Computational Geometry 53 (2016) 27–35

Contents lists available at ScienceDirect

Computational Geometry: Theory and Applications

www.elsevier.com/locate/comgeo

Tighter estimates for ϵ-nets for disks

Norbert Bus [a], Shashwat Garg [b], Nabil H. Mustafa [a,*,1], Saurabh Ray [c]

[a] *Université Paris-Est, Laboratoire d'Informatique Gaspard-Monge, Equipe A3SI, ESIEE Paris, France*
[b] *IIT Delhi, India*
[c] *Computer Science, New York University, Abu Dhabi, United Arab Emirates*

ARTICLE INFO

Article history:
Received 17 December 2014
Accepted 20 December 2015
Available online 29 December 2015

Keywords:
Epsilon nets
Delaunay triangulations
Disks
Hitting sets

ABSTRACT

The geometric hitting set problem is one of the basic geometric combinatorial optimization problems: given a set P of points, and a set $\mathcal{D}$ of geometric objects in the plane, the goal is to compute a small-sized subset of P that hits all objects in $\mathcal{D}$. In 1994, Bronnimann and Goodrich [5] made an important connection of this problem to the size of fundamental combinatorial structures called ϵ-nets, showing that small-sized ϵ-nets imply approximation algorithms with correspondingly small approximation ratios. Very recently, Agarwal and Pan [2] showed that their scheme can be implemented in near-linear time for disks in the plane. Altogether this gives $O(1)$-factor approximation algorithms in $\tilde{O}(n)$ time for hitting sets for disks in the plane.

This constant factor depends on the sizes of ϵ-nets for disks; unfortunately, the current state-of-the-art bounds are large – at least $24/\epsilon$ and most likely larger than $40/\epsilon$. Thus the approximation factor of the Agarwal and Pan algorithm ends up being more than 40. The best lower-bound is $2/\epsilon$, which follows from the Pach–Woeginger construction [32] for halfplanes in two dimensions. Thus there is a large gap between the best-known upper and lower bounds. Besides being of independent interest, finding precise bounds is important since this immediately implies an improved linear-time algorithm for the hitting-set problem.

The main goal of this paper is to improve the upper-bound to $13.4/\epsilon$ for disks in the plane. The proof is constructive, giving a simple algorithm that uses only Delaunay triangulations. We have implemented the algorithm, which is available as a public open-source module. Experimental results show that the sizes of ϵ-nets for a variety of data-sets are lower, around $9/\epsilon$.

1. Introduction

The minimum hitting set problem is one of the most fundamental combinatorial optimization problems: given a range space $(P, \mathcal{D})$ consisting of a set P and a set $\mathcal{D}$ of subsets of P called the *ranges*, the task is to compute the smallest subset $Q \subseteq P$ that has a non-empty intersection with each of the ranges in $\mathcal{D}$. This problem is strongly NP-hard. If there are no restrictions on the set system $\mathcal{D}$, then it is known that it is NP-hard to approximate the minimum hitting set within a logarithmic factor of the optimal [34]. The problem is NP-complete even for the case where each range has exactly two

* Corresponding author.
E-mail addresses: busn@esiee.fr (N. Bus), garg.shashwat@gmail.com (S. Garg), mustafan@esiee.fr (N.H. Mustafa), saurabh.ray@nyu.edu (S. Ray).

1 The work of Nabil H. Mustafa in this paper has been supported by the grant ANR SAGA (JCJC-14-CE25-0016-01).

http://dx.doi.org/10.1016/j.comgeo.2015.12.002

points since this problem is equivalent to the vertex cover problem which is known to be NP-complete [22,18]. A natural occurrence of the hitting set problem occurs when the range space $\mathcal{D}$ is derived from geometry – e.g., given a set P of n points in $\mathbb{R}^2$, and a set $\mathcal{D}$ of m triangles containing points of P, compute the minimum-sized subset of P that hits all the triangles in $\mathcal{D}$. Unfortunately, for most natural geometric range spaces, computing the minimum-sized hitting set remains NP-hard. For example, even the (relatively) simple case where $\mathcal{D}$ is a set of unit disks in the plane is strongly NP-hard [21]. Therefore fast algorithms for computing provably good approximate hitting sets for geometric range spaces have been intensively studied for the past three decades (e.g., see the two recent PhD theses on this topic [16,17]).

Computing hitting sets for disks in the plane has been the subject of a long line of research. The case when all the disks have the same radius is easier, and has been studied in a series of works: Călinescu et al. [8] proposed a 108-approximation algorithm, which was subsequently improved by Ambhul et al. [3] to 72. Carmi et al. [9] further improved that to a 38-approximation algorithm, though with the running time of $O(n^6)$. Claude et al. [13] were able to achieve a 22-approximation algorithm running in time $O(n^6)$. More recently Fraser et al. [14] presented a 18-approximation algorithm in time $O(n^2)$. Mustafa et al. [28] showed a QPTAS for the dual problem of covering points by weighted disks and pseudo-disks in the plane.

