Doctopic: Algorithms, automata, complexity and games ARTICLE IN PRESS TCS:10979

Theoretical Computer Science ••• (••••) •••–•••

Contents lists available at ScienceDirect

Theoretical Computer Science

www.elsevier.com/locate/tcs

Covering points with convex sets of minimum size ☆

Sang Won Bae [a], Hwan-Gue Cho [b], William Evans [c], Noushin Saeedi [c], Chan-Su Shin [d,*]

[a] Kyonggi University, Suwon, Republic of Korea
[b] Pusan National University, Busan, Republic of Korea
[c] University of British Columbia, Vancouver, Canada
[d] Hankuk University of Foreign Studies, Yongin, Republic of Korea

ARTICLE INFO

Article history:
Received 13 May 2016
Accepted 11 November 2016
Available online xxxx

Keywords:
Covering points
Convex sets
Convex hulls
Area
Perimeter
Computational geometry

ABSTRACT

For a set P of n points in the plane, we present algorithms for finding two bounded convex sets that cover P such that the total area or perimeter of the convex sets is minimized in $O(n^4 \log n)$ and $O(n^2 \log n)$ time, respectively. The former is the first result for minimum total area, and the latter is an improvement on the fastest previous algorithm for minimum total perimeter, which runs in $O(n^3)$ time [24]. We also extend our algorithms to find $k \geqslant 2$ convex sets minimizing area in $O(n^{2k(k-1)} \log n)$ time. The algorithms can be applied to detect road intersections from the GPS trajectories of moving vehicles for automated map generation or partial clustering.

1. Introduction

Let P be a set of n points in the plane. We want to find two bounded convex sets whose union contains P and whose size is minimized. We measure the size of a convex set by its area or perimeter. We consider two different optimization versions; one to minimize the sum of their sizes, called *min-sum* optimization, and the other to minimize their maximum size, called *min-max* optimization.

The optimal two bounded convex sets (A, B) covering P are, in fact, two convex hulls on a partition (P_1, P_2) of P, that is, $A = \text{conv}(P_1)$ and $B = \text{conv}(P_2)$, where $\text{conv}(P)$ denotes the convex hull of P. Thus the problem is equivalent to finding a partition (P_1, P_2) of P such that $\mu(\text{conv}(P_1)) + \mu(\text{conv}(P_2))$ is minimized, or $\max(\mu(\text{conv}(P_1)), \mu(\text{conv}(P_2)))$ is minimized, where $\mu(C)$ is either the area, $\text{area}(C)$, or the perimeter, $\text{peri}(C)$, of C. For the degenerate case that a convex hull is a line segment, its area is zero, and its perimeter is defined as twice the length of the segment. The problem can extend to covering P with $k > 2$ convex hulls.

It is natural to characterize the geometric structure of two convex hulls with minimum size, for instance, the existence of two optimal hulls which can be separated by a line. A partition (P_1, P_2) of P is said to be *linearly separable* if there is a line ℓ such that P_1 lies on one side of ℓ and P_2 lies on the other side. Note that one of P_1 and P_2 may be an empty

☆ S.W. Bae is supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2015R1D1A1A01057220). W. Evans and N. Saeedi are supported by NSERC Discovery Grant (RGPIN-2016-03856). C.-S. Shin is supported by Research Grant of Hankuk Univ. of Foreign Studies.

* Corresponding author.
E-mail addresses: swbae@kgu.ac.kr (S.W. Bae), hgcho@pusan.ac.kr (H.-G. Cho), will@cs.ubc.ca (W. Evans), noushins@cs.ubc.ca (N. Saeedi), cssin@hufs.ac.kr (C.-S. Shin).

http://dx.doi.org/10.1016/j.tcs.2016.11.014
0304-3975/

Please cite this article in press as: S.W. Bae et al., Covering points with convex sets of minimum size, Theoret. Comput. Sci. (2016), http://dx.doi.org/10.1016/j.tcs.2016.11.014

subset. If there is an optimal partition (P_1, P_2) which is linearly separable, then we can find its separation line by looking at all lines defined by two points from P and their associated partitions. Since we have $O(n^2)$ partitions, and compute two convex hulls for each partition in $O(n \log n)$ time, we can compute an optimal partition in $O(n^3 \log n)$ time, possibly $O(n^3)$ time [24] by pre-sorting the points angularly. Thus it is useful to determine whether the optimal partition is linearly separable or not.

For the min-sum optimization, the perimeter-optimal partition is linearly separable [8], but the area-optimal partition is not; Fig. 7(a) is an example of this. For the min-max optimization, both are not linearly separable; Fig. 7(a) is again an example of this for the area-optimal partition, and also an example for the perimeter-optimal partition, which will be explained later. If the optimal partition in not linearly separable, finding the partition becomes much harder, and an efficient algorithm must exploit other properties that hold for every optimal partition.

