MIC'99

III Metaheuristics International Conference

organized by

PUC-Rio - Catholic University of Rio de Janeiro

Angra dos Reis, Brazil

July 19-22, 1999

Table of contents

TABU SEARCH APPLICATION FOR POINT FEATURES CARTOGRAPHIC LABEL PLACEMENT PROBLEM

MISSAE YAMAMOTO[1] , LUIZ ANTÔNIO NOGUEIRA LORENA[1] AND GILBERTO CÂMARA NETO[1]

[1]Brazilian Institute of Space Research (INPE)
C.P. 515 - CEP 12201-970 - São José dos Campos - SP - Brazil
E-mails: `missae@dpi.inpe.br`, `gilberto@dpi.inpe.br` and `lorena@lac.inpe.br`

Abstract: The generation of better label placement configurations in maps is a problem that comes up in automated cartographic production. The objective of a good label placement is to show the geographic position of the features with their corresponding texts clearly, respecting the cartographic conventions, with an esthetic and a harmonious quality when presenting the information. We approached the label placement problem from a combinatorial optimization point of view. Our research consisted in the evaluation of a tabu search (TS) optimization algorithm applied to cartographic label placement. The TS implemented in SCARTA, a cartographic production software, in development by the Image Processing Division (DPI/INPE), proved to be an efficient algorithm, in real and random test cases. When compared with alternative techniques such as simulated annealing, genetic algorithm (GA)with mask and others described in literature, the TS had the best performance in quality. We concluded that TS is a recommended method to solve cartographic label placement problem of point features, due to its simplicity, practicality, efficiency and good performance along with its ability to generate quality solutions in acceptable computational time.

1 Introduction

Cartographic label placement refers to the name insertion process in maps. This process has showed one of the most challenging problems for the geographical information systems (GIS), computational systems that manipulate geographical data. Positioning the names requires that overlap among names be avoided, that cartographic conventions and preference be obeyed, that unambiguous association be achieved between each name and its corresponding feature and that a high level of harmony and quality be achieved.

In cartography, three different label-placement types are identified: labeling of point features (cities, schools, hospital, mountain peaks ...), line features (rivers, roads, ...), and area features (countries, states, oceans, ...). In this article we are only concerned with the placement of labels for point features that can be thought of as a combinatorial optimization problem. Certain aspects such as potential label positions and cartographic preference for placing labels must be defined. Potential label positions of each point feature is the set of potential positions where the label may be placed. Figure 1.1 shows a set of four potential label positions for a point feature. Each box indicates a region in which a label may be placed and the value inside each box corresponds to the order of preference for placing a label. More desirable positions are indicated for lower values.

The label placement is a combinatorial optimization problem of difficult solution and [7] showed that the point features label placement (PFLP) problem is NP-hard. Accordingly, we have need of heuristics and metaheuristics that do not seek the exact solution, but a sufficiently good solution in cost terms.

Several heuristics and metaheuristics have been used for years to solve the PFLP problem, such as exhaustive search, greedy algorithms, discrete gradient descent, Hirsch's algorithm [4], Lagrangean

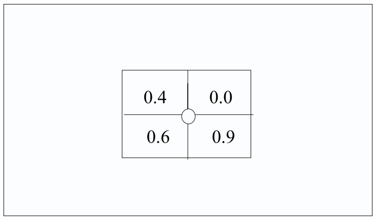

Figure 1.1 : A set of potential label positions and their cartographic preference

relaxation [12, 13], Simulated Annealing [2, 1, 9] and others. They are reviewed by [2]. A GA with mask is described in [10]. In [11], some of the techniques above were reviewed and a new algorithm was proposed to solve the problem. In this article we propose a TS algorithm to solve the PFLP problem. The rest of the article is described as follows. In Section 2 we introduce the TS optimization algorithm to solve the PFLP problem. In Section 3 we present and discuss the results obtained of TS application in real data and in a standard set of randomly generated points suggested in literature. The comparison of the TS algorithm with the other algorithms described in the literature is also presented in this section. Our conclusions are presented in Section 4.

2 TABU SEARCH FOR PFLP

TS is a heuristic procedure proposed by Fred Glover to solve combinatorial optimization problems. The basic idea is avoiding that the search for best solutions stops when a local optimal solution is found [3, 6].

The algorithm used to solve label placement problem calculates conflicts in two ways (Figure 2.1): Number of overlappings (NOVER) = 4 (P1=0, P 2=1, P 3=2, P 4=1) and Number of labels that overlap (NLOVER) = 3. The NOVER calculate the conflicts of each point. It is used to choose neighborhood points and to select next candidate points. The NLOVER calculate the number of overlapping on layout, to be used on the computation of neighborhood and tabu list sizes. It is the number of overlapping labels of final layout that is showed to the user.

