
5	57708	3.91%	0.00%
7	39363	1.47%	1.19%
8	32461	0.06%	0.06%
10	24159	0.35%	0.00%
13	16948	0.49%	0.00%
15	14203	2.54%	0.00%
17	12379	0.98%	0.00%
20	10365	0.73%	0.00%
		Data 100b	
5	41942	6.01%	4.55%
7	29014	3.04%	0.00%
10	18797	1.40%	0.00%
13	13993	0.00%	0.00%
15	11621	2.00%	1.69%
17	10198	1.07%	0.00%
20	8478	0.53%	0.00%
		Data 100c	
10	18699	0.00%	0.00%
15	11860	0.40%	0.00%
20	8347	1.83%	0.00%
		Data 200	
	Best Known		
10	48912	0.68%	0.00%
15	31153	2.80%	0.00%
20	23475	0.09%	0.00%