CAHIER DU LAMSADE

200

Décembre 2002

On the approximability of Minimum labelled spanning

trees when each color appears at most r times

J. Monnot

On the approximability of Minimum labeled spanning trees when each color appears at most r times

Jérôme Monnot,

CNRS-LAMSADE, UMR 7024, Université Paris-Dauphine,

Place du Maréchal De Lattre de Tassigny, 75775 Paris Cedex 16,

France, e_mail: monnot@lamsade.dauphine.fr

Sur l'approximabilité du problème de l'arbre couvrant minimum dans un graphe arête-coloriée dont chaque couleur apparaît au plus r fois

Résumé

Cet article traite du problème suivant: étant donné un graphe $G = (V,E)$ simple connexe dont les arêtes sont coloriées, on cherche à construire un arbre couvrant utilisant un nombre minimum de couleurs. Nous prouvons qu'un algorithme de recherche locale garantit le rapport d'approximation $(r+1)/2$ lorsque chaque couleur apparaît au plus r-fois. De plus, nous montrons que cette restriction est **APX-complet** lorsque $r \geq 3$. Finalement, nous prouvons que la généralisation où chaque arête a une liste de couleurs admissibles n'est pas un problème plus difficile à approximer que celui où chaque arête a une seule couleur.

Mots-clés: Arbre couvrant ; Approximation ; Optimisation locale ; APX-complétude.

On the approximability of Minimum labeled spanning trees when each color appears at most r times

Abstract

In this paper we consider the problem of constructing spanning trees, using a minimum number of different colors in a simple graph where the edges are labeled with colors. We prove that when each color appears at most r times, a local search algorithm yields a $(r + 1)/2$-approximation. Moreover, we show that this restriction is **APX-complete** for $r \geq 3$. Finally, we show that the generalization of this problem where each edge can be colored with a color among a list of allowed colors is equivalent to approximate.

Keywords: Spanning tree; Approximation; Local optimization; APX-completeness.

1   Introduction

In this paper we consider the problem of finding a spanning tree with minimum number of colors. Given a simple connected graph on n vertices and colors on edges (denote by $L(e)$ the color of edge e), we look for a spanning tree using a minimum number of colors. We refer to this problem as the *Minimum Labeled Spanning Tree* (in short $MinLST$). Indeed, this problem is clearly equivalent to the problem selecting a subset of colors of minimum size such that the subgraph induced by these colors is connected and span the vertex set. In the following we take this description. We also consider a natural generalization of this problem where we can choose a color for an edge e among a list of possible colors (i.e., $L(e) \in List(e)$). This generalization will be called *Minimum List Labeled Spanning Tree* (in short $MinListLST$). It is easy to see that $MinListLST$ is clearly equivalent to $MinLST$ on multigraphs. $MinLST$ has been studied firstly by Chang and Leu [3] and then by Krumke and Wirth [5], and very recently by Wan *et al.* [7]. The first authors [3] showed that this problem is **NP-hard** and it has applications in network design. The second authors [5] gave a $(2lnn + 1)$-approximation and proved that $MinLST$ cannot be approximated within $(1 - \varepsilon)lnn$ for any $\varepsilon > 0$ unless $\mathbf{NP} \subseteq \mathbf{DTIME}(\mathbf{n^{O(loglogn)}})$. Finally, the last authors [7] showed that $MinLST$ is H_{n-1}-approximable where H_n is the n-th harmonic number.

In this paper, we first prove that $MinLST$ is $(r+1)/2$-approximable when each color appears at most r-times by using a local search algorithm. Then, we show that this restriction is **APX-complete** even if the maximum degree of the graph is three and each color appears at most three times. Finally, we show that $MinListLST$ is not more difficult to approximate than $MinLST$.

2   Approximation results

In $MinLST$, an instance is given by $I = (G,L)$ where $G = (V,E)$ is a connected undirected graph without loops and multiple edges and L is a function from E to $\{1,\dots,q\}$ whereas in $MinListLST$, an input is given by $I = (G,List)$ where $G = (V,E)$ is a connected graph and $List$ is a multi-function from E to $\{1,\dots,q\}$. For any subset E' of edges, we denote by $L_{E'} = \cup_{e \in E'} L(e)$ (for $MinListLST$, $L(e) \in List(e)$) and by $G[E']$ (resp. $G[A]$ for $A \subseteq \{1,\dots,q\}$) the subgraph induced by the edge subset E' (resp., by the edge subset $L^{-1}(A) = \{e \in E : L(e) \in A\}$). The goal of these two problems consists in finding a connected subgraph $G[E']$ of G spanning V such that $|L_{E'}|$ is minimum.