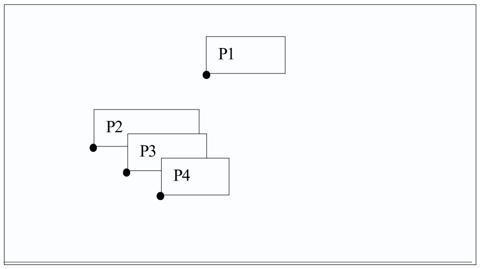

Figure 2.1 : Conflict evaluation

The TS implementation involves seven components. They are resumed in Table 2.1. The chosen initial configuration was to label each point with its best cartographic preference position. The code was implemented using a Sun Sparc 20 workstation, UNIX solaris version 2.5, C++ compiler version 4.0.1, GIS SCARTA and SPRING version 3.0 [8].

Table 2.1: TS implementation

TS component	**Short description**
Pre-computation	compute all potential overlappings between label positions.
Objective function	Min $\sum_{i=1}^{np}(\alpha_1 conflict(i) + \alpha_2 preference(i))$ where: np is number of points; conflict(i) = NOVER of the point i; preference(i) = cartographic preference of the active labels in the point i; α_1 = weight for overlapping labels; α_2 = weight for cartographic preference.
Neighborhood	Pairs (Pi,Li), where Pi are points having large costs, and Li their corresponding active label.
Configuration changes	Swap of active labels of points in neighborhood. Fixing the configuration of large decrease on objective function.
Tabu list	Last visited points and active labels.
Aspiration criteria	By objective in global form and by default.
Long term memory	Normalized frequency for the number of times a point is visited. Updated after 50 consecutive iterations. Costs in points penalized with normalized frequency(i) (on neighborhood): α_1 conflict(i) + α_2 preference(i) + normalized frequency(i)

3 Results

3.1 Real problem

In order to verify the TS algorithm performance in regard to real dataset - natural distribution of points - we used one dataset available in [5]. The set consist of 128 point features from regions of the USA map. Each text label varies in length depending on the name of the cities it represents, turning then the test more realistic. The considered area in geographical coordinates is: longitude (O 123 0 0 O 73 0 0), latitude (N 24 0 0 N 51 0 0) and projection LAMBERT/HAYFORD.

We made tests using different values of α_1 (that handle the level of consideration of the NOVER), α_2 (that handle the level of consideration of the cartographic preference), character height and scale. The parameters considered for tests are: tabu list size = 7 + INT (0.25 * NLOVER), neighborhood size = 1 + INT (0.05 * NLOVER), number of iterations for recalculations = 50, potential label positions = 8 and maximum number of iterations allowed = 5000. The results of the tests are reported on Table 3.2.

Table 3.2: RESULTS OF THE TESTS

layout	**USA1**	**USA2**	**USA3**	**USA4**	**USA5**	**USA6**
$\alpha 1$	1	1	1	3	3	3
$\alpha 2$	10	5	1	1	1	1
character height	1 mm	1 mm	1 mm	1 mm	1.5 mm	1.5 mm
scale	1:23000000	1:23000000	1:23000000	1:23000000	1:23000000	1:20000000
iteration	324	140	28	28	583	4812
NLOVER	26	4	0	0	6	2

When the cartographic preference weight is large regarding the weight of the NOVER, the NLOVER is large, but it is possible to verify that the labels try to occupy the best available cartographic position. On the other hand, as cartographic preference weight decreases and the weight for NOVER increases, the NLOVER decreases, but it is possible to see that the labels occupy any one of potential positions.

Next, we made tests for 1.0 mm (USA4) and 1.5 mm (USA5) character height. The results showed that the character size affected considerably the final layout of the map (Table 3.2). Finally we verified that the scale affects considerably the final layout of the map. Tests for scales 1:23000000 (USA5) and 1: 20000000 (USA6) are on the Table 3.2.

3.2 Comparative analysis

The works [2] and [10] compared several algorithms using standard sets of randomly generated points: grid size of 792 by 612 units, fixed size label of 30 by 7 units and page size of 11 by 8.5 inch.

We used the standard sets of randomly generated points and simulated the same conditions as described by [2] to compare our TS algorithm with the others algorithms of the literature: number of the points: n = 100, 250, 500, 750, 1000; for each problem size n, we generate 25 different configurations with random placement of point feature using different seeds; for each problem size n, we calculate the average percentage of labels placed without conflict of the 25 trials; no penalty was attributed for label that extended beyond the boundary of the region; 4 potential label positions were considered; cartographic preference was not considered; without point selection (we do not allow to omit certain points and their labels even trough it is impossible to solve overplots); the parameters used for tabu search are: tabu list size = 7 + INT (0.25 * NLOVER), neighborhood size = 1 + INT (0.05 * NLOVER) and number of iterations for recalculation = 50.

