

MITSUBISHI ELECTRIC RESEARCH LABORATORIES
<http://www.merl.com>

Empirical Testing of Algorithms for Variable-Sized Label Placement

Jon Christensen, Stacy Friedman, Joe Marks, Stuart Shieber

TR97-13 December 1997

Abstract

We report an empirical comparison of different heuristic techniques for variable-sized point-feature label placement.

This work may not be copied or reproduced in whole or in part for any commercial purpose. Permission to copy in whole or in part without payment of fee is granted for nonprofit educational and research purposes provided that all such whole or partial copies include the following: a notice that such copying is by permission of Mitsubishi Electric Research Laboratories, Inc.; an acknowledgment of the authors and individual contributions to the work; and all applicable portions of the copyright notice. Copying, reproduction, or republishing for any other purpose shall require a license with payment of fee to Mitsubishi Electric Research Laboratories, Inc. All rights reserved.

Copyright © Mitsubishi Electric Research Laboratories, Inc., 1997
201 Broadway, Cambridge, Massachusetts 02139

MERL – A MITSUBISHI ELECTRIC RESEARCH LABORATORY
<http://www.merl.com>

Empirical Testing of Algorithms for Variable-Sized Label Placement

Jon Christensen	Stacy Friedman
Painted Word, Inc.	Oracle Corporation
Joe Marks	Stuart Shieber
MERL	Harvard University

TR-97-13 October 1997

Abstract

We report an empirical comparison of different heuristic techniques for variable-sized point-feature label placement.

This work may not be copied or reproduced in whole or in part for any commercial purpose. Permission to copy in whole or in part without payment of fee is granted for nonprofit educational and research purposes provided that all such whole or partial copies include the following: a notice that such copying is by permission of Mitsubishi Electric Information Technology Center America; an acknowledgment of the authors and individual contributions to the work; and all applicable portions of the copyright notice. Copying, reproduction, or republishing for any other purpose shall require a license with payment of fee to Mitsubishi Electric Information Technology Center America. All rights reserved.

Copyright © Mitsubishi Electric Information Technology Center America, 1997
201 Broadway, Cambridge, Massachusetts 02139

1. First printing, TR97-13, October 1997. Also in *Proc. of the 13th Annual Symposium on Computational Geometry*, Nice, France, June 1997, pp. 415-417.

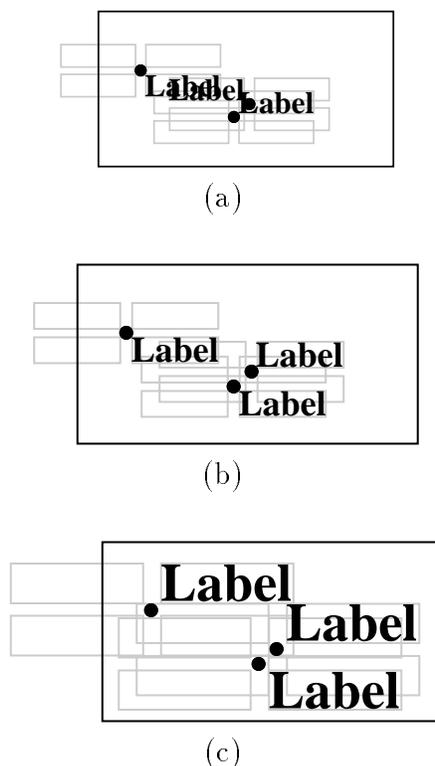


Figure 1: Examples of fixed- and variable-size label placement. A map of three points (a) shows the four potential label positions for each point and depicting an especially poor choice of label positions. Viewed as a fixed-size label placement problem, the points can be labeled well as in (b). Viewed as a variable-size label placement problem, the size of the labels may be increased by some 40 percent and still labeled without overlap (c) by moving the label of the upper left point.

The development of methods for placing labels on map features is a central problem in automated cartography. A well-studied problem of this sort is the *point-feature label placement* (PFLP) problem: Given a set of points and a set of candidate label positions adjacent to each point find the choice of label position for each point that minimizes the total number of label-label and label-point overlaps. Sample labelings of this sort are depicted in Figure 1. This NP-hard problem has been attacked by a number of researchers. Thorough empirical testing has shown that a heuristic algorithm based on simulated annealing (SA) outperforms all previously published practical algorithms for this problem [1].