In this paper, we consider the restriction of $MinLST$, denoted by $MinLST_r$ where for any $i = 1,\dots,q$, $|L^{-1}(i)| = |\{e : L(e) = i\}| \leq r$. Even this restriction is **NP-hard** for any $r \geq 3$ and the proof results from the APX-completeness given in the next section.

We study the approximability of $MinLST_r$ by a local search algorithm. The local optimization with performance ratio guarantees has been studied for instance by Ausiello and Protasi [2] or Arkin and Hassin [1]. In this paper, the neighborhood can be defined in terms of the number of used colors. In the following, a connected subgraph $G[E']$ is called a k-local optimum if we cannot delete the edges of $G[E']$ using at most k colors and add some other edges using strictly less than k colors and obtain another connected subgraph. Thus for instance, a 1-local optimum is a connected subgraph $G[E']$ using a minimal number of color (i.e., $\forall i \in L_{E'}$, the graph $G[L_{E'} \setminus \{i\}]$ is not connected). Note that if $G[E']$ is a 1-local optimum then, any spanning tree T of $G[E']$ satisfies $L_T = L_{E'}$.

It is easy to see that a 1-local optimum is a r-approximation. In the following, we show that a 2-local optimum is a $(r+1)/2$-approximation and that this ratio is tight. A 2-local optimum can be found by the following algorithm:

$[2 - 0PT]$

Input : A connected graph $G = (V,E)$ and a function $L : E \longrightarrow \{1, \ldots, q\}$;
Output : A connected subgraph $G[E']$;

        Start with an arbitrary connected subgraph $G[E']$ which span V;

        While there exists $A \subseteq L_{E'}$ with $|A| \leq 2$ and $B \subseteq \{1, \ldots, q\} \setminus L_{E'}$ with $|B| \leq 1$

        such that $G[(L_{E'} \cup B) \setminus A]$ is connected and span V do

            $G[E'] := G[(L_{E'} \cup B) \setminus A]$;

        End while

It is easy to see that the time-complexity of this algorithm is $O(n^5)$ where $n = |V|$ if we start with a connected subgraph $G[E']$ using $O(n)$ edges.

Theorem 2.1 *The* $[2-0PT]$ *algorithm is a* $\frac{r+1}{2}$*-approximation for* $MinLST_r$*. Moreover, this ratio is tight.*

Proof: Let $I = (G,L)$ be an instance of $MinLST_r$ where $G = (V,E)$ is a connected graph with n vertices and let $G[E^*]$ be an optimum solution of I. We denote by $G[E']$ the solution found by the algorithm $[2-0PT]$ and by $\{i_1, \ldots, i_k\}$ the set of colors used once in the subgraph $G[E']$ and by e_j the edge of E' with color i_j (i.e., $L_{E'}(e_j) = i_j$ and $|L_{E'}^{-1}(i_j)| = 1$ for any $j = 1, \ldots, k$).

Let T and T^* be two spanning trees of $G[E']$ and $G[E^*]$ respectively. Since $G[E']$ and $G[E^*]$ are in particular a 1-local optimum, we have $L_T = L_{E'}$ and $L_{T^*} = L_{E^*}$. Now, we work with T and T^*.

Since each color in $L_T \setminus \{i_1, \ldots, i_k\}$ is used at least 2 times and T has $n-1$ edges then, we have the following inequality for the number of colors used by the tree T:

$$|L_T| \leq k + \frac{n-1-k}{2} = \frac{n-1+k}{2} \tag{2.1}$$

On the other hand, since each color appears at most r times in T^* and T^* has $n-1$ edges, we deduce:

$$|L_{T^*}| \geq \frac{n-1}{r} \tag{2.2}$$

Moreover, we also have the inequality:

$$|L_{T^*}| \geq k \tag{2.3}$$

Assume that the previous inequality is not true and denote by $T_1, \ldots, T_{k+1}$ the $k+1$ subtrees build from T after the deletion of edges $e_1, \ldots, e_k$. Since $|L_{T^*}| < k$, there exists one color in $G[E^*]$ that uses at least two edges and such that each of these edges has one endpoint in T_i and the other one in T_j. Thus, from T we can add these two edges and delete two edges in $e_1, \ldots, e_k$ and obtain a new tree T' such that $|L_{T'}| < |L_T|$. Thus, we obtain a contradiction with the definition of a 2-local optimum.