So far, besides ad-hoc approaches, there are two systematic lines along which all progress on the hitting-set problem for geometric ranges has relied on: rounding via ϵ-nets, and local-search. The local-search approach starts with any hitting set $S \subseteq P$, and repeatedly decreases the size of S, if possible, by replacing k points of S with $\leq k-1$ points of $P \setminus S$. Call such an algorithm a k-local search algorithm. It has been shown [30] that a k-local search algorithm for the hitting set problem for disks in the plane gives a PTAS. Unfortunately the running time of their algorithm to compute a $(1+\epsilon)$-approximation is $O(n^{O(1/\epsilon^2)})$. Very recently Bus et al. [6] were able to improve the analysis and algorithm of the local-search approach to design a 8-approximation running in time $O(n^{2.33})$. However, at this moment, a near-linear time algorithm based on local-search seems beyond reach. We currently do not even know how to compute the most trivial case, namely when $k=1$, of local-search in near-linear time: given the set of disks D, and a set of points P, compute a *minimal* hitting set in P of D.

Rounding via ϵ-nets Given a range space $(P, \mathcal{D})$ and a parameter $\epsilon > 0$, an ϵ-net is a subset $S \subseteq P$ such that $D \cap S \neq \emptyset$ for all $D \in \mathcal{D}$ with $|D \cap P| \geq \epsilon n$. The famous ϵ-net theorem of Haussler and Welzl [20] states that for range spaces with VC-dimension d, there exists an ϵ-net of size $O(d/\epsilon \log d/\epsilon)$; this bound was later improved to $O(d/\epsilon \log 1/\epsilon)$ and which was shown to be optimal in general [23]. Sometimes, weighted versions of the problem are considered in which each $p \in P$ has some positive weight associated with it so that the total weight of all elements of P is 1. The weight of each range is the sum of the weights of the elements in it. The aim is to hit all ranges with weight more than ϵ. The condition of having finite VC-dimension is satisfied by many geometric set systems: disks, half-spaces, k-sided polytopes, r-admissible set of regions etc. in $\mathbb{R}^d$. For certain range spaces, one can further improve the bound of the ϵ-net theorem [36,10,12,29,24,35,31]. An important case is ϵ-net for disks in the plane, for which there are several proofs showing the existence of $O(1/\epsilon)$-sized nets [33].

In 1994, Bronnimann and Goodrich [5] proved the following interesting connection between the hitting-set problem, and ϵ-nets: let $(P, \mathcal{D})$ be a range-space for which we want to compute a minimum hitting set. If one can compute an ϵ-net of size c/ϵ for the ϵ-net problem for $(P, \mathcal{D})$ in polynomial time, then one can compute a hitting set of size at most $c \cdot \text{Opt}$ for $(P, \mathcal{D})$, where Opt is the size of the optimal (smallest) hitting set, in polynomial time. A shorter, simpler proof was given by Even et al. [15]. Both these proofs construct an assignment of weights to points in P such that the total weight of each range $D \in \mathcal{D}$ (i.e., the sum of the weights of the points in D) is at least $(1/\text{Opt})$-th fraction of the total weight. Then a $(1/\text{Opt})$-net with these weights is a hitting set. Until very recently, the best such rounding algorithms had running times of $\Omega(n^2)$, and it had been a long-standing open problem to compute a $O(1)$-approximation to the hitting-set problem for disks in the plane in near-linear time. In a recent break-through, Agarwal and Pan [2] presented an algorithm that is able to do the required rounding efficiently for a broad set of geometric objects. In particular, they are able to get the first near-linear algorithm for computing $O(1)$-approximations for hitting sets for disks.

Bounds on ϵ-nets The result of Agarwal and Pan [2] opens the way, for the first time, for near linear-time algorithms for the geometric hitting set problem. The catch is that the approximation factor depends on the sizes of ϵ-nets for disks; despite over seven different proofs of $O(1/\epsilon)$-sized ϵ-nets for disks, the precise bounds are not very encouraging. The paper containing the earliest proof, Matousek et al. [27], was over twenty-two years ago and thus summarized their result:

> *"Note that in principle the ϵ-net construction presented in this paper can be transformed into a deterministic algorithm that runs in polynomial time, $O(n^3)$ at worst. However, we certainly would not advocate this algorithm as being practical. We find the resulting constant of proportionality also not particularly flattering."* [27]

So far, the best constants for the ϵ-nets come from the proofs in [33] and [19]. Denote by $f(\alpha)$ the best known bound on the size of an α-net for lower halfspaces in $\mathbb{R}^3$. A lifting of the problem of disks to $\mathbb{R}^3$ gives an ϵ-net problem with lower halfspaces in $\mathbb{R}^3$. The former paper constructs $\frac{1}{4}$-nets for $4/\epsilon$ independent sub-problems, resulting in ϵ-nets of size $\frac{4}{\epsilon} f(\frac{1}{4})$ for halfspaces in $\mathbb{R}^3$. The latter paper presents five proofs for the existence of linear size ϵ-nets for halfspaces in $\mathbb{R}^3$. The best constant for disks is obtained by using their first proof, obtaining a bound of $\frac{4}{\epsilon} f(\alpha)$ where $\alpha < \frac{1}{3}$. Thus, by using