1.1. Related work

Partitioning P into $k \geqslant 2$ subsets under some optimization criterion is a clustering problem. Our criteria, which are based on properties of the convex hulls of the partitions (clusters), are most closely related to investigations of this clustering problem from a particular geometric perspective [4,8,17,24].

Capoyleas et al. [8] first showed that any pair of two clusters among k clusters with min-sum or min-max diameters or radii are linearly separable. Based on this fact, they presented algorithms for finding optimal k-clustering in $O(n^{6k-11} \log n)$ time for any $k > 2$. Their argument can be applied to find two clusters for minimizing the sum of perimeters of their convex hulls, so the min-sum perimeter problem on P can be solved in the same time for $k > 2$. For $k = 2$, the min-max and min-sum diameter problems were solved in $O(n \log n)$ time [4] and $O(n \log^2 n / \log \log n)$ time [17], respectively. The min-sum perimeter problem was solved in $O(n^3)$ time [24] by simply checking all bipartitions of P.

Hershberger et al. [18] summarized the shape of streaming point data using limited memory by covering the points approximately with a bounded number of convex hulls, called a “ClusterHull”. They used a combined cost function $\text{area}(H) + c \cdot \text{peri}(H)$ for each convex hull H, and clustered the points into k convex hulls to minimize total cost, where c is a constant, and k depends on the memory size $m \ll n$. This is an “on-line” algorithm to cover the streaming data approximately, so it is essentially different than our off-line algorithm to cover the static data exactly.

Many researchers have focused on the problem of finding a subset $P' \subseteq P$ of size k whose convex hull has the minimum area or perimeter [1,5,11–13,15,19]. The minimum-area subset P' of size k can be found in $O(kn^3)$ time [13] and $O((k^3 + \log n)n^2)$ time [12], while the minimum-perimeter subset P' can be found in $O(k^2 n \log n + k^4 n)$ time [11].

Yet another related line of work addresses convex hull construction from imprecise point data [16,23,25]. Löffler and van Kreveld [23] gave algorithms for selecting the points from imprecise points modeled as line segments or squares such that the convex hull of the selected points has a maximum/minimum area/perimeter. They solved the minimum area and minimum perimeter problems for squares in $O(n^2)$ and $O(n^7)$ time. No polynomial time algorithm is known for the minimum area problem on the line segments, nor is the problem known to be NP-hard.

Covering a point set P by geometric objects is a widely investigated topic. Perhaps the most related results consider covering P by two objects such as disks [9], rectangles [6,21], and squares [20,26]. These problems search for two optimal objects whose union covers P and the larger area is minimized. This is a typical min-max optimization. The main difference with our problem is that the two optimal objects are determined by a constant number of points in P, which always guarantees polynomial algorithms.

1.2. Our results

We first present an algorithm to cover an n-point set P in the plane with two convex sets (or hulls) of minimum total area, which runs in $O(n^4 \log n)$ time. This is the first result on the minimum total area. The key property that we will use is that if two optimal convex hulls intersect at their boundaries, then they can intersect at most four times. This property can also be used to find k convex hulls of minimum total area that cover P for any $k \geqslant 2$ in time $O(n^{2k(k-1)} \log n)$.

We next describe an algorithm to find two convex hulls covering P with minimum total perimeter in $O(n^2 \log n)$ time, which improves the existing $O(n^3)$ time algorithm [24]. For $k \geqslant 3$, Capoyleas et al. [8] give a $O(n^{6k-11} \log n)$ time solution. Improving this result is an open question.

We also consider min-max perimeter problem. We show, unlike the min-sum perimeter problem, that there is a point set P for which any optimal partition with min-max perimeter is not linearly separable. Interestingly, this observation makes the min-max perimeter version much harder than the min-sum one.

1.3. Applications

A recent advance in mobile services using GPS information triggered the generation of massive GPS tracking data of moving vehicles (including hurricane tracking data and animal movement data), which allows at a low cost the generation of new maps or refinement of existing maps [2,3,10,14,22,27]. One of the issues in such applications is to identify the 3-way or 4-way intersections of the road precisely. If the car trajectories are traced near the intersection within some time interval, then the change of convex hulls that cover a point set (car locations at a certain time) can be used to determine where

Please cite this article in press as: S.W. Bae et al., Covering points with convex sets of minimum size, Theoret. Comput. Sci. (2016), http://dx.doi.org/10.1016/j.tcs.2016.11.014