ARTICLE IN PRESS

Discrete Applied Mathematics ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Discrete Applied Mathematics

journal homepage: www.elsevier.com/locate/dam

1.5-approximation algorithm for the 2-Convex Recoloring problem[☆]

Reuven Bar-Yehuda [a], Gilad Kutiel [a,*], Dror Rawitz [b]

[a] Department of Computer Science, Technion, Haifa 32000, Israel
[b] Faculty of Engineering, Bar Ilan University, Ramat-Gan 52900, Israel

ARTICLE INFO

Article history:
Received 17 January 2016
Received in revised form 19 December 2016
Accepted 12 January 2017
Available online xxxx

Keywords:
Approximation algorithms
Convex recoloring
Greedy

ABSTRACT

Given a graph $G = (V, E)$, a coloring function $\chi : V \to C$, assigning each vertex a color, is called *convex* if, for every color $c \in C$, the set of vertices with color c induces a connected subgraph of G. In the Convex Recoloring problem a colored graph G_χ is given, and the goal is to find a convex coloring χ' of G that *recolors* a minimum number of vertices. In the weighted version each vertex has a weight, and the goal is to minimize the total weight of recolored vertices. The 2-Convex Recoloring problem (2-CR) is the special case, where the given coloring χ assigns the same color to at most two vertices. 2-CR is known to be NP-hard even if G is a path.

We show that weighted 2-CR cannot be approximated within any ratio, unless P = NP. On the other hand, we provide an alternative definition of (unweighted) 2-CR in terms of maximum independent set of paths, which leads to a natural greedy algorithm. We prove that its approximation ratio is $\frac{3}{2}$ and show that this analysis is tight. This is the first constant factor approximation algorithm for a variant of CR in general graphs. For the special case, where G is a path, the algorithm obtains a ratio of $\frac{5}{4}$, an improvement over the previous best known approximation. We also consider the problem of determining whether a given graph has a convex recoloring of size k. We use the above mentioned characterization of 2-CR to show that a problem kernel of size $4k$ can be obtained in linear time and to design an $O(|E|) + 2^{O(k \log k)}$ time algorithm for parameterized 2-CR.

1. Introduction

Let $G = (V, E)$ be a graph and let $\chi : V \to C$ be a coloring function,[1] assigning each vertex in V a color in C. We say that χ is a *convex coloring* of G if, for every color $c \in C$, the vertices with color c induce a connected sub-graph of G. Fig. 1 shows an example of convex and non-convex colorings. In the Convex Recoloring problem (abbreviated CR), we are given a colored graph G_χ, and we wish to find a recoloring of a minimum number of vertices of G, such that the resulting coloring is convex. That is, the goal is to find a convex coloring χ', that minimizes the size of the set $\{v : \chi(v) \neq \chi'(v)\}$. The t-Convex Recoloring problem (t-CR) is the special case, in which the given coloring assigns the same color to at most t vertices in G. Fig. 2 depicts an input and a possible output for the 2-Convex Recoloring problem.

☆ A preliminary version was presented at the 26th International Workshop on Combinatorial Algorithms (IWOCA), 2015.
* Corresponding author.
E-mail addresses: reuven@cs.technion.ac.il (R. Bar-Yehuda), gkutiel@cs.technion.ac.il (G. Kutiel), dror.rawitz@biu.ac.il (D. Rawitz).

1 W.l.o.g. we assume that $\text{img}(\chi) = C$.

http://dx.doi.org/10.1016/j.dam.2017.01.008

(a) Convex. (b) Non-convex.

Fig. 1. Example of convex and non-convex colorings of a graph.

(a) An example of input to the problem. (b) A possible output with cost of 2.

Fig. 2. The 2-Convex Recoloring problem.

Related work. The Convex Recoloring problem (CR) in trees was introduced by Moran and Snir [10] and was motivated by its relation with the concept of *perfect phylogeny*. They proved that the problem is NP-hard [6], even when the given graph is a path. Later, Kanj et al. [6] showed that 2-CR is also NP-hard on paths. Applications of CR in general graphs, such as multicast communication, were described by Kammer and Tholey [5]. Many variants of the problem have been intensively investigated. The differences between one variant to another can be related to

- The structure of the given graph G. The given graph can be a path, a tree, a bounded treewidth graph, a general graph, and others.
- Constraints on the coloring function χ. In a t-coloring at most t colors are used to color a graph, while in a t-CR instance at most t vertices are colored using the same color.
- Type of weight function. In the weighted case, each vertex is associated with a weight, and the weight of a recoloring is the total weight of recolored vertices. In the unweighted case the weight of the solution is the number of recolored vertices. In a third variant, referred to as *block recoloring* [5], a cost is incurred for a color c if at least one vertex of color c was recolored.

Since CR was shown to be NP-hard, it was natural to try to design both approximation algorithms and parameterized algorithms.

Moran and Snir [9] presented a 2-approximation algorithm for CR in paths and a 3-approximation algorithm for CR in trees. Both algorithms work for the problem with weights. Bar-Yehuda, Feldman and Rawitz [2] improved the latter by providing a $(2+\varepsilon)$-approximation algorithm for CR in trees. This result was later extended to bounded treewidth graphs by Kammer and Tholey [5]. Recently, Lima and Wakabayashi [8] gave a $\frac{3}{2}$-approximation algorithm for unweighted 2-CR in paths.

On the negative side, Kammer and Tholey [5] proved that, if vertex weights are either 0 or 1, then 2-CR has no polynomial time approximation algorithm with a ratio of size $(1-o(1))\ln\ln n$ unless NP $\subseteq$ DTIME$(n^{O(\log\log n)})$. In Section 2 we show that this variant of the problem cannot be approximated at all. Campêlo et al. [4] showed that, for $t \geq 2$, CR is NP-hard on t-colored grids. They also proved that there is no polynomial time approximation algorithm within a factor of $c \ln n$ for some constant $c > 0$, unless P = NP, for unweighted CR in bipartite graphs with 2-colorings.

Moran and Snir [10] presented an algorithm for CR whose running time is $O(n^4 \cdot k(\frac{k}{\log k})^k)$, where k is the number of recolorings in an optimal solution. Razgon [12] gave a $2^{O(k)}\text{poly}(n)$ time algorithm for CR in trees. Ponta et al. [11] designed several algorithms for different variants of CR in trees. For the unweighted case, they gave a $O(3^b \cdot b \cdot n)$ time algorithm, where b is the number of colors that do not induce a connected subtree, and it is bounded from above by $2k$. Bar-Yehuda, Feldman and Rawitz [2] provided an algorithm with an upper bound of $O(n^2 + n \cdot k2^k)$ on the running time, but it is not hard to verify that this bound can be improved to $O(n \cdot k2^k)$. Bodlaender et al. [3] showed that CR admits a kernel of size $O(k^2)$ in

Please cite this article in press as: R. Bar-Yehuda, et al., 1.5-approximation algorithm for the 2-Convex Recoloring problem, Discrete Applied Mathematics (2017), http://dx.doi.org/10.1016/j.dam.2017.01.008.