Theoretical Computer Science 555 (2014) 43–54

Contents lists available at ScienceDirect

Theoretical Computer Science

www.elsevier.com/locate/tcs

A competitive strategy for distance-aware online shape allocation

Sándor P. Fekete *, Jan-Marc Reinhardt, Nils Schweer

Department of Computer Science, TU Braunschweig, Germany

ARTICLE INFO

Available online 7 March 2014

Keywords:
Clustering
Average distance
Online problems
Optimal shape of a city
Space-filling curves
Competitive analysis

ABSTRACT

We consider the following online allocation problem: Given a unit square S, and a sequence of numbers $n_i \in [0,1]$ with $\sum_{j=0}^{i} n_j \leqslant 1$; at each step i, select a region C_i of previously unassigned area n_i in S. The objective is to make these regions compact in a distance-aware sense: minimize the maximum (normalized) average Manhattan distance between points from the same set C_i. Related location problems have received a considerable amount of attention; in particular, the problem of determining the "optimal shape of a city", i.e., allocating a *single* n_i has been studied. We present an online strategy, based on an analysis of space-filling curves; for continuous shapes, we prove a factor of 1.8092, and 1.7848 for discrete point sets.

1. Introduction

Many optimization problems deal with allocating point sets to a given environment. Frequently, the problem is to find compact allocations, placing points from the same set closely together. One well-established measure is the average L_1 distance between points. A practical example occurs in the context of grid computing, where one needs to assign a sequence of jobs i, each requiring an (appropriately normalized) number n_i of processors, to a subset C_i of nodes of a large square grid, such that the average communication delay between nodes of the same job is minimized; this delay corresponds to the number of grid hops [11], so the task amounts to finding subsets with a small average L_1, i.e., *Manhattan* distance. Karp et al. [8] studied the same problem in the context of memory allocation.

Even in an offline setting without occupied nodes, finding an optimal allocation for one set of size n_i is not an easy task; as shown in Fig. 1, the results are typically "round" shapes. If a whole sequence of sets have to be allocated, packing such shapes onto the grid will produce gaps, causing later sets to become disconnected, and thus leads to extremely bad average distances. Even restricting the shapes to be rectangular is not a remedy, as the resulting problem of deciding whether a set of squares (which are minimal with respect to L_1 average distance among all rectangles) can be packed into a given square container is NP-hard [10]; moreover, disconnected allocations may still occur.

In this paper, we give a first algorithmic analysis for the *online* problem. Using an allocation scheme based on a space-filling curve, we establish competitive factors of 1.8092 and 1.7848 for minimizing the maximum average Manhattan distance within an allocated set.

* Corresponding author.
E-mail addresses: s.fekete@tu-bs.de (S.P. Fekete), j-m.reinhardt@tu-bs.de (J.-M. Reinhardt), n.schweer@tu-bs.de (N. Schweer).

http://dx.doi.org/10.1016/j.tcs.2014.02.050

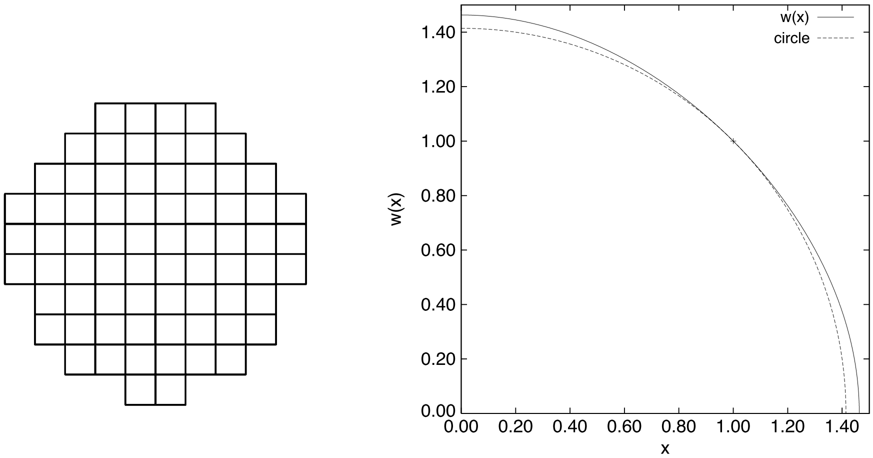

Fig. 1. Finding optimal individual shapes. (Left) An optimal shape composed of $n = 72$ grid cells, according to [5]. (Right) The optimal limit curve $w(x)$, according to [2].

1.1. Related work

Compact location problems have received a considerable amount of attention. Krumke et al. [9] have considered the *offline* problem of choosing a set of n vertices in a weighted graph, such that the average distance is minimized. They showed that the problem is NP-hard (even to approximate); for the scenario in which distances satisfy the triangle inequality, they gave algorithms that achieve asymptotic approximation factors of 2. For points in two-dimensional space with distances measured according to the Manhattan metric, Bender et al. [2] gave a simple 1.75-approximation algorithm, and a polynomial-time approximation scheme for any fixed dimension.

The problem of finding the "optimal shape of a city", i.e., a shape of given area that minimizes the average Manhattan distance, was first considered by Karp, McKellar, and Wong [8]; independently, Bender, Bender, Demaine, and Fekete [1] showed that this shape can be characterized by a differential equation for which no closed form is known. For the case of a finite set of n points that needs to be allocated to a grid, Demaine et al. [5] showed that there is an $O(n^{7.5})$-time dynamic-programming algorithm, which allowed them to compute all optimal shapes up to $n = 80$. Note that all these results are strictly offline, even though the original motivation (register or processor allocation) is online.

Space-filling curves for processor allocation with our objective function have been used before, see Leung et al. [11]; however, no algorithmic results and no competitive factor was proven. Wattenberg [16] proposed an allocation scheme similar to ours, for purposes of minimizing the maximum *Euclidean diameter* of an allocated shape. Like other authors before (in particular, Niedermeier et al. [12] and Gotsman and Lindenbaum [7]), he considered a measure called *c-locality*: for a sequence $1, \dots, i, \dots, j, \dots$ of points on a line, a space-filling mapping $h(.)$ will guarantee $L_2(h(i), h(j)) < c\sqrt{|j-i|}$, for a constant c that is $\sqrt{6} \approx 2.449$ for the Hilbert curve, and 2 for the so-called H-curve. One can use c-locality for establishing a constant competitive factor for our problems; however, given that the focus is on bounding the worst-case distance ratio for an embedding instead of the average distance, it should come as no surprise that the resulting values are significantly worse than the ones we achieve. On a different note, de Berg, Speckmann, and van der Weele [4] consider treemaps with bounded aspect ratio. Other related work includes Dai and Su [3].

1.2. Our results

We give a first competitive analysis for the online shape allocation problem within a given bounding box, with the objective of minimizing the maximum average Manhattan distance. In particular, we give the following results.

- We show that for the case of continuous shapes (in which numbers n_i correspond to area), a strategy based on a space-filling Hilbert curve achieves a competitive ratio of 1.8092.
- For the case of discrete point sets (in which numbers n_i indicate the number of points that have to be chosen from an appropriate $N \times N$ orthogonal grid), we prove a competitive factor of 1.7848.
- We sketch how these factors may be further improved, but point out that a Hilbert-based strategy is no better than a competitive factor of 1.3504, even with an improved analysis.
- We establish a lower bound of 1.144866 for *any* online strategy in the case of discrete point sets, and argue the existence of a lower bound for the continuous case.

The rest of this paper is organized as follows. In Section 2, we give some basic definitions and fundamental facts. In Section 3, we provide a brief description of an allocation scheme based on a space-filling curve. Section 4 provides more