Bul. Stiint. Univ. Baia Mare, Ser. B, Matematică-Informatică, Vol. XVIII(2002), Nr.1, 95 - 104 Because the C. IST problem is N. hard - a very unlikely that this problem huld be

Dedicated to Costică MUSTĂŢĀ on his 60th anniversary

APPROXIMATION RESULTS FOR THE GENERALIZED MINIMUM SPANNING TREE PROBLEM

Petrică Claudiu POP

Abstract. We consider the Generalized Minimum Spanning Tree problem denoted by GMST. It is known that the GMST problem is NP-hard. Throughout this paper we distinguish between so-called positive results and negative results in the area of approximation theory. We present an in-approximability result for the GMST problem and under special assumptions we give an approximation algorithm for the problem. 2. A negative result for the GMST prof

MSC: 90C11, 90C27, 05C05, 90B10.

Keywords: approximation algorithms, combinatorial optimization, generalized minimum spanning trees, LP relaxation.

The Generalized Minimum Spanning Tree Problem is defined on an undirected graph $G = (V_n E)$ with the nodes partitioned into m node sets called clusters. Let |V| = n and $K = \{1, 2, ..., m\}$ be the node index of the clusters. Then, $V = V_1 \cup V_2 \cup ... \cup V_m$ and $V_l \cap V_k = \emptyset$ for all $l, k \in K$ such that $l \neq k$. We assume that edges are defined only between nodes belonging to different clusters and each edge $e = (i, j) \in E$ has a nonnegative cost Theorem 2. their the assumbtion 2 + N.E. more was a supravilmation algorithm

The GMST is the problem of finding a minimum cost tree spanning a subset of nodes which includes exactly one node from each cluster. We will call a tree containing exactly one node from each cluster a generalized spanning tree. A not san bound at A. E.

The GMST problem was introduced by Myung, Lee and Tcha [6]. Feremans, Labbé and Laporte in [1] present several integer formulations of the GMST problem and compare them in terms of their linear programming relaxations, and in [2] they study the polytope The following result was proven by Myung et at. in [6].

Theorem 1. The GMSTP is NP-hard.

Because the GMST problem is NP-hard it is very unlikely that this problem could be solved by a polynomial time algorithm. At the expense of reducing the quality of the solution by relaxing some of the requirements, we can get often speed-up in the complexity. This leads us to the following definition: AT AT AT AT AT AT AT A STATE OF THE PROPERTY OF AN ADDITIONAL

Definition 1. (Approximation algorithms)

Let X be a minimization problem and $\alpha > 1$. An algorithm APP is called an α approximation algorithm for problem X, if for all instances I of X it delivers in polynomial
time a feasible solution with objective value APP(I) such that

$$APP(I) \le \alpha OPT(I)^{-1}$$
 (1)

where by APP(I) and OPT(I) we denoted the values of an approximate solution and that of an optimal solution for instance I, respectively.

Abstract, We consider the CMS1 problem is NP hard. Throughout this paper are

The value α is called the performance guarantee or the worst case ratio of the approximation algorithm APP. The closer α is to 1 the better the algorithm is.

2. A negative result for the GMST problem

For some hard combinatorial optimization problems it is possible to show that they don't have an approximation algorithm unless P = NP. In order to give a result of this form it is enough to show that the existance of an α -approximation algorithm would allow one to solve some decision problem, known to be NP-complete, in polynomial time.

Applying this scheme to the GMST problem we obtain an in-approximability result. This result is a different formulation in terms of approximation algorithms of a result provided by Myung et al. [6] which says that even finding a near optimal solution for the GMST problem is \mathcal{NP} -hard. The proof of this result similar with the proof provided in [6].

Theorem 2. Under the assumption $P \neq NP$, there is no α -approximation algorithm for the GMST problem image sort two influences and included a problem image.

3. An Approximation Algorithm for Bounded Cluster Size

The As we have seen in the previous section there exists no α -approximation algorithm for the GMST problem under the assumption $\mathcal{P} \neq \mathcal{NP}$ respect upon that the extratal modification of the GMST problem under the assumptions and the following assumptions:

A1: the graph has bounded cluster size, i.e. $|V_k| \le \rho$, for all k = 1,...m

A2: the cost function is strict positive and satisfies the cost function!triangle inequality triangle inequality, i.e. $c_{ij} + c_{jk} \ge c_{ik}$ for all $i, j, k \in V$,

a polynomial approximation algorithm for GMST problem is possible.

In this section under the above assumptions we present an approximation algorithm for the GMST problem with performance ratio 2ρ . The approximation algorithm is constructed following the ideas of Slavik [8] where the Generalized Traveling Salesman Problem and Group Steiner Tree Problem have been treated.

3.1 An integer programming formulation of GMST problem

We define for each edge $\{i, j\}$ and each node i the binary variables:

$$x_{ij} = \left\{ \begin{array}{l} 1 \;\; \text{if edge } \{i,j\} \; \text{is included in the selected subgraph} \\ \\ 0 \;\; \text{otherwise} \\ \text{meldorq ToMST not multipost A not tamix or qq A nA 2.3.} \end{array} \right.$$

$$y_i = \begin{cases} 1 & \text{if node } i \text{ is included in the selected subgraph and emission of the selection of the sel$$

The GMST problem can be formulated as the following integer programming problem: Problem IP1:ml tear entition victor dairy (8 V) = 50 places entition of the artificial problem in the contract of the contract

We use here the standard shorthand notations: for every subset S of Vala angue . 8 $E(S) = \{(i, j) \in E | i, j \in S\}, x(E(S)) = \sum_{e \in E(S)} x_e, y(S) = \sum_{j \in S} y_j \text{ and as usual the } x_e = x_e$ cutset $\delta(S)$ is defined by

$$\delta(S) = \left\{ \{i,j\} \in E \mid i \in S \text{ and } j \not \in S \right\}.$$

Condition (2) guarantees that a feasible solution contains exactly one vertex from every cluster. Condition (3) guarantees that any feasible solution is a connected subgraph. Condition (4) simply assures that any feasible solution has m-1 edges and due to the fact that the cost function is non-negative this constraint is redundant.

Consider now the linear programming relaxation of the integer programming formulation of the GMST problem. In order to do that, we simply replace conditions (5) and (6) in IP1 by new conditions: Group Straner Tree Problem jone been prouted.

meldang TSM(1) in anothermore gamma-report regions at 1.8 (7)
$$0 \le x_e \le 1, \quad \text{for all } e \in E. \tag{7}$$
 replaced at 1, when for all $i \in V$.

$$0 \le z_i \le 1$$
, for all $i \in V$.

3.2 An Approximation Algorithm for GMST problem

We assume that the assumptions A1 and A2 hold.

The algorithm for approximating the optimal solution of the GMST problem is as follows:

- Algorithm "Approximate the GMST problem"
- Input: A complete graph G = (V, E) with strictly positive cost function on the edges satisfying the triangle inequality, and with the nodes partitioned into the clusters $V_1, ..., V_m$ with bounded size, $|V_k| \le \rho$.
- Output: A tree $T \subset G$ spanning some vertices $W' \subset V$ which includes exactly one vertex from every cluster, that approximates the optimal solution to the GMST problem.
 - 1. Solve the linear programming relaxation of the problem IP1 and let $(z^*, x^*, Z_1^*) = ((z_i^*)_{i=1}^n, (x_i^*)_{i \in E}, Z_1^*)$ be the optimal solution.
 - 2. Set $W^* = \left\{i \in V | z_i^* \ge \frac{1}{\rho}\right\}$ and consider $W' \subset W^*$ with the property that W' has exactly one vertex from each cluster, and find a minimum spanning tree $T \subset G$ on the subgraph G' generated by W'.
 - Output APP = cost(T) and the generalized spanning tree T. and a red source. $E(S) = \{(e_i) | \in E_{i+1} \in S\}$ with $\{(e_i) | = \sum_{i \in S} e_i \in V(S) = \sum_{i \in S} e_i \text{ and as usual the}$

Even though the linear programming relaxation of the problem IPL has exponentially many constraints, it can still be solved in polynomial time either using ellipsoid method with a min-cut max-flow oracle [3] or using Karmakar's algorithm [5] since the linear programming relaxation can be formulated "compactly" (the number of constraints polynomially bounded) see [7].

3.3 Auxiliary results

In order to establish upper bounds on the performance ratio of the above algorithm, we now present some auxiliary results. Let now $W \subset V$ and consider the following linear program:

Problem LP2: was tall anottalianned and be measured. The drad and all stall amendons

These award appropriate
$$Z_2^{\bullet}(W) = \min$$

$$\sum_{e \in E} \frac{4^e 4 \sum_{e \in E} t_e \text{ near the ferrition are to solve all } \sum_{e \in E} t_e \text{ stone}(1)$$

s.t.
$$x(\delta(S)) \ge 1$$
, $S \subset V$, s.t. $W \cap S \ne \emptyset \ne W_{\parallel} \setminus S_{\text{sinleaded}}(9)$
 $x(\delta(i)) = 0$, $i \in V \setminus W$

$$0 \le x_e \le 1, e \in E$$
. (11)

Replacing constraints (11) with the integrality constraints $x_e \in \{0, 1\}$, the formulation obtained is the formulation of the minimum tree spanning the subset of nodes $W \subset V$ Consider the following relaxation of the problem LP2.

Problem LP3:

$$Z_3^*(W) = \min \qquad \sum_{e \in E} c_e x_e$$

$$s.t. \quad x(\delta(S)) \ge 1, \ S \subset V, \ s.t. \ W \cap S \ne \emptyset \ne W \setminus S$$

$$0 \le x_e, \qquad e \in E. \tag{12}$$

Thus we omitted constraint (10) and relaxed constraint (11).

The following result is a straightforward consequence of the parsimonius property (see [4]) if we choose $r_{ij} = 1$, if $i, j \in W$, and 0 otherwise, and $D = V^{*}(W)$

Lemma 1. The optimal solution values to problems LP2 and LP3 are the same, that is

$$Z_2^*(W) = Z_3^*(W).$$

Even though the following problem: other manufactors of the medical of the standard private the following problem: It can stand to be solved an polynomial time state using eligible to the medical of the formulation of the f

$$Z_4 = \min \qquad \qquad \sum_{e \in E} c_e x_e \qquad \qquad \text{allower yieldown} A. E. C.$$

in it regular to the value of the section
$$s.t.$$
 $x(\delta(S)) \ge 1$, $S. \subseteq V_0$, $s.t.$ $S \ne \emptyset \ne V_0$ then in the section (13)

result gravation of traditions has
$$\exists x_k \in \{0,1\}_{i=1}^n$$
 , $e \in E$, with the constraint world)

Clearly, it is the integer programming formulation of the MST (minimum spanning tree) problem. Let LP4 be the LP relaxation of this formulation, that is we simply replace the constraint (14) by the constraint $0 \le x_c \le 1$, for all $c \in E$.

Denote by Z_4^* the value of the optimal solution of the LP4. The following known result for minimum spanning trees holds:

Proposition 1.

$$L^{T}(V) \leq (2 - \frac{2}{|V|})Z_{4}^{*}$$

where $L^T(V)$ denotes the cost of the minimum spanning tree on V.

Proof: See for example [8].

Let $W \subset V$, then Proposition 5 can be easily modified to obtain:

Proposition 2.

Proof: Let (x_c) be a feasible solution to LP2. If $e \notin E(W) = \{(i,j) \mid i,j \in W\}$ implies $x_k = 0$ and using Proposition 5 we prove the inequality.

4. Performance Bounds G bos parwind to 0 bos W a C Vi I = or second ow it

Let $(y^*, x^*, Z_1^*) = ((y_1^*)_{c=1}^n, (x_c^*)_{c \in E}, Z_1^*)$ be the optimal solution to the LP relaxation fo the GMST problem. Define

$$\hat{x}_e = \rho x_e^*$$

One can easily generalize the algorithm and its analysis to the case when, in addition to distance between edges, there is a $\cos \frac{1}{2} \le \frac{1}{2} \sqrt{4}i$. Proposition this near the GMST problem can be formulated selicity allowing integer programs assert the GMST problem can be formulated.

W = $\left\{i \in V | y_i^* \geq \frac{1}{\rho}\right\} = \left\{i \in V | \widehat{y_i} = 1\right\}$. Because we need only one vertex from every cluster we delete extra vertices from W and consider $W' \subset V$ such that $|W'| = \mathbf{m}$ and W' consists of exactly one vertex from every cluster.

Since LP1 is the LP relaxation of the problem IP1, we have regertial suff of TAO substantial positions of the problem IP1, we have Z_1 and Z_2 and Z_3 and Z_4 and Z_4 and Z_5 and Z_5 and Z_6 are also an example of Z_6 and Z_6 and Z_6 are also an example of Z_6 and Z_6 and Z_6 and Z_6 are also an example of Z_6 and Z_6 and Z_6 are also an example of Z_6 and Z_6 and Z_6 are also an example of Z_6 and Z_6 and Z_6 are also an example of Z_6 and Z_6 and Z_6 are also an example of Z_6 and Z_6 are also an example of Z_6 and Z_6 are also an example of Z_6 and Z_6 and Z_6 are also an example of Z_6 and Z_6 are also an example of Z_6 and Z_6 are also an example of Z_6 and Z_6 and Z_6 are also an example of Z_6 and Z_6 are also an example of Z_6 and Z_6 are also an example of Z_6 and Z_6 and Z_6 are also an example of Z_6 and Z_6 are also are also an example of Z_6 and Z_6 are also are also an example of Z_6 and Z_6 are also are

Now let us show that $(\widehat{x}_e)_{e \in E_i}$ is a feasible solution to LP4. Indeed, $\widehat{x}_e \geq 0$ for all $e \in E$, hence condition (12) is satisfied. Let $S \subset W$ be such that $W' \cap S \neq \phi \neq W' \setminus S$ and choose some $i \in W' \cap S$. Hence $\widehat{y}_i = 1$ and $y_i^* \geq \frac{1}{\rho}$. Then we have

Under the same assumption of the GMST problem that calculation of the GMST problem that calculations is solution of the GMST problem that calculations is solution of the GMST problem of

by definition of \widehat{x}_e and the fact that the x_e^* are solution to LP1. Hence the \widehat{x}_e satisfy constraint (9) in LP3.

Therefore, and M must conseque out this "if — W inferiors has $\left\{\frac{1}{n} \leq \frac{1}{n} | V \geq \frac{1}{2}\right\} = \frac{1}{n} M$ and $\sum_{i=1}^{n} APP^{int} = \frac{1}{n} L^{T}(W^{i}) \leq \left(2 \ln \frac{2}{|W^{i}|}\right) Z_{3}^{2} = \left(2 \ln \frac{2}{|W^{i}|}\right) P_{4}^{2} \leq \left(2 \ln \frac{2}{|W^{i}|}\right) \sum_{i=1}^{n} |c_{i}\hat{x}_{i}|$ and $\sum_{i=1}^{n} \frac{2}{|W^{i}|} P_{4}^{2} = \frac{2$

And since $W'' \subset V$, what is $\mathbf{m} = |W'| \leq |V| = n$, we have proved the following. In baseline survey results and defining (x_0, x_1) as we are the section. The survey results and defining (x_0, x_1) as $\mathbf{m} = \mathbf{m}$.

Theorem 3. The performance ratio of the algorithm "Approximate GMST problem" for approximating the optimum solution to the GMST problem satisfies:

$$\frac{APP}{OPT} \le (2 - \frac{2}{\pi})\rho.$$

mode of (W)

One can easily generalize the algorithm and its analysis to the case when, in addition to distances between edges, there is a cost, say d_i , associated with each vertex $i \in V$.

In this case the GMST problem can be formulated as the following integer program:

where most verses one
$$vOPT = \min_{x \in E} \sup_{z \in E} c_i x_x + \sum_{i \in V} d_i z_i$$
 [$l \in [v, 1] \setminus \{1\} \setminus$

Suppose that $(\overline{x}, \overline{z})$ is an optimal solution. Then the optimal value OPT of this integer program consists of two parts:

A so the sol 0
$$\leq$$
 π , be shall 191 or mannes obtained a $\pi_{SS}(\mathbb{Z})$ and while ar not work and choice $\mathbb{Z} := \sum_{i \in \mathbb{Z}} d_i \overline{I_i}$ and $I_i \in \mathbb{Z}$ and choice example over $I_i \in \mathbb{Z}$ band $I_i \in \mathbb{Z}$. Hence $g_i = V = 0$ and $f_i \in \mathbb{Z}$ and $f_i \in \mathbb{Z}$ is an expectation of $f_i \in \mathbb{Z}$.

Under the same assumptions A1 and A2, the algorithm for approximating the optimal solution of the GMST problem in this case, is as follows:

- Solve the linear programming relaxation of the previous integer program and let (z*, x*) = ((z*,)*n*, (x*,)*e∈E) be the optimal solution.
- Set W* = {i ∈ V | z_i^* ≥ 1/ρ} and consider W' ⊂ W* with the property that W' has exactly one vertex from each cluster, and find a minimum spanning tree T ⊂ G on the subgraph G' generated by W'.
- Output APP = ccost(T) + vcost(T) and the generalized spanning tree T.

where by ccost(T) and vcost(T) we denoted the cost of the tree T with respect to the edges, respectively to the nodes.

Regarding the performance bounds of this approximation algorithm, using the same auxiliary results and defining (\hat{x}_e, \hat{z}_i) as we did at the beginning of this subsection, the following inequalities hold:

All amounts (\hat{x}_e, \hat{z}_i) indicates and to other communicated all (\hat{x}_e, \hat{z}_i) in the matter (\hat{x}_e, \hat{z}_i) is a monomorphism of (\hat{x}_e, \hat{z}_i) .

$$L^{T}(W') \le \rho(2 - \frac{2}{n})L_{OPT},$$

$$V^{T}(W') \le \rho V_{OPT}$$
,

where $L^T(W')$, $V^T(W')$ denote the cost of the tree T spanning the nodes of W' with respect to the edges, respectively to the nodes and as before $W' \subset W^* = \left\{ i \in V \mid z_i^* \geq \frac{1}{\rho} \right\}$, such that |W'| = m and W' consists of exactly one vertex from every cluster. For the approximation algorithm proposed in this case, the following holds:

$$\begin{array}{ll} APP &= L^T(W^l) + V^T(W^l) \leq \rho(2-\frac{2}{n})L_{OPT} + \rho V_{OPT} \\ &\leq & \rho(2-\frac{2}{n})(L_{OPT} + V_{OPT}) = \rho(2+\frac{2}{n})OPT. \end{array}$$

REFERENCES

- C. Feremans, M. Labbe and G. Laporte, A comparative Analysis of Several Formulations for the Generalized Minimum Spanning Tree Problem, Technical Report IS-MG 22, Universite Libre de Bruxelles, Belgium, 2000.
- [2] C. Feremans, M. Labbe and G. Laporte, Polyhedral Analysis of the Generalized Minimum Spanning Tree Problem, Technical Report IS-MG 1, University Libre de Bruxelles, Belgium, 2000.
- [3] M. Grotschel, L. Lovasz and A. Scrijver, The ellipsoid method and its consequences in combinatorial optimization, Combinatorica, 1, 169-197, 1981.
- [4] M.X. Goemans and D.J. Bertsimas, Survivable networks, linear