

# Propriedades do Problema da Floresta Geradora $k$ -Rotulada

Pedro Jorge de Abreu Figueredo<sup>1</sup>, Manoel Campêlo<sup>1</sup>

<sup>1</sup>Mestrado e Doutorado em Ciência da Computação – Universidade Federal do Ceará  
Campus do Pici, Bl. 910, 60.440-900 – Fortaleza – CE – Brasil

{pedrojorge, mcampelo}@lia.ufc.br

**Abstract.** *We consider the  $k$ -Labeled Spanning Forest problem, which consists in finding a spanning forest of an edge-labeled graph with the least number of trees and at most  $k$  colors. We derive two characterizations of the problem and show strategies to convert an instance into an equivalent one. These results impact the development of solution approaches. We also identify some polynomial cases.*

**Resumo.** *Consideramos o problema da Floresta Geradora  $k$ -Rotulada, que consiste em encontrar uma floresta geradora de um grafo colorido em arestas com a menor quantidade de árvores e no máximo  $k$  cores. Derivamos duas caracterizações para o problema e mostramos estratégias para converter uma instância em outra equivalente. Esses resultados impactam o desenvolvimento de abordagens de solução. Identificamos também alguns casos polinomiais.*

## 1. O problema

Um grafo colorido em arestas (GCA) é um grafo simples e não direcionado com uma rotulação das arestas, podendo ser representado pela tupla  $G = (V, E, L, l)$ , onde  $V$  é o conjunto de vértices,  $E$  o conjunto de arestas e  $l : E \rightarrow L$  é uma função que atribui um rótulo  $l(e)$  do conjunto de rótulos (cores)  $L$  a cada aresta  $e \in E$ .

Dado um GCA  $G$  e um inteiro positivo  $k$ , o problema de Floresta Geradora  $k$ -Rotulada (FGkR) consiste em encontrar uma floresta geradora de  $G$  com a menor quantidade de árvores (componentes) e no máximo  $k$  cores distintas [Cerulli et al. 2014].

Esse problema está fortemente associado àquele da Árvore Geradora Minimamente Rotulada (AGMR), cujo objetivo é encontrar uma árvore geradora com a menor quantidade de cores de um GCA. Em [Chang and Shing-Jiuan 1997] provou-se que a versão de decisão do AGRM é um problema NP-Completo, e esse resultado pode ser estendido para o FGkR, através de uma redução direta apresentada por [Cerulli et al. 2014].

Ainda há muito poucos trabalhos relacionados ao FGkR. Os principais, devido a [Cerulli et al. 2014], [Consoli and Pérez 2015] e [Consoli et al. 2017], propõem meta-heurísticas para o problema, adaptando algoritmos similares desenvolvidos para o AGMR. [Cerulli et al. 2014] também apresentam um algoritmo enumerativo de *backtracking*.

Neste trabalho, apresentamos duas caracterizações para o FGkR, que fundamentam, por exemplo, diferentes formulações de programação inteira. Além disso, tomando por base o trabalho de [Silva 2018] para o AGMR, mostramos estratégias para converter uma instância em outra equivalente, que permitem modificar a estrutura do grafo, impactando na eficiência de métodos de solução. Finalmente, identificamos alguns casos polinomiais do problema.

## 2. Caracterizações do problema

Seja  $G = (V, E, L, l)$  um GCA. Dados  $E' \subseteq E$  e  $L' \subseteq L$ , sejam  $l(E') = \{l(e) \mid e \in E'\}$  os rótulos das arestas em  $E'$ ,  $E(L') = \{e \in E \mid l(e) \in L'\}$  as arestas com rótulos em  $L'$  e  $G[L'] = (V, E(L'), L', l')$  o subgrafo gerador de  $G$  induzido por  $L'$ , onde  $l'$  é a função  $l$  restrita a  $E(L')$ . Denota-se por  $\mathcal{F}_G(k)$  o conjunto de todas as florestas geradoras de  $G$  com no máximo  $k$  rótulos distintos e por  $\mathcal{W}(G)$  o número de componentes conexas de  $G$ .

