Labeling Points with Weights

Sheung-Hung Poon* Chan-Su Shin† Tycho Strijk‡ Alexander Wolff§

Abstract

Annotating maps, graphs, and diagrams with pieces of text is an important step in information visualization that is usually refered to as label placement. We define nine label-placement models for labeling points with axis-parallel rectangles given a weight for each point. There are two groups; fixed-position models and slider models. We aim to maximize the weight sum of those points that receive a label.

We first compare our models by giving bounds for the ratios between the weights of maximum-weight labelings in different models. Then we present algorithms for unit-height labels. We give an $O(n \log n)$-time factor-2 approximation algorithm for fixed-position models and the first algorithm for labeling weighted points with sliding labels. Its approximation factor is $(2 + \varepsilon)$ and its runtime in $O(n^2/\varepsilon)$ for any $\varepsilon > 0$.

1 Introduction

Label placement is one of the key tasks in the process of information visualization. In diagrams, maps, technical or graph drawings, features like points, lines, and polygons must be labeled to convey information. The interest in algorithms that automate this task has increased with the advance in type-setting technology and the amount of information to be visualized. For an extensive bibliography about label placement see [8]. The ACM Computational Geometry Impact Task Force report [2] denotes label placement as an important research area.

This paper deals with one of the most basic label-placement problems, namely labeling points with axis-parallel rectangles. There is an abundance of publications on this problem, see [8]. However, with two exceptions this is the first paper that gives approximation algorithms for points with weights, which is extremely important for practical applications. The only two other approximation algorithms for weighted label placement are the following. First, Iturriaga [5] gave a factor-$O(\log n)$ approximation algorithm for finding a maximum weight independent set (MWIS) in a set of n axis-parallel rectangles. Second, Erlebach et al. improved this result for squares by giving a polynomial-time approximation scheme (PTAS) for the weighted case [3].

Van Keveld et al. [7] forged the term of *slider models* where a label can slide along one or several edges under the constraint that it touches the point it labels, see Figure 1. This is opposed to *fixed-position models* that allow only a constant number of *label candidates* per point. Van Kreveld et al. compared a number of fixed-position and slider models with respect

*Dept. of Comp. Science, HKUST, Hong Kong, hung@cs.ust.hk

†Dept. of Comp. Science, KAIST, Korea, cssin@jupiter.kaist.ac.kr

‡Dept. of Comp. Science, Utrecht University, The Netherlands, tycho@cs.ruu.nl

§Inst. of Mathmatics and Comp. Science, Greifswald University, Germany, awolff@mail.uni-greifwald.de

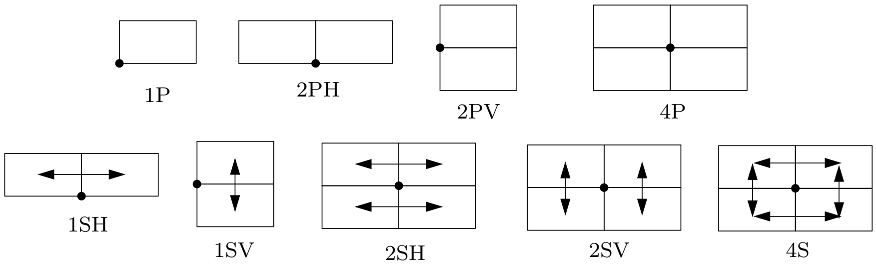

Figure 1: The labeling models that we consider in this paper.

to how many more points can be labeled in one model than in another using *unit square* labels. Figure 1 shows all nine fixed-position models and slider models that we will consider in this paper. In that figure, each (unit-height) rectangle stands for a feasible label position. An arrow between two rectangle indicates that additionally all label position are feasible that arise when moving one rectangle on a straight line onto the other.

For each of their labeling models, van Kreveld et al. gave a PTAS and a fast factor-2 approximation algorithm for labeling unweighted points with unit-height rectangles. They also did an experimental comparison that showed that algorithms for sliding labels perform especially well on dense point sets such as scatterplots. Other applications with dense point sets include drill-hole maps or electrophoresis gels, where usually hundreds of points must be labeled and thus automation is especially beneficial. We extend the results of van Keveld et al. by taking weights into account. Our main result is the first algorithm for labeling weighted points with sliding labels.

2 Comparing labeling models

Let M_1 and M_2 be any two different labeling models from Figure 1. Given a finite set P of points in the plane, each point with a weight, let $W_M(P)$ denote the maximum sum of weights of those points that can be labeled without intersections given labeling model M. Then the (M_1, M_2)-*ratio* is defined as $\Psi(M_1, M_2) = \lim_{n\to\infty} \max_{|P|=n} \frac{W_{M_1}(P)}{W_{M_2}(P)}$.

In order to bound this ratio simultaneously for several pairs of labeling models with similar properties, we use definitions similar to those for the unweighted problems in [7]. Let v be a unit vector parallel to the y-axis. We say that M_1 can be *flipped* into M_2 by v (see bold arrows in Figure 2) if any label position in M_1 that is not allowed in M_2 can be translated by v into a valid label position in M_2. M_1 can be *one-way slid* into M_2 by v (dashed arrows) if any label position in M_1 can be translated by σv into a valid label position in M_2 for some $\sigma \in [0, 1]$. M_1 can be *two-way slid* into M_2 by v (dotted arrows) if any label position in M_1 can be translated by σv for some $\sigma \in [-1, 1]$ into a valid label position in M_2 such that a corner of the label coincides with the point to be labeled.

In the full paper we show that the (M_1, M_2)-ratio equals 2 if M_1 can be flipped or two-way slid into M_2 and is at most 3 if M_1 can be one-way slid into M_2. These and our other results are summarized in Figures 2 and 3. The numbers that are attached to the arcs between two models M_1 and M_2 give the (M_1, M_2)-ratio; intervals specify lower and upper bounds. Some of the logarithmic bounds in Figure 3 are non-trivial.

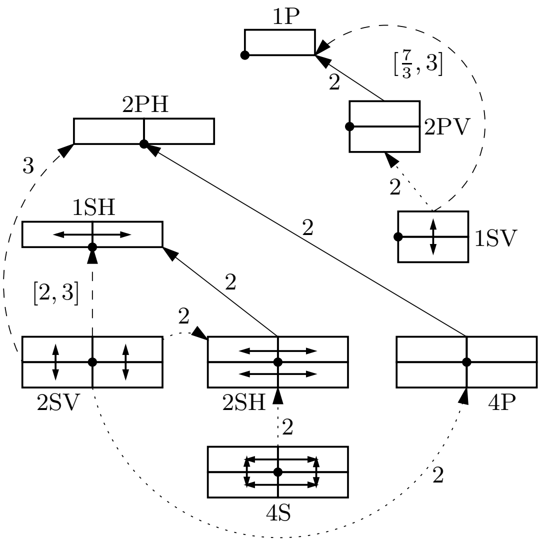

Figure 2: Constant ratios between different labeling models.

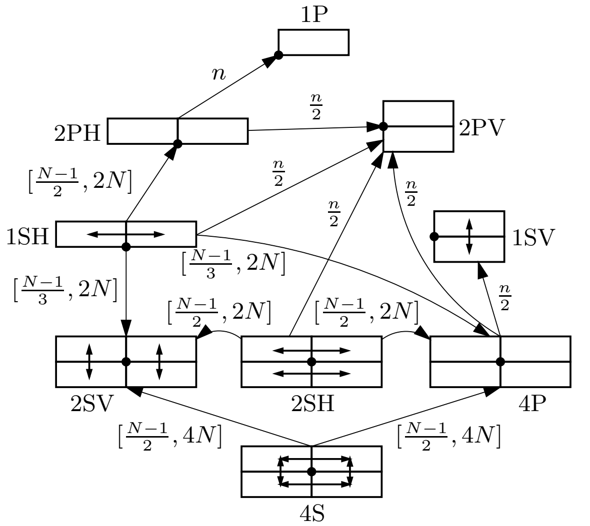

Figure 3: Ratios that cannot be bounded by constants. N is shorthand for $\log n$.

3 Approximation algorithms for unit-height labels

We first consider a 1d-problem, namely the problem *1d-1P* of finding a MWIS of n (topologically open) intervals given on the x-axis. The problem is exactly the 1d-version of 1P, and it can be solved in $O(n \log n)$ time by a simple dynamic programming algorithm [4]. The 1d-version *1d-2P* corresponding to 2PH is as follows. Given a set of n points, each with a weight and an interval length, find a MWIS from the $2n$ intervals that either start or end at one of the input points. We use open intervals but here we make them intersect artificially if they belong to the same point. This can be achieved by a symbolic comparison rule, which allows to use the same algorithm [4], although it assumes disjoint interval endpoints. The 2d-problems can be approximated by using line-stabbing as in [7].

Theorem 1 *The weighted fixed-position labeling problems* 1P, 2PH, 2PV*, and* 4P *can be factor-2 approximated in $O(n \log n)$ time.*

Next we consider algorithms for sliding labels. Again we first tackle the corresponding 1d-problem. Given a set of n points $x_1, \ldots, x_n$ on the x-axis, each with a weight w_i and an interval length l_i, the problem *1d-1S* consists of maximizing the weight sum of those points that can be labeled by intervals of the prescribed length such that (the closure of) each interval contains its point and no two intervals intersect. This is in fact a special case of the *single-machine throughput maximization* problem [1]. It is special in that in our case the execution window $[x_i - l_i, x_i + l_i]$ of each job is exactly twice the job length l_i. With a similar symbolic comparison rule as above the algorithm in [1, Theorem 9] yields a factor-$(1 + \varepsilon)$ approximation for 1d-1S. Again we use line-stabbing for the 2d-problems. This even works when vertical sliding is allowed, but the proof becomes more involved.

Theorem 2 *The weighted sliding problems* 1SH, 1SV, 2SH, 2SV*, and* 4S *can be* $(2 + \varepsilon)$*-approximated in $O(n^2/\varepsilon)$ time for any $\varepsilon > 0$.*

4 Discussion

1d-1S can be solved exactly by dynammic programming if the number of different weights is bounded, but we do not know how this can be done in the general case. While we have a PTAS for sliding unit-square labels we do not know how to extend this result to unit-height rectangles.

For arbitrary axis-parallel rectangular labels we have a factor-$(3+\varepsilon)\lceil \log_2 \beta \rceil$ approximation algorithm, where β is the ratio of maximum and minimum label height. For $\beta > 11$ this improves a practical factor-$(1 + \beta)$ approximation algorithm for labeling *unweighted* points with sliding labels [6].

References

[1] Piotr Berman and Bhaskar DasGupta. Improvements in throughput maximization for real-time scheduling. In *Proc. 32nd ACM Symp. on Theory of Computing (STOC'00)*, 2000. To appear. Full version available at `ftp://dimacs.rutgers.edu/pub/dimacs/TechnicalReports/TechReports/1999/99-52.ps.gz`.

[2] Bernard Chazelle et al. Application challenges to computational geometry: CG impact task force report. Technical Report TR-521-96, Princeton University, April 1996. `http://www.cs.princeton.edu/~chazelle/taskforce/CGreport.ps`.

[3] Thomas Erlebach, Klaus Jansen, and Eike Seidel. Polynomial-time approximation schemes for geometric graphs. In *Proceedings of the 12th ACM-SIAM Symposium on Discrete Algorithms (SODA'01)*, Washington, DC, 7–9 January 2001. To appear.

[4] Ju Yuan Hsiao, Chuan Yi Tang, and Ruay Shiung Chang. An efficient algorithm for finding a maximum weight 2-independent set on interval graphs. *Information Processing Letters*, 43(5):229–235, October 1992.

[5] Claudia Iturriaga. *Map Labeling Problems.* PhD thesis, University of Waterloo, 1999.

[6] Tycho Strijk and Marc van Kreveld. Practical extensions of point labeling in the slider model. In *Proc. 7th ACM Symposium on Advances in Geographic Information Systems*, pages 47–52, 5–6 November 1999.

[7] Marc van Kreveld, Tycho Strijk, and Alexander Wolff. Point labeling with sliding labels. *Computational Geometry: Theory and Applications*, 13:21–47, 1999.

[8] Alexander Wolff and Tycho Strijk. A map labeling bibliography. `http://www.math-inf.uni-greifswald.de/map-labeling/bibliography/`, 1996.