We have transcribed the results of other PFLP algorithms from [10] to Table 3.3, and also the average of percentages of labels placed without conflict got by our TS algorithm. Regarding the optimization algorithms of the literature, the tabu search showed superior results. In Table 3.3 the columns show the average of the percentage of labels placed without conflict for 100, 250, 500, 750 and 1000 points, considering different algorithms of the literature. The lines show the average of the percentage of labels placed without conflict got by the optimization algorithms tested on [2] (random placement, greedy-depth first, gradient descent, 2-opt gradiente descent, 3-opt gradient descent, Hirsch, Zoraster and simulated annealing), on [10] (GA without masking and GA with masking) and on [11] the tabu search.

4 Conclusion

The PFLP is a problem of practical importance for geoprocessing and automated cartography. One of our purposes is to evaluate a TS optimization algorithm applied to the PFLP problem. The computational tests using the standard set of randomly generated points with the same conditions as described by [2] and [10] showed that TS had the best performance in quality. The other purpose was to evaluate TS algorithm performance with regard to real dataset in order to verify its behavior regarding the natural occurrence point feature distributions. The results showed that higher the level of consideration for cartographic preference, larger is the algorithm difficult to get the state of no conflict among labels, but at the same time, the label can reach best cartography positions. With regard to text character height, the algorithm showed quite sensible and it was possible to verify that text character size affect considerably the final quality of the map. Finally, it is possible to say that the TS performance is good, even if applied to real datasets with natural clustering of the point features distributions with label of variable length.

Table 3.3: PFLP ALGORITHMS RESULTS, USING A STANDARD SET OF RANDOMLY GENERATED DATASETS (adapted from [10], page 273)

algorithms	**100 points**	**250 points**	**500 points**	**750 points**	**1000 points**
Tabu Search	100.00	100.00	99.26	96.76	90.00
GA with masking	100.00	99.98	98.79	95.99	88.96
GA without masking	100.00	98.40	92.59	82.38	65.70
Simulated Annealing	100.00	99.90	98.30	92.30	82.09
Zoraster	100.00	99.79	96.21	79.78	53.06
Hirsch	100.00	99.58	95.70	82.04	60.24
3-Opt Gradient Descent	100.00	99.76	97.34	89.44	77.83
2-Opt Gradient Descent	100.00	99.36	95.62	85.60	73.37
Gradient Descent	98.64	95.47	86.46	72.40	58.29
Greedy-depth First	95.12	88.82	75.15	58.57	43.41
Random Placement	84.56	65.63	44.06	29.06	19.53

Bibliography

[1] J. Christensen, J. Marks, and S. Shieber. *Graphics Gems IV*, chapter Placing Text Labels on Maps and Diagrams. London, Academic Press, Cambridge, Mass., 1994.

[2] J. Christensen, J. Marks, and S. Shieber. An empirical study of algorithms for point-feature label placement. *ACM Transactions on Graphics*, 14:203–232, 1995.

[3] F. Glover and M. Laguna. *Tabu Search.* Kluwer Publishers, Boston, 1997.

[4] S. A. Hirsch. An algorithm for automatic name placement around point data. *American Cartographer*, 9:5–17, 1982.

[5] D. E. Knuth. *The stanford graphBase, a platform for combinatorial computing.* Addison-Wesley, New York, 1993.

[6] M. Laguna. A guide to implementing tabu search. *Investigacin Operativa*, 4:5–25, 1994.

[7] J. Marks and S. Shieber. The computational complexity of cartographic label placement. Technical report, Center for Research in Computing Technology, Harvard University, 1993.

[8] Brazilian Institute of Space Research (INPE)/Division of Digital Image Processing (DPI). Spring -inpe's integrated image processing and geographic information system. http://sputnik.dpi.inpe.br/spring/, 1998.

[9] E. Shawn, J. Christensen, J. Marks, and S. Shieber. A general cartographic labeling algorithm. *Cartographica*, 33:111–999, 1997.

[10] O. V. Verner, R. L. Wainwright, and D. A. Schoenefeld. Placing text labels on maps and diagrams using genetic algorithms with masking. *INFORMS Journal on Computing*, 9:266–275, 1997.

[11] M. Yamamoto. Tabu search application for point features cartographic label placement problem. Master's thesis, Applied computing - Brazilian Institute of Space Research (INPE), São José dos Campos, SP, Brazil, 1998. (in portuguese).

[12] S. Zoraster. Integer programming applied to the map label placement problem. *Cartographica*, 23:16–27, 1986.

[13] S. Zoraster. The solution of large 0-1 integer programming problems encountered in automated cartography. *Operations Research*, 38:752–759, 1990.