Note que a condição de ser floresta pode ser relaxada na definição do FGkR, permitindo-se ciclos na solução. Dado que a partir desta é possível encontrar uma floresta equivalente (mesmas componentes conexas) em tempo polinomial, retirando-se arestas suficientes para quebrar os ciclos. Sendo assim, [Cerulli et al. 2014] redefinem o problema.

**Proposição 1.** *FGkR( $G, k$ ) equivale a encontrar um subconjunto de rótulos  $L^* \subseteq L$  tal que  $|L^*| \leq k$  e  $\mathcal{W}(G[L^*]) \leq \mathcal{W}(G[L'])$ ,  $\forall L' \subseteq L, |L'| \leq k$ .*

Na proposição acima pode-se ainda trocar a condição  $|L^*| \leq k$  por  $|L^*| = \min(k, |L|)$ , diferentemente de quando se exige que  $G[L^*]$  seja uma floresta.

Usando o fato de que o número de componentes numa floresta é a diferença entre o número de vértices e arestas, mostramos que:

**Proposição 2.**  *$F^*$  é uma solução ótima para FGkR( $G, k$ ) se, e somente se,  $F^* \in \mathcal{F}_G(k)$  e  $|E(F^*)| \geq |E(F)|$ ,  $\forall F \in \mathcal{F}_G(k)$ .*

Como consequência da Proposição 2, podem ser obtidas formulações de programação inteira para o FGkR, usando restrições para eliminação de ciclos, enquanto o número de arestas escolhidas (respeitando o limite de cores) é maximizado [Figueredo 2020]. Há vários tipos de restrições para eliminar ciclos [Miller et al. 1960, Padberg and Wolsey 1983], que levam a formulações diversas.

Adicionalmente, mostramos que FGkR corresponde a uma versão ponderada do problema AGMR com restrição adicional sobre um subconjunto de cores. Para isso, definimos o GCA estendido  $G^+$ , a partir da inclusão de um vértice universal  $s$ . Precisamente,  $G^+ = (V \cup \{s\}, E \cup \{(s, v) : v \in V\}, L \cup C, l^+)$ , onde  $s \notin V$ ,  $C = \{c_v : v \in V\}$  com  $C \cap L = \emptyset$ , e  $l^+$  é tal que  $l^+(e) = l(e)$ , se  $e \in E$ , e  $l^+(e) = c_v$ , se  $e = (s, v)$ .

**Proposição 3.** *FGkR( $G, k$ ) equivale a encontrar um subgrafo gerador conexo  $G^+[L^* \cup C^*]$ , com  $L^* \subseteq L$  e  $C^* \subseteq C$ , tal que  $|L^*| \leq k$  e  $|C^*|$  seja mínimo.*

Atribuindo peso 0 às cores originais (em  $L$ ) e 1 às novas cores (em  $C$ ), a Proposição 3 estabelece uma redução entre FGkR e o problema de encontrar uma árvore rotulada com peso de cores mínimo, usando no máximo  $k$  cores originais. Essa relação permite adaptar, para o FGkR, formulações de programação inteira para o AGMR, como por exemplo aquelas propostas por [Silva 2018]. Em particular, essas formulações usam restrições de eliminação de cortes para garantir conectividade. Adicionalmente, pode-se usar restrições de fluxo para esse fim [Figueredo 2020].

## 3. Equivalência entre instâncias

[Silva 2018] mostram algumas maneiras de transformar uma instância do AGMR em outra equivalente. Aqui apresentamos adaptações dessas transformações para o FGkR. Tais estratégias são úteis para reduzir o tamanho de uma instância ou modificá-la de modo a obter uma estrutura do grafo que seja melhor explorada por alguma técnica.