So, using inequalities (2.1), (2.2) and (2.3), we obtain:

$$|L_T| \leq \frac{n-1}{2} + \frac{k}{2} \leq \frac{r}{2}|L_{T^*}| + \frac{1}{2}|L_{T^*}| = \frac{r+1}{2}|L_{T^*}| \tag{2.4}$$

We show in the following that this ratio is tight. Let $I = (G,L)$ be an instance with: $G = T \cup T^*$ where $T = \{(x_i,x_{i+1}),(y_i,y_{i+1}) : 0 \leq i \leq r-2\} \cup \{(x_0,v_0),(y_0,v_0)\}$ and $T^* = \{(v_0,x_i),(v_0,y_i) : 1 \leq i \leq r-1\} \cup \{(x_0,x_{r-1}),(y_0,y_{r-1})\}$. The solutions T and T^* are described in Fig. 1. Moreover, we consider the labeled function defined by: $L(x_i,x_{i+1}) = L(y_i,y_{i+1}) = i+1$, $L(v_0,x_{i+1}) = r+2 = L(x_0,x_{r-1})$, $L(v_0,y_{i+1}) = r+3 = L(y_0,y_{r-1})$ for $0 \leq i \leq r-2$, $L(x_0,v_0) = r$ and $L(y_0,v_0) = r+1$. The solutions T and T^* satisfy $|L_T| = r+1$ and $|L_{T^*}| = 2$. $\qquad\square$
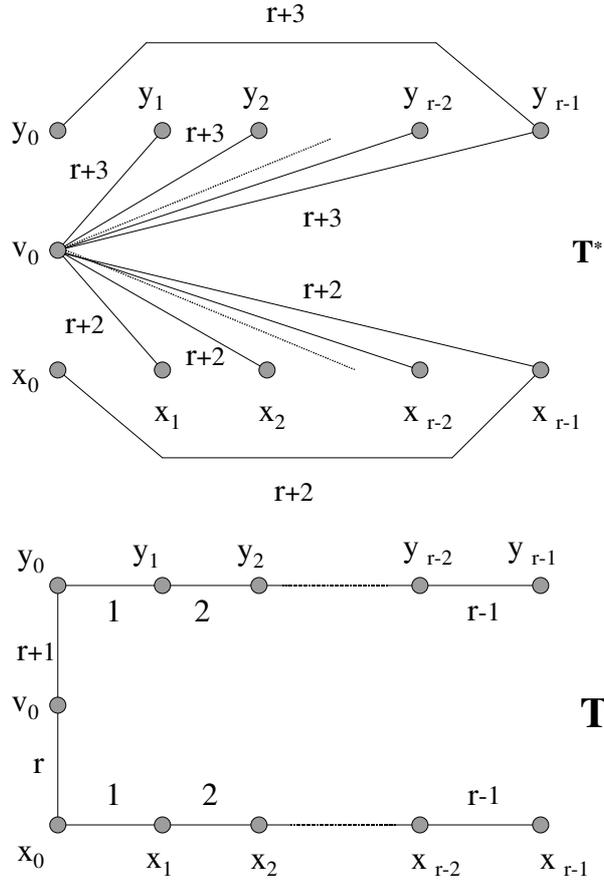
FIG. 1 – *The solutions T^* and T.*

We end this section by proving that $MinListLST$ is not more difficult to approximate than $MinLST$. Indeed, we show that $MinListLST$ and $MinLST$ are equivalent from approximation point of view. Note that this result does not hold for $MinLSTP_r$ and $MinListLST_r$.

