Computing minimum diameter color-spanning sets is hard

Rudolf Fleischar,*,1, Xiaoming Xub

a School of Computer Science and IIPL, Fudan University, Shanghai, China
b Department of Computer Science, Duke University, Durham, USA

Abstract

We show that the minimum diameter color-spanning set problem is NP-hard for $L_p$ metric, $1 \leq p < \infty$.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

Assume we have a set of resources of one of several different types. We want to solve a task that requires simultaneous use of one resource of each type. There is a communication delay between any pair of resources. How should we allocate the resources so as to minimize the maximum delay between any two of our selected resources? This problem arises in large computer networks with different types of servers (think of a large company trying to pool resources to solve a certain computational task). It also arises in spatial databases, where it has recently been studied by Zhang et al. [6]. For example, we may want to search for a holiday location that features skiing, sailing, golfing, and shopping, all within short distance of each other.

The formal model. We now give a formal definition of the problem. Let $S$ be a set of $n$ points in the Euclidean plane.

Each point is colored with one of $k$ colors, for a fixed parameter $k \geq 1$. $S$ may be a multiset, which means a point in the plane may be colored simultaneously with several colors (in which case $S$ contains several points with the same coordinates but with different colors). We call a subset of $k$ points in $S$ of distinct colors a rainbow set. The minimum diameter color-spanning set problem (MDCS) is the problem of finding a rainbow set of smallest diameter.

A related concept is that of a color-spanning disk which is a disk containing at least one point of each color. Fig. 1 illustrates these definitions.

Previous work. Zhang et al. proposed an $O(n^5)$-time algorithm for MDCS based on a brute-force enumeration of all possible rainbow sets. Their algorithm was implemented in...

Fig. 1. An instance of MDCS with three colors. The minimum diameter color-spanning disk $B$ encloses the dashed triangle, while the solid triangle spans the minimum diameter rainbow set (which has a smaller diameter than $B$).
Our results. We are concerned with the complexity of solving \textit{MDCS}. We show that the problem is NP-hard for $L_p$ metric for $1 < p < \infty$. We do not know whether the problem is \textit{W[1]}-hard or APX-hard. For $L_1$ and $L_\infty$ metrics, \textit{MDCS} can easily be solved by computing a minimum diameter color-spanning disk \cite{5}.

Related work. We first studied the algorithmic complexity of \textit{MDCS} in a previous paper \cite{5}. Several other color-spanning set problems had been studied earlier by Abellanas et al. \cite{1}. They called rainbow sets color-spanning sets and gave efficient algorithms for the minimum diameter color-spanning disk problem and the smallest color-spanning rectangle problem \cite{2} (finding a smallest rectangle containing at least one point of each color). Fan et al. \cite{4} recently studied several other color-spanning problems; they gave an efficient randomized algorithm to compute a maximum diameter color-spanning set, and they showed it is NP-hard to compute a largest closest pair color-spanning set and a planar minimum color spanning tree.

2. Hardness of \textit{MDCS}

To show NP-hardness of \textit{MDCS} we first give the decision version of \textit{MDCS}: Given an instance of \textit{MDCS} and a positive number $d$, decide whether there exists a rainbow set of diameter at most $d$. This problem is in NP (we can compare the square of all pairwise point distances with $d^2$, avoiding square root calculations). Hardness of the decision problem implies hardness of the optimization problem.

\textbf{Theorem 1.} The decision version of \textit{MDCS} is NP-hard for any $L_p$ metric, for $1 < p < \infty$.

\textbf{Proof.} We show the hardness of \textit{MDCS} for $L_p$ metric, for any $1 < p < \infty$, by reduction from 3SAT. We first give the proof under the assumption that we can exactly compute coordinates of points on a circle. We will then show how to approximate these coordinates with low precision rational numbers without affecting the correctness of the proof.

Let $F$ be a Boolean formula in conjunctive normal form with $n$ variables $x_1, \ldots, x_n$ in $m$ clauses $c_1, \ldots, c_m$ of size at most three. To construct an instance $I$ of \textit{MDCS}, we consider an $L_p$ disk of diameter $1$. Let $C$ be the boundary of that disk. For each variable $x_i$, we select two antipodal slots $s_i$ and $\bar{s}_i$ on $C$, corresponding to the positive literal $x_i$ and the negative literal $\bar{x}_i$, respectively. These slots can be placed arbitrarily, but the slots for different variables should not be located on the same point. For example, we could create $2n$ equally-spaced slots on $C$.

For each clause $c_j$ we create a new color $col_{ij}$. If $x_i$ appears in $c_j$, we place a point of color $col_{ij}$ on slot $s_i$. Similarly, if $\bar{x}_i$ appears in $c_j$, we place a point of color $col_{ij}$ on slot $\bar{s}_i$. Note that several points can coincide if a literal appears in several clauses. Finally, we set $d = 1 - \epsilon$ for a small $\epsilon$ defined later. See Fig. 2 for an example of the construction in $L_2$ metric.

We now argue that $F$ is satisfiable if and only if $I$ contains a rainbow set of diameter at most $d$. If $F$ admits a satisfying assignment, then each clause $c_j$ contains at least one true literal $x_j$ (or $\bar{x}_j$). We add the corresponding point of color $col_{ij}$ at slot $s_j$ (or $\bar{s}_j$) to the set $R$. Then, $R$ is a rainbow set of diameter at most $d$.

If $I$ contains a rainbow set $R$ of $m$ points of diameter at most $d$, then $R$ cannot contain both $s_i$ and $\bar{s}_i$, for any $i$ (because $d < 1$). Therefore, $R$ induces a truth assignment for the variables $x_i$. If $s_i \in R$, we set $x_i = 1$, otherwise we set $x_i = 0$. Since $R$ contains one point of each color, each clause contains at least one true literal, i.e., $F$ is satisfied.

We now discuss the problem of approximating the slot coordinates on $C$ with low precision rationals. For example, if the slots are evenly spaced around the circle, then we can set $\epsilon$ to be any value strictly smaller than $\frac{1}{2n}$ and approximate the slot positions by choosing an arbitrary point inside a disk centered at the slot with diameter at most $\frac{1}{2}$. \hfill \square

3. Conclusions

We have shown that \textit{MDCS} is NP-hard for $L_p$ metrics, $1 < p < \infty$. Our proof does not show APX-hardness because the $\epsilon$ used in the proof can become arbitrarily small for large formulas. Another open problem is to study FPT (fixed parameter tractable) algorithms for \textit{MDCS}, for example with parameter $k$, the number of colors. \textit{MDCS} is polynomial-time solvable for two colors, while our NP-hardness reduction constructs an instance with a large number of colors. Which value of $k$ is the threshold between easy and hard?

Acknowledgements

We thank the reviewers for their helpful criticism.

References

\begin{thebibliography}{9}
\end{thebibliography}

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{fig2}
\caption{The NP-hardness reduction for $F = (x_1 \land \bar{x}_2 \land \bar{x}_4) \lor (x_2 \land x_3 \land x_4) \lor (x_1 \land x_3 \land x_4)$ in $L_2$ metric. The literals of the three clauses are colored white, gray, and black, respectively. Note that the slot $s_4$ is colored white and black.}
\end{figure}
