Information Processing Letters 135 (2018) 53–56

Contents lists available at ScienceDirect

Information Processing Letters

www.elsevier.com/locate/ipl

Approximation algorithms for color spanning diameter

Mohammad Reza Kazemi [a], Ali Mohades [a,*], Payam Khanteimouri [b]

[a] Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran
[b] Department of Computer Science, Faculty of Mathematical Sciences, Alzahra University, Tehran, Iran

ARTICLE INFO

Article history:
Received 15 May 2017
Received in revised form 23 January 2018
Accepted 24 January 2018
Available online 7 February 2018
Communicated by B. Doerr

Keywords:
Approximation algorithms
Inapproximability
Diameter
Color spanning set

ABSTRACT

Minimum Diameter Color Spanning Set (MDCSS) on a given set of colored points is the problem of selecting one point from each color such that the diameter of the selected points gets minimized. In this paper, we present some approximation algorithms and show some results on approximability of this problem in low and high dimensions.

1. Introduction

Given a set of colored points, one may be interested in finding Minimum Diameter Color Spanning Set (MDCSS). The MDCSS problem is to compute a subset with minimum diameter where it contains at least one point from each color. It is a case of the more general problem of computing a color spanning set such that an extent measure gets minimized (or maximized).

Since the radius of the minimum enclosing ball for any point set is at most a constant factor of the diameter, therefore, the Minimum Color Spanning Ball (MCSB) approximates the MDCSS. Recently, a linear time approximation scheme has been introduced for the MCSB [7] that is a constant factor approximation for the MDCSS. Also it is known that the MDCSS problem is NP-Hard even for the points in the plane [3]. Moreover, a polynomial time approximation scheme has been presented [4] that runs in $O(2^{\frac{1}{\epsilon^d}} \cdot \epsilon^{-2d} \cdot n^3)$ time. We aim to focus on and improve it as a part of our contributions.

In this paper, for the MDCSS problem, we propose a $(1+\epsilon)$-approximation algorithm that runs in $2^{O(\gamma \log \gamma)} \times O(n \log n)$ time for MDCSS where d is the dimension and $\gamma = \epsilon^{\frac{d-1}{2}}$. We also demonstrate that by taking the Exponential Time Hypothesis (ETH), there is not any $(1+\epsilon)$-approximation algorithm running in $2^{o(\gamma)}\text{poly}(n)$ time for this problem. In high dimensional spaces, when the dimension is considered as an input parameter, we argue that there is not any polynomial time approximation algorithm with ratio $\sqrt{2}-\epsilon$ for any $\epsilon > 0$ unless P = NP. Furthermore, by using coreset for computing the minimum enclosing ball [2], we show that there is a $(\sqrt{2}+\epsilon)$-approximation algorithm running in $O(dn^{\lceil 1/\sqrt{2}\epsilon \rceil + 1})$ time.

2. MDCSS problem in low dimensions

In this section, we approximate MDCSS in low dimensions and give a lower bound for any approximation algorithm. Throughout this section, we consider the dimension d as a constant.

* Corresponding author.
E-mail addresses: mr.kazemi@aut.ac.ir (M.R. Kazemi), mohades@aut.ac.ir (A. Mohades), p.khanteimouri@aut.ac.ir (P. Khanteimouri).

https://doi.org/10.1016/j.ipl.2018.01.007
0020-0190/

Let P be a set of n points in $\mathbb{R}^d$ partitioned into k colors. A color spanning set is a subset that contains at least one point from each color. Let also $\mathcal{CS}$ be the collection of all color spanning subsets of P. The MDCSS problem is to compute a color spanning set $S^* \in \mathcal{CS}$ in which $\Delta^* = \mathrm{Diam}(S^*) = \min_{s \in \mathcal{CS}} \mathrm{Diam}(S)$. Note that $\mathrm{Diam}(S) = \max_{p,q \in S} |p - q|$.