Theorem 2.2 *$MinLST$ and $MinListLST$ are equivalent.*

Proof: Since $MinListLST$ is a generalization of $MinLST$, this problem is at least as difficult to approximate as $MinLST$. Let us prove that it is not more difficult. Let $I = (G, List)$ be an instance of $MinListLST$ where $G = (V, E)$ is a simple connected graph on n vertices and m edges and $|List_E| = q$; w.l.o.g., we assume that, $\forall e \in E$, $|List(e)| \leq m$ and there exists an edge e such that $|List(e)| \geq 2$. For any $i \leq q$, we construct an instance $I_i = (G', L_i)$ of $MinLST$ as follows: the graph $G' = (V', E')$ is common for instances $I_i = (G', L_i)$. Starting with G, we replace any edge $e = (v, w) \in E$ by the gadget F_e: assume that $List(e) = \{i_1, \ldots, i_p\}$, we add $p - 1$ vertices $A_e = \{a_{e,2}, \ldots, a_{e,p}\}$ and the edge set is $\{(v, a_{e,j}), (a_{e,j}, w) : j = 2, \ldots, p\} \cup \{(v, w)\}$. Finally, for any $j = 2, \ldots, p$, $L_i(v, a_{e,j}) = i$, $L_i(a_{e,j}, w) = i_j$ and $L_i(v, w) = i_1$. The gadget F_e is described in Fig. 2. Now, let $G[E^*]$ be an optimal solution of $MinListLST$ on I and assume that $i \in L_{E^*}$. Consider the instance I_i and construct a connected subgraph $G'[E_i^*]$ as follows: in the first step, for any edge $e = (v, w) \in E$ with $|List(e)| \geq 2$, we add $E_e = \{(v, a_{e,j}) : j = 2, \ldots, p\}$ (the edges colored by color i). Then, for $e \in E^*$ with $i_j = L(e) \in List(e)$, we add the edge (a_{e,i_j}, w) if
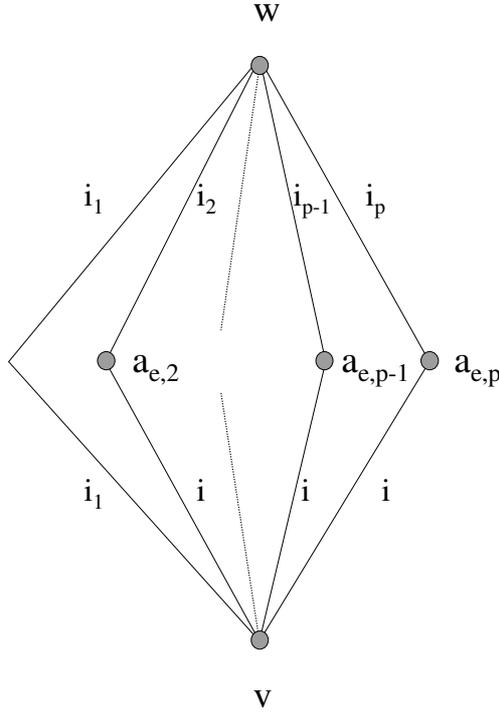
3

FIG. 2 – *The gadget F_e for the edge $e = (v,w)$ and instance $I_i = (G',L_i)$.*

$j \geq 2$ else we add the edge $e = (v,w)$. We obtain a connected subgraph $G'[E_i^*]$ of I_i verifying:

$$|L_{E_i^*}| = |L_{E^*}| \qquad (2.5)$$

Reciprocally, let $G'[E_i']$ be a connected subgraph of I_i. Observe that $G'[E_i']$ necessarily contains color i since we have assumed that there exists an edge e such that $|List(e)| \geq 2$. We construct a connected subgraph $G[E_i]$ on I satisfying $E_i = \{e : \ (v,a_{e,i_j}) \in E_i' \text{ and } (a_{e,i_j},w) \in E_i' \text{ for some } j \geq 2\} \cup \{e = (v,w) : \ (v,w) \in E_i'\}$. Moreover, if (v,a_{e,i_j}) and (a_{e,i_j},w) are in E_i', we take for the edge $e = (v,w)$ the color $L(e) = L_i(a_{e,i_j},w) = i_j$ and the color $L(e) = L_i(v,w) = i_1$ for the edge $e = (v,w) \in E_i'$. We obtain the inequality:

$$|L_{E_i}| \leq |L_{E_i'}| \qquad (2.6)$$

Lastly, we take $G[E'] = \min\{|L_{E_i}| : \ i = 1,\dots,q\}$ and the expected result follows from equality (2.5) and inequality (2.6). □

Unfortunately, the theorem 2.2 holds only for constant approximation. Thus, a performance ratio $\rho(z)$ for $MinLST$ where z is a parameter of the instance (like maximum degree or number of the vertices) does not give a performance ratio $\rho(z)$ for $MinListLST$ (the reduction between minimum dominating set problem and minimum set cover problem has the same property for $z = \Delta(G)$). Nevertheless, by using the result of Wan *et al.* [7], we can derive the following corollary:

Corollary 2.3 *$MinListLST$ is $H_{n+m(\ell_{max}-1)-1}$-approximable where $n = |V|$, $m = |E|$ and $\ell_{max} = max_{e \in E}|List(e)|$.*

4

3 Inapproximation results

In this section, we prove that $MinLST_r$ has no approximation scheme (i.e., $|L_{E'}| \leq (1 + \epsilon)|L_{E^*}|$ for any $\epsilon > 0$) unless **P=NP**. For this, we use the *Minimum Vertex Cover* problem (in short $MinVC$) that is defined as follows: given a connected graph $G = (V,E)$, we want to find a vertex subset V' with minimum size such that $\forall e = (v,w) \in E$, $v \in V'$ or $w \in V'$. $MinVC_r$ for $r \geq 3$ is the restriction of $MinVC$ where the graph G also verifies $d_G(v) \leq r$, $\forall v \in V$ and $\Delta(G) = max_{v \in V} d_G(v) \geq 3$ (since the cases VC_1 and VC_2 are polynomial).

Proposition 3.1 *For any $r \geq 3$, $MinLST_r$ is **APX-complete** even for bipartite graphs with maximum degree r.*

Proof: We construct an L-reduction (see Papadimitriou and Yannakakis [6]) from $MinVC_r$ to $MinLSTP_r$ and since $MinVC_r$ is **MaxSNP-complete** [6], we will obtain the expected result.

Let $G = (V,E)$ where $V = \{v_1, \ldots, v_n\}$ and $E = \{e_1, \ldots, e_m\}$ be an instance of $MinVC_r$. We polynomially transform G into $I = (H,L)$, instance of $MinLST_r$ in the following way: H contains the whole vertex set of G and each edge $e_k = (v_i, v_j) \in E$ is replaced by the following instance (F_k,L) where $F_k = (V_k,E_k)$ with $V_k = \{v_i, N_{k,i}, N_k, N_{k,j}, N'_{k,i}, N'_k, N'_{k,j}, v_j\}$, $E_k = \{(v_i,N_{k,i}), (N_{k,i},N_k), (N_k,N_{k,j}), (N_{k,j},v_j), (N_{k,i},N'_{k,i}), (N'_{k,i},N'_k), (N'_k,N'_{k,j}), (N'_{k,j},N_{k,j})\}$ and for $i < j$, $L(v_i,N_{k,i}) = L(N_{k,i},N'_{k,i}) = L(N'_{k,i},N'_k) = n + k$, $L(v_j,N_{k,j}) = L(N_{k,j},N'_{k,j}) = L(N'_{k,j},N'_k) = n + m + k$, $L(N_{k,i},N_k) = i$ and $L(N_{k,j},N_k) = j$.
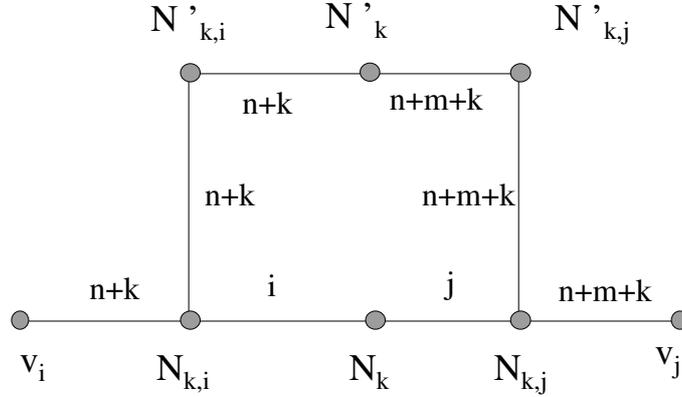
The gadget (F_k,L) is described in Fig. 3

FIG. 3 – *The gadget (F_k,L) for the edge $e_k = (v_i,v_j)$ with $i < j$.*

This instance (H,L) verifies $|V(H)| = n + 6m$, $|E(H)| = 8m$ and $|L_{E(H)}| = n + 2m$, $\Delta(G) = \Delta(H)$ where $\Delta(G)$ is the maximum degree of G and each color appears at most $\Delta(G)$ times. Moreover, the graph H is connected since G is connected as well and H is bipartite. Finally, remark that the labels $n + k$ and $n + m + k$ only occur in the gadget F_k.