Recently, Wagner and Wolff [3] have explored a variant of the PFLP problem. In their problem, which we call *variable-sized point-feature label placement* (VPFLP), the goal is to find the largest label scale at which the set of points can be labeled with zero overlaps; that

is, only perfect labelings are allowed, but the size of the labels is allowed to vary uniformly. (See Figure 1.) Wagner and Wolff present an algorithm for the specific case of VPFLP in which each point has four candidate label positions (4-VPFLP). Wagner and Wolff do not compare their own algorithm for this problem with those of other researchers. In this note, we report on empirical testing of Wagner and Wolff’s algorithm (which we will refer to as WW) for 4-VPFLP with the SA algorithm. We find that SA performs essentially identically to WW, though it is slower. However, the generality of the SA algorithm (explored, for instance, by Edmondson et al. [2]) means that, unlike WW, SA can be applied to a far broader range of VPFLP variants. Finally, we address the question of whether the clever preprocessing method used in WW might be advantageously applied to PFLP by using it as a preprocessing step for SA, and show that it provides no advantage.

The WW algorithm for VPFLP works by performing a simple binary search over scales; at each candidate scale, a test is performed to determine whether a perfect labeling exists. To perform this *admissibility test*, any algorithm for PFLP may be used, but Wagner and Wolff take advantage of the fact that finding a good but imperfect solution to this embedded PFLP problem is profligate; they propose instead a very fast heuristic that determines quickly whether a perfect labeling exists. Their admissibility test is particularized to the case where each point has only four candidate label positions. It works by (i) a *preprocessing phase* to eliminate candidate positions that cannot occur in a perfect labeling; (ii) a *greedy heuristic phase* to eliminate all but two candidate positions for each point, thereby reducing the problem to 2-PFLP; (iii) a *solution phase* to solve the 2-PFLP problem by reduction to 2-SAT and solution of the 2-SAT problem. Wagner and Wolff claim that their admissibility test performs well in solving 4-VPFLP problems.

We tested the WW algorithm for 4-VPFLP in comparison with binary search using other admissibility tests, in particular: the SA variant mentioned above, a random-descent method (DE) also proposed by Christensen et al. [1], and a baseline random-placement method (RND). Figure 2 shows the result of these tests on a set of problems at a wide range of densities generated randomly according to the methods of Christensen et al. [1]. Label density, varying along the x axis, is given by the number of labels of size 30×7 on a map of dimensions 792×612 , ranging from 100 to 1000 in multiples of 100. At each label density, for each method tested, runs were made on 50 maps. Scale is presented on the y axis by the computed maximal scale relative to a benchmark “space-filling scale”, the scale at which the area of all the labels just equals the area of the map. This serves as a very weak upper bound on the optimal scale. The box plot shows the mean value over the 50 runs with a crosshair. The top middle and bottom lines in the box mark the 75th, 50th (median), and 25th percentiles, and the top and bottom extensions mark the 90th and 10th percentiles.

The graph shows the clear and unsurprising superiority of WW and SA over the two alternatives; the repeated use of an admissibility test in each VPFLP problem means that small improvements in a PFLP algorithm are leveraged into large improvements in a VPFLP algorithm. Figure 3 graphs the average ratio of SA label size to WW label size at different densities. The ratio stays extremely close to 1, with neither algorithm clearly superior. The differences are not significant according to a one-tail paired two-sample t-test over all 500 maps tested ($t = .42$, $df=499$, $P = .337$). Similar significance levels obtain for tests at each label density.