2.1. Approximation algorithm

We sketch the algorithm as follows. We first consider a decision version of the problem in which we are given a parameter Δ in addition to the set of input points P, and the goal is to determine whether $\Delta^* \le \Delta$. To approach this, we present an algorithm that gives a slightly loose answer. Strictly speaking, for the case where $\Delta^* < (1-\epsilon)\Delta$ or $\Delta^* > (1+\epsilon)\Delta$, the algorithm certainly makes the correct decision. Else, Δ is close enough to Δ^* and can be considered as a good approximation. Finally, we use a binary search on WSPD pair distances to compute a $(1+\epsilon)$-approximation.

Now, we have some definitions. Let us consider the set of *grid points* $\mathrm{Z} = \{ax | x \in \mathbb{Z}^d\}$ where a is the *grid side*. For any point $z = (z_1, \cdots, z_d) \in \mathrm{Z}$, the set of *grid cells* incident to z is $\{[z_1 - a, z_1], [z_1, z_1 + a]\} \times \cdots \times \{[z_d - a, z_d], [z_d, z_d + a]\}$. We define the parameter a such that $a = \Theta(\epsilon\Delta)$ and the diameter of any grid cell is at most $\epsilon\Delta/4$. Then, we round any point to its closest grid point. Let $c(z)$ be the set of colors of points that are rounded to the grid point z. We can restate the problem in the following way; The MDCSS problem on a multi colored grid Z is to find a subset Z^* with minimum diameter such that $\bigcup_{z \in Z^*} c(z)$ covers all the colors.

Since any point in P has a unique color, there is a color with frequency at most $\frac{n}{k}$. Let us denote this color by c and the set of its points by P^c. For any point $p \in P^c$, G_p denotes the set of all grid points which are incident to at least one grid cell that intersects the ball centered at p with radius Δ. G_p contains $O(1/\epsilon^d)$ points. Since the distance of each point from its rounded variant is at most $\frac{\epsilon\Delta}{4}$, it is adequate to solve MDCSS for G_p and then take the minimum over all $p \in P^c$. Indeed, by solving MDCSS over the rounded points, the additive error is at most $\epsilon\Delta/2$. Let $\Delta^*(G_p)$ be the solution of MDCSS on G_p. Because of rounding error, we know that there is a point $p \in P^c$ where $\Delta^* - \frac{\epsilon}{2}\Delta \le \Delta^*(G_p) \le \Delta^* + \frac{\epsilon}{2}\Delta$. Thus, if $\Delta^*(G_p) < (1 - \frac{\epsilon}{2})\Delta$ or $\Delta^*(G_p) > (1 + \frac{\epsilon}{2})\Delta$, we make a correct decision. Otherwise, $(1-\epsilon)\Delta \le \Delta^* \le (1+\epsilon)\Delta$ and we consider Δ as a $(1+\epsilon)$-approximation.

A naive algorithm to solve MDCSS over G_P is to test all subsets. The number of the subsets is $2^{O(1/\epsilon^d)}$ and checking whether a subset is color spanning takes $O(k/\epsilon^d)$ time. Since P^c has at most $\frac{n}{k}$ points, the algorithm runs in $O(n)2^{O(1/\epsilon^d)}$ total time. Note that $2^{O(1/\epsilon^d)}$ dominates $\mathrm{poly}(\frac{1}{\epsilon})$.

To reduce the running time, instead of checking all subsets, we guess an ϵ-kernel of the solution. A subset $Q \subseteq P$ is an ϵ-kernel if for any direction $\theta \in \mathbb{R}^d$,

$$\max_{x,y \in P} |\langle \theta, x - y \rangle| \le (1+\epsilon) \max_{x,y \in Q} |\langle \theta, x - y \rangle|;$$

where $\langle ., . \rangle$ indicates the dot product of two vectors. In fact, in any direction the width of an ϵ-kernel is a $(1+\epsilon)$-approximation of the width of the point set. Considering an ϵ-kernel suffices to approximate all extent measures including the diameter with ratio $(1+\epsilon)$. Agarwal et al. [1] demonstrated that for any point set in $\mathbb{R}^d$ there is an ϵ-kernel of size $O(1/\epsilon^{\frac{d-1}{2}})$ and it is also the worst case optimal. Let us denote the size of such an ϵ-kernel by γ.

