The districting problem

Citation for published version (APA):

Document status and date:
Published: 01/01/2009

Document Version:
Publisher’s PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:
• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:
www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:
openaccess@tue.nl
providing details and we will investigate your claim.
The Districting Problem

Esther M. Arkin∗ Irina Kostitsyna† Joseph S.B. Mitchell∗ Valentin Polishchuk‡
Girishkumar R. Sabhnani§

1 Introduction

We consider the following districting problem: Given a simple polygon partitioned into simple polygonal sub-districts (see Fig. 1) each with a given weight, merge the sub-districts into a minimum number of districts so that each district is a simple polygon and the total weight of the sub-districts in any district is at most a given number M. We consider also the “dual” version of the problem, in which the objective is to minimize the maximum district weight, subject to a given bound on the total number of districts.

Related Work The problem has its roots in political districting for voting where the districts may have restrictions on the number of sub-districts, size, total population, etc.; see, for example, the survey by Tasnádi [4]. Our motivation comes from an air traffic management problem in which sub-districts correspond to Fixed Posting Areas (or “sub-sectors”) with weights representing a measure of controllers’ “workload”, and the goal of the merging is to provide a balanced partitioning of the workload among a set of airspace sectors. In related work, Bloem et al. [1] analyze greedy heuristics for merging underutilized airspace sectors to conserve air traffic control resources. The districting problem is also related to problems in bin packing; in our case, however, the shape of the bins is not fixed (only their capacity is fixed), and the items are not allowed to be moved, but only to be grouped.

2 Hardness

A simple reduction from PARTITION (see Fig. 2) shows that the problem is weakly NP-hard; in fact, it is hard to distinguish between the cases in which the optimal solution has 2 vs. 3 districts. Hence, the problem is weakly hard to approximate to within \((3/2 - \varepsilon)\), for any \(\varepsilon > 0\).

Moreover, we show that the districting problem is strongly NP-hard by a reduction similar to the one in [3] that shows the hardness of array partitioning. The 5/4 hardness of approximation from [3] applies to the dual problem of minimizing the maximum weight of a district. With slight modifications to the construction, we also prove hardness (and 5/4 hardness of approximation for the dual version) for the following special cases of the districting problem: (i) that in which all districts are required to be convex, and (ii) that in which districts are allowed to be multi-connected (simple polygons with holes).
3 Approximation Algorithms

Our approximations are based on properties of the dual graph, $G$, of the input subdivision into sub-districts.

3.1 The dual graph $G$ is Hamiltonian

Assume that $G$ has a Hamiltonian path. Then, a simple clustering method follows the Hamiltonian path, greedily adding the sub-districts to a district, until the district’s weight is about to exceed the bound $M$, at which point we begin a new district. The total weight of any two consecutive districts thus obtained is more than $M$. Thus, the average district weight is more than $M/2$, so the number of districts is at most twice the optimal.

Even though the resulting districts are not necessarily simple polygons (Fig. 3), the total number of holes is bounded; indeed, since each hole is itself a district (or a group of districts), the number of holes is at most the number of districts. For each hole we can break the surrounding district into two districts, charging the increase in the number of districts to the hole. Thus, all holes can be removed if we allow the number of districts to (at most) double. This yields a 4-approximation for the districting problem with the constraint of having simply connected districts.

![Figure 3: Clustering sub-districts along a Hamiltonain path may result in non-simple districts.](image)

3.2 The dual graph $G$ is not Hamiltonian

If $G$ has no Hamiltonian path, then we turn instead to a low-degree spanning tree, $T$, of $G$. (One can compute in polynomial time a spanning tree that has degree at most 1 greater than the degree of a minimum-degree spanning tree [2].) In particular, we obtain a $2\Delta$-approximation, where $\Delta$ is the maximum degree of $T$. We root the tree $T$ and start from the leaves, clustering sub-districts into districts as we work our way towards the root. Specifically, for each node $v$ we merge $v$’s children in order of increasing subtree weight, until the district weight is about to exceed the threshold $M$. The average weight of the resulting districts is at least $M/\Delta$; thus, the number of districts is at most $\Delta$ times optimal. As described above, we can make all districts simply connected by removing holes, causing the number of districts to at most double.

Acknowledgements

We thank George Hart and other members of the Algorithms Reading Group at Stony Brook for helpful discussions. This research is partially supported by NSF (CCF-0528209, CCF-0729019), NASA Ames, and Metron Aviation. V. Polishchuk is supported in part by Academy of Finland grant 118653 (ALGODAN).

References