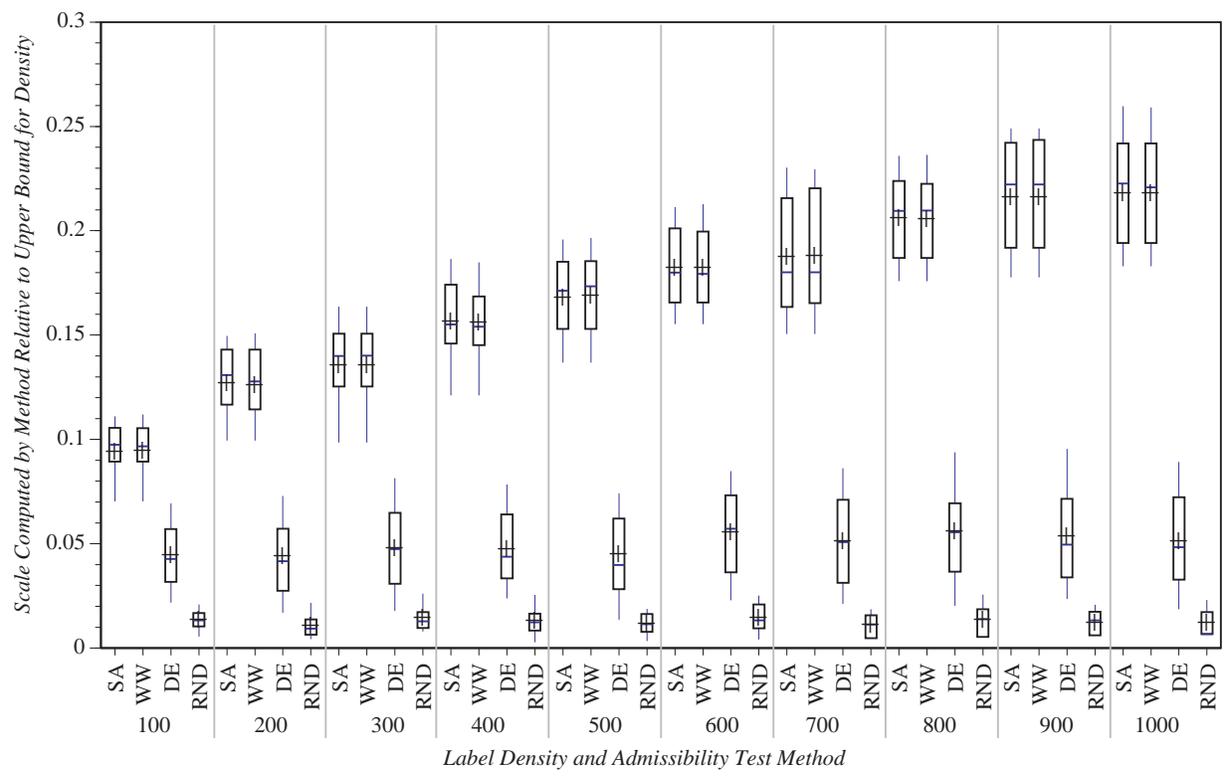


Figure 2: Box plot of four admissibility test methods used to solve 4-VPFLP at various densities.

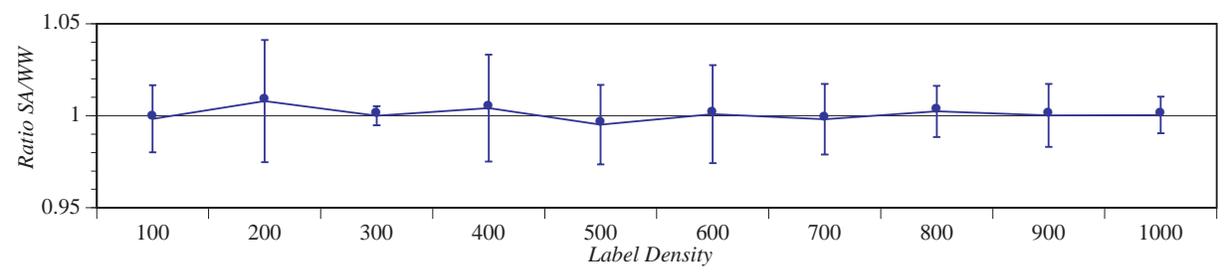


Figure 3: Mean and standard error of ratios of SA-computed scale to WW-computed scale for 4-VPFLP.

On average, the SA-based VPFLP algorithm required about three times more time than the WW-based one. This is not surprising, as the WW admissibility test is designed to fail quickly for PFLP problems with no perfect labeling, whereas SA must run through an entire annealing process before it can test whether the final labeling generated is perfect. Thus, for the 4-VPFLP problem, WW is preferable to SA.

On the other hand, the fact that the SA performance is essentially identical to that of WW means that SA is as good a method for determining admissibility as WW for the four-candidate variant. We can thus expect that for other versions of VPFLP for which WW cannot be applied, such as versions with more than four candidate positions per point, a priori preferences among positions, or with linear and area features, SA will perform extremely well. As extensions beyond the capability of WW are usual in cartographic applications, the efficacy of SA for variable-sized label placement problems is welcome news.

