Theoretical Computer Science 593 (2015) 154–159

Contents lists available at ScienceDirect

Theoretical Computer Science

www.elsevier.com/locate/tcs

Note

Revisiting a randomized algorithm for the minimum rainbow subgraph problem

Chen Yuan, Haibin Kan *

Shanghai Key Lab of Intelligent Information Processing, Fudan University, China

ARTICLE INFO

Article history:
Received 24 July 2014
Received in revised form 13 May 2015
Accepted 28 May 2015
Available online 3 June 2015
Communicated by P. Widmayer

Keywords:
Randomized algorithm
Minimum rainbow subgraph

ABSTRACT

The minimum rainbow subgraph problem arises in bioinformatics. The graph is given as an edge-colored undirected graph. Our goal is to find a subgraph with minimum number of vertices such that there is exactly one edge from each color class. The currently best approximation ratio achieved by a deterministic approximation algorithm is $O(\Delta)$. (Here Δ is the max degree of a graph.) In [4] he proposes a randomized algorithm which achieves an approximation ratio of $O(\sqrt{\Delta \ln \Delta})$. However, we find that there is a flaw in his probability analysis which renders this approximation ratio invalid. We present a simple example to show why his analysis does not work. Instead, we propose an alternative analysis for his randomized algorithm. Our estimate shows that this randomized algorithm may achieve approximation ratio of $O(\Delta)$ in general. However, if the number of colors is $\Theta(n\Delta^r)$ for some positive $r \leq 1$, his randomized algorithm can beat the bound of $O(\Delta)$. Moreover, through our analysis, we also find that if we impose an extra constraint on the color function, the bound $O(\sqrt{\Delta \ln \Delta})$ still holds.

1. Introduction

The rainbow subgraph problem is important in combinatorics [3]. We first formalize this problem.

Definition 1 *(Minimum Rainbow Subgraph).* (See [4].) Given an undirected graph $G = (V, E)$, and a color function $col : E \to \{1, \ldots, p\}$, a rainbow subgraph of G is a graph $G' = (V', E')$ with $V' \subseteq V$ and $E' \subseteq E$ such that for any $i \in \{1, \ldots, p\}$, there is exactly one edge $e \in E'$ with $col(e) = i$. We aim to find such a graph with minimum number of vertices.

It is interesting to consider the algorithmic aspect of this problem. There are several deterministic algorithms achieving $O(\Delta)$ approximation [1,2,5]. Popa [4] proposes a randomized algorithm based on linear programming. He claims that this algorithm achieves approximation ratio of $O(\sqrt{\Delta \ln \Delta})$ which is clearly better than any known deterministic algorithm. However, there is a flaw in their probability analysis. After we fix this flaw, the approximation ratio of this randomized algorithm may still be $O(\Delta)$ in general.

We follow the notation in [4]. $\{1, \ldots, n\}$ is the vertex set of the input graph, m is the number of edges in G and Δ is the maximum degree of G. We say that a color w is covered by a subgraph $G' = (V', E')$ of G if there exists an edge $e \in E'$

* Corresponding author.
E-mail address: hbkan@fudan.edu.cn (H. Kan).

http://dx.doi.org/10.1016/j.tcs.2015.05.042

Algorithm 1 A randomized algorithm for the minimum rainbow subgraph problem.

INPUT: Graph $G = (V, E)$ colored with p colors and a parameter $q \in \mathbb{R}_+$.
$V' \leftarrow \emptyset$
Step 1: Solve the LP relaxation problem.
Step 2: For $i = 1$ to n add vertex i to V' with probability $min(1, q \cdot v_i)$
Step 3: For each uncovered color w:
Select arbitrarily an edge (i, j) colored with w and add vertex i and j to V'.
Let $E' \leftarrow (V' \times V') \cap E$. If there are two or more edges colored with the same color in E', keep only one of them.

such that $col(e) = w$. For a color w, f_w is the number of edges in G colored with w. p is the total number of colors in this graph.

At the end of this section, we state the main result of Popa [4] as a comparison.

Theorem 1. *(See [4].) There is a randomized algorithm achieving approximation ratio $O(\sqrt{\Delta \ln \Delta})$ when the number of colors p satisfies $p \geq n$.*

2. Randomized algorithm

For completeness, we give the description of his randomized algorithm. This algorithm contains two parts. The first part is concerned with integer programming and its relaxation. The goal of this integer program is to determine exactly which vertex we should select. Since it is NP-complete to solve integer linear programs, we have to relax the constraints and use probability to select vertices. The concrete integer linear program reads as follows:

$$\begin{aligned} &\text{minimize} && \sum_{i=1}^{n} v_i \\ &\text{subject to} && \\ & && \sum_{e \in E: col(e)=i} x_e \geq 1 \quad 1 \leq i \leq p \\ & && v_i \geq x_e \text{ for all edges } e \text{ incident to } i \\ & && x_e, v_i \in \{0, 1\} \end{aligned}$$

Here, variable v_i corresponds to vertex i and x_e corresponds to edge e. If $v_i = 1$, vertex i is part of the solution. Similarly, if $x_e = 1$, edge e is part of the solution. Our aim is to minimize the number of vertices. The LP relaxation is obtained by allowing variables x_e, v_i to have any positive value.

$$\begin{aligned} &\text{minimize} && \sum_{i=1}^{n} v_i \\ &\text{subject to} && \\ & && \sum_{e \in E: col(e)=i} x_e \geq 1 \quad 1 \leq i \leq p \\ & && v_i \geq x_e \text{ for all edges } e \text{ incident to } i \\ & && x_e, v_i \geq 0 \end{aligned}$$

A solution to this integer program is exactly the answer of minimum rainbow subgraph problem. Therefore, the solution of the LP relaxation is no larger than the optimal solution of the minimum rainbow subgraph problem. Popa [4] proposes a randomized algorithm (Algorithm 1) based on this LP relaxation.

Then, he claims:

Theorem 2. *(See [4].) Algorithm 1 is a randomized polynomial time $q + \frac{\Delta}{e^{pq^2/n\Delta}}$ approximation for the minimum rainbow subgraph problem.*

However, we find that there is a flaw in his probability analysis. Let us consider the following simple example. Here, we set parameter $q = 1$.

Example 1. K_3 is a complete graph with vertex set $\{1, 2, 3\}$ and edge set $\{(1, 2), (1, 3), (2, 3)\}$. Each edge belongs to the same color class w. In other words, there is only one color in K_3 and hence we have $p = 1$. The LP relaxation of this