Um GCA  $G = (V, E, L, l)$  é dito monocromático se todas as suas arestas possuem a mesma cor, ou seja,  $|l(E)| = 1$ . Para cada  $l \in L$ ,  $G[\{l\}]$  é, portanto, um subgrafo monocromático (induzido por  $l$ ). Cada componente conexa de um subgrafo monocromático  $G[\{l\}]$ , para qualquer  $l \in L$ , é dita uma componente monocromática de  $G$ . Note que componentes monocromáticas relativas a uma mesma cor são disjuntas, porém podem se interceptar quando geradas por cores distintas.

À luz da Proposição 1, uma solução para o FGkR pode ser vista como a união de componentes monocromáticas. Por outro lado, se arestas da cor de uma componente monocromática fossem acrescentadas ou removidas da componente, mantendo-a conexa, a solução original ainda seria ótima para o novo grafo. Nesse sentido, percebe-se que modificações podem ser feitas no grafo original sem alterar a solução ótima do problema. Em particular, a seguinte propriedade é uma adaptação daquela derivada em [Silva 2018] no contexto do AGMR.

**Proposição 4.** *Se  $G' \subseteq G$  é um subgrafo gerador livre de ciclo monocromático, maximal em arestas, então toda solução ótima para  $FGkR(G', k)$  é ótima para  $FGkR(G, k)$ .*

Na verdade, o resultado acima pode ser estendido, como segue.

**Proposição 5.** *Se  $G$  e  $G'$  são dois GCAs com as mesmas componentes monocromáticas, então o problema  $FGkR(G, k)$  e  $(G', k)$  são equivalentes.*

Pela proposição acima, contanto que as componentes monocromáticas sejam preservadas, é possível criar e remover arestas em  $G$ , restritas a uma mesma componente monocromática, sem mudar o problema. Quando tais modificações são feitas, podem aparecer arestas paralelas, de cores diferentes, entre um mesmo par de vértices, transformando assim o grafo num multigrafo. Essas transformações garantem que é possível modificar a densidade de arestas do grafo, gerando uma instância equivalente. Por outro lado, a depender da estratégia de solução empregada, a densidade pode afetar a eficiência do algoritmo. Isso acontece por exemplo com algumas formulações de programação matemática para o problema.

Outra estratégia para modificar uma instância, olhando agora para relações entre cores diferentes, baseia-se na noção de dominância, introduzida por [Silva 2018] no contexto do AGMR. Aqui, ela é adaptada para o FGkR.

Diz-se que um subconjunto de rótulos  $L' \subset L$  domina um rótulo  $c \in L \setminus L'$  (ou  $c$  é dominado por  $L'$ ), se  $\mathcal{W}(G[L']) = \mathcal{W}(G[L' \cup \{c\}])$ . Em particular, as arestas de uma cor  $c$  dominada por um subconjunto unitário  $L'$  podem ser removidas do grafo, gerando uma instância equivalente. De forma mais geral, essa noção de dominância pode ser empregada junto com processos iterativos para construir soluções para FGkR. Em outros termos, se  $L'$  é um conjunto de rótulos já escolhidos e domina uma cor  $c$ , esta pode ser descartada na formação da solução. Essa estratégia pode levar à fixação de variáveis em uma árvore de branch-and-bound, por exemplo.

#### 4. Casos polinomiais

[Cerulli et al. 2014] provaram que o FGkR é NP-Difícil. Aqui identificamos alguns casos polinomiais do problema.