The preprocessing phase of WW is conservative in the sense that if a perfect labeling of the original set of candidate positions exists, then one will still exist after the preprocessing phase has eliminated its candidates. A natural question is whether the preprocessing phase might therefore be used to eliminate some potential candidate positions as a preprocess to SA so that SA might find better solutions to the PFLP problem. Of course, if the preprocess reports that no perfect labeling exists (as it well might), we need to decide what to do. Figure 4 shows the result of a test of algorithms using the preprocess to augment SA. (The methodology is identical to that used by Christensen et al. [1] so that the results are directly comparable.) The line marked “WW+SA halt on fail detect” stops eliminating candidate positions as soon as the preprocess concludes that no perfect labeling exists. The line marked “WW+SA, thorough” continues to eliminate candidate positions thereafter. This more aggressive use of the preprocess when used to prune candidate positions for the SA algorithm is clearly inferior. The former method, though it does not perform better than SA, performs only very slightly worse, and provides some slight speedup. Thus, the use of the preprocessing does not seem to provide a significant advantage to SA.

In summary, the empirical study reported on here demonstrates that SA performs as well as the WW 2-SAT heuristic on 4-VPFLP though it is considerably slower. Use of the WW preprocessing heuristic does not seem to benefit SA. The generality of SA makes it the method of choice for variable-sized label placement problems beyond 4-VPFLP.

Acknowledgements

This paper is based upon work supported in part by the National Science Foundation under Grant Number IRI-9350192 and by a gift from MERL — A Mitsubishi Electric Research Laboratory.

References

- [1] Jon Christensen, Joe Marks, and Stuart M. Shieber. An empirical study of algorithms for point feature label placement. *ACM Transactions on Graphics*, 14(3):203–232, July 1995.

- [2] Shawn Edmondson, Jon Christensen, Joe Marks, and Stuart M. Shieber. A general cartographic labeling algorithm. *Cartographica*, 1997. To appear.
- [3] Frank Wagner and Alexander Wolff. A practical map labeling algorithm. *Journal of Computational Geometry*. To appear.

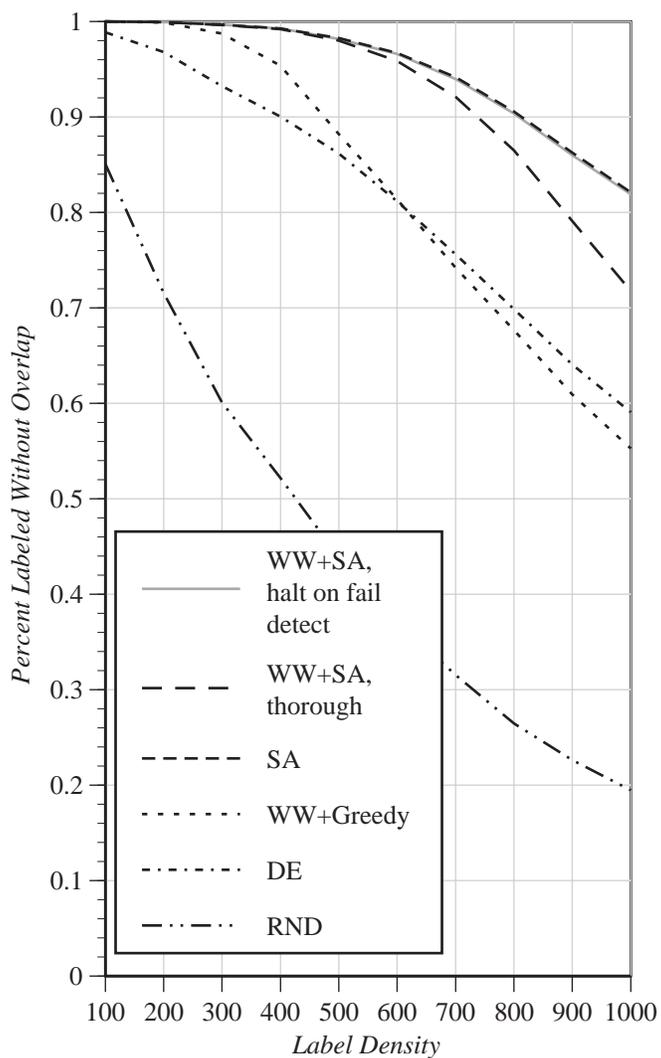


Figure 4: Performance of various algorithms on PFLP problems at various densities, including SA with and without WW preprocessing.