Since an ϵ-kernel approximately preserves the diameter, it suffices to guess an ϵ-kernel of the solution of the MDCSS problem. For any subset $Q \subset G_p$, let $\mathcal{CH}^+(Q)$ denote the set of all grid points which are in the convex hull of Q or incident with a cell that intersects the convex hull of Q. Thus, for any subset Q of size $\gamma = O(1/\epsilon^{\frac{d-1}{2}})$, we consider $\mathcal{CH}^+(Q)$ as a candidate for the solution of MDCSS. This procedure is as complex as selecting γ points from $O(1/\epsilon^d)$. Now, we first estimate the number of ways to select γ things from γ^2.

$$\binom{\gamma^2}{\gamma} = 2^{\log \gamma^2! - \log \gamma! - \log(\gamma^2 - \gamma)!}.$$

Then, by using Stirling's formula;

$$\begin{aligned}
&\ln \gamma^2! - \ln \gamma! - \ln(\gamma^2 - \gamma)! \\
&= \gamma^2 \ln \gamma^2 - \gamma^2 - \gamma \ln \gamma + \gamma - (\gamma^2 - \gamma)\ln(\gamma^2 - \gamma) \\
&\quad + (\gamma^2 - \gamma) + O(\ln \gamma) \\
&= \gamma^2 \ln\left(\frac{\gamma^2}{\gamma^2 - \gamma}\right) + \gamma \ln(\gamma - 1) + O(\ln \gamma) = O(\gamma \log \gamma).
\end{aligned}$$

Note that $\gamma^2 \ln(\frac{\gamma^2}{\gamma^2 - \gamma}) = O(\gamma)$. It is easy to see that $\binom{O(1/\epsilon^d)}{\gamma}$ is also $2^{O(\gamma \log \gamma)}$. This is a significant improvement since $\gamma^2 = O(1/\epsilon^d)$. Therefore, we approximately solve the decision problem in $O(n)2^{O(\gamma \log \gamma)}$ time.

Now, all we need to approximate the MDCSS is to preform a binary search on all pair distances as candidates for the solution. Because the number of pair distances is $\Omega(n^2)$, instead, we use *well separated pair decomposition* (WSPD) to avoid a quadratic time. For a point set P, an ϵ-WSPD is a collection of pairs of subsets with some properties. We use an important property of WSPD here which states that for any pair $x, y \in P$ there is a pair $(A_i, B_i) \in \epsilon$−WSPD such that for any $x' \in A_i, y' \in B_i$, $(1-\epsilon)d(x', y') \le d(x, y) \le (1+\epsilon)d(x', y')$. In other words, the distance of any pair in P is approximated by the distance of a pair in ϵ-WSPD with ratio $1+\epsilon$. There are several algorithms to compute an ϵ-WSPD of size $O(n/\epsilon^d)$ in $O(n \log n)$ time [5].

We first construct an $\frac{\epsilon}{4}$-WSPD, and then use a binary search on all of the pair distances. Since we have already $\frac{\epsilon}{4}$-WSPD, there is a pair distance Δ such that $(1 - \frac{\epsilon}{4})\Delta \le \Delta^* \le (1 + \frac{\epsilon}{4})\Delta$. Thus, if we build the grid points so that $\Delta^* - \frac{\epsilon}{4}\Delta \le \Delta^*(G_p) \le \Delta^* + \frac{\epsilon}{4}\Delta$, then a $(1+\epsilon)$-approximation can be computed by a binary search procedure. Hence, the following theorem holds.

Theorem 1. *There is a $(1+\epsilon)$-approximation for MDCSS running in $O(n \log n)2^{O(\gamma \log \gamma)}$ time where $\gamma = 1/\epsilon^{\frac{d-1}{2}}$.*