**Proposição 6.** *Suponha  $G$  conexo. Se  $k_1$  é o valor de uma solução viável para  $AGMR(G)$ , então essa solução é ótima para  $FGkR(G, k)$ , para todo  $k \geq k_1$ .*

Como consequência do resultado acima, note que, se existe um algoritmo com fator de aproximação  $\alpha$  para AGMR e  $k^*(G)$  é o valor ótimo para AGRM( $G$ ), então  $FGkR(G, k)$  é polinomial para  $k \geq \alpha k^*(G)$ . Vale destacar que a heurística MVCA para o AGMR tem fator de aproximação  $1 + \ln(|V| - 1)$  [Wan et al. 2002]. Quer dizer, como esperado, a partir de um certo valor de  $k$ , o problema torna-se fácil. A seguir mostramos que o mesmo acontece até um certo valor de  $k$ .

**Proposição 7.** *Suponha  $L = \{c_1, \dots, c_{|L|}\}$  com  $|E(c_i)| \geq |E(c_{i+1})|$  para  $i = 1, \dots, |L| - 1$ . Então  $FGkR(G, k)$  é polinomial, para todo  $k$  tal que  $G[c_1, \dots, c_k]$  seja acíclico.*

O resultado acima decorre da Proposição 2, observando que  $G[c_1, \dots, c_k]$  é uma solução ótima. Além disso, como podemos sempre eliminar os ciclos monocromáticos de  $G$  (ver Proposição 4), uma consequência direta é:

**Corolário 8.**  *$FGkR(G, 1)$  é polinomial.*

Observe que a Proposição 7 estabelece um valor  $k$  até onde o problema é garantidamente polinomial. Esse valor está relacionado ao número de cores em um ciclo de  $G$ . Definimos, assim, a cintura colorida de um GCA  $G$ , denotada por  $g_c(G)$ , como o menor número de cores distintas em um ciclo de  $G$ , ou  $+\infty$  caso  $G$  seja acíclico. Esse conceito é similar ao de cintura de um grafo  $G$ , que é o comprimento do menor ciclo de  $G$ , ou  $+\infty$  caso  $G$  seja acíclico. Claramente,  $g_c(G) \leq g(G)$ . Decorre da Proposição 7 que:

**Corolário 9.**  *$FGkR(G, k)$  é polinomial, para todo  $k < g_c(G)$ .*

**Corolário 10.** *Se  $G$  é acíclico, então  $FGkR(G, k)$  é polinomial para qualquer  $k$ .*

## Referências

- Cerulli, R., Fink, A., Gentili, M., and Raiconi, A. (2014). The  $k$ -labeled spanning forest problem. *Procedia - Social and Behavioral Sciences*, 108:153 – 163.
- Chang, R.-S. and Shing-Jiuan, L. (1997). The minimum labeling spanning trees. *Information Processing Letters*, 63(5):277 – 282.
- Consoli, S. and Pérez, J. A. M. (2015). Variable neighbourhood search for the  $k$ -labelled spanning forest problem. *Electronic Notes in Discrete Mathematics*, 47:29 – 36.
- Consoli, S., Pérez, J. A. M., and Mladenović, N. (2017). Comparison of metaheuristics for the  $k$ -labeled spanning forest problem. *International Transactions in Operational Research*, 24(3):559–582.
- Figueredo, P. J. (2020). O problema da floresta geradora  $k$ -rotulada. Master's thesis, Universidade Federal do Ceará, Brasil.
- Miller, C. E., Tucker, A. W., and Zemlin, R. A. (1960). Integer programming formulation of traveling salesman problems. *J. ACM*, 7(4):326–329.
- Padberg, M. W. and Wolsey, L. A. (1983). Trees and cuts. In Berge, C., Bresson, D., Camion, P., Maurras, J., and Sterboul, F., editors, *Combinatorial Mathematics*, volume 75 of *North-Holland Mathematics Studies*, pages 511 – 517.
- Silva, T. G. d. (2018). *The Minimum Labeling Spanning Tree and Related Problems*. PhD thesis, Universidade Federal Fluminense (Brazil), Université d'Avignon et des Pays de Vaucluse (France).
- Wan, Y., Chen, G., and Xu, Y. (2002). A note on the minimum label spanning tree. *Information Processing Letters*, 84(2):99 – 101.