Let V' be a vertex cover of G and f a function that associates at each edge of G, a vertex of V'. We construct a connected subgraph $H[E']$ of H as follows: in first step, for each gadget F_k with $i < j$, we add the edge subset $E_k = \{(v_i,N_{k,i}), (N_{k,i},N'_{k,i}), (N'_{k,i},N'_k), (N'_k,N'_{k,j}), (N'_{k,j},N_{k,j}), (N_{k,j},v_j)\}$. In the second step, if $f(e_k) = v_i$ then, we add the edge $(N_{k,i},N_k)$. We obtain a connected subgraph $H[E']$ of H spanning $V(H)$ and satisfying the equality:

$$|L_{E'}| = 2m + |V'| \tag{3.1}$$

5

Reciprocally, let $H[E']$ be a connected subgraph of H which span $V(H)$. By construction, we have $\{n+1,\ldots,2m+n\} \subset L_{E'}$ and in order to span the vertices N_k, $k = 1,\ldots,m$, we have to use $N_{k,i}$ or $N_{k,j}$. Thus, the set $V' = \{v_i : i \in L_T \setminus \{n+1,\ldots,2m+n\}\}$ is a vertex cover of G and we have:

$$|V'| = 2m - |L_{E'}| \tag{3.2}$$

Finally, let us denote by V^* and $H[E^*]$ an optimal solution of $MinVC$ and $MinLST$, respectively. So, we have:

1. $|L_{E^*}| = 2m + |V^*| \le (2r+1)|V^*|$ since $\Delta(G) \le r$
2. $|V'| - |V^*| = |L_{E'}| - |L_{E^*}|$ using equalities (3.1) and (3.2)

and the **MaxSNP-completeness** follows. $\qquad\square$

It is easy to verify that $MinLST$ is polynomial time solvable when the maximum degree is $\Delta(G) \le 2$. Moreover, we conjecture that $MinLST_r$ is polynomial for $r \le 2$. On the other hand, when $r \ge 3$, we can summarize the results on the complexity of $MinLST_r$ using the known results on $MinVC$ [4] and the proposition (3.1):

Theorem 3.2 *For any $r \ge 3$, the following hold:*

1. *$MinLST_r$ is **APX-complete** if G is triangle-free.*
2. *$MinLST_r$ is **APX-complete** if G*
3. *$MinLST_r$ is **NP-hard** if G is planar and its maximum degree is r.*

Proof: for the items 1 and 3, the proofs are immediately deduced from the proposition (3.1).

For the item 2, we start with a r-regular graph, instance of $MinVC$ and we apply the reduction of proposition (3.1). Moreover for any $k \le m$ and for each vertex of (F_k, L) with degree 2 (i.e., $N_k, N'_{k,i}, N'_k, N'_{k,j}$), we add a gadget that consists of a graph verifying that each degree is r, except for one vertex with a degree equal to one (there exists such graphs with $O(r)$ vertices) and a new label for each edge. $\qquad\square$

4   Conclusion

In future research, it will be interesting to study the complexity of $MinLST$ depending on $\Delta(G)$ and deciding if $MinLST \in$ **APX** when $\Delta(G)$ is bounded. Moreover, it will be also interesting to improve the bound $(r+1)/2$ for $MinLST_r$ by using a k-local optimum for $k \ge 3$.

Références

[1] E. M. ARKIN and R. HASSIN, *On local search for weighted k-set packing*, Mathematics of Operations Research **23** (1998), no. 3, 640–648.

[2] G. AUSIELLO and M. PROTASI, *Local search, reducibility and approximability of NP-optimization problems*, Information Processing Letters **54** (1995), no. 2, 73–79.

[3] R-S. CHANG and S-J. LEU, *The minimum labeling spanning trees*, Information Processing Letters **63** (1997), no. 5, 277–282.

[4] P. CRESCENZI and V. KANN, *A compendium of NP-optimization problems*, Available on www address: http://www.nada.kth.se/∼viggo/problemlist/compendium.html, 1998.

[5] S. O. KRUMKE and H-C. WIRTH, *On the minimum label spanning tree problem*, Information Processing Letters **66** (1998), no. 2, 81–85.

[6] C. H. PAPADIMITRIOU and M. YANNAKAKIS, *Optimization, approximation, and complexity classes*, Proc. STOC 88 (1988), 229–234.

[7] Y. WAN, G. CHEN, and Y. XU, *A note on the minimum label spanning tree*, Information Processing Letters **84** (2002), no. 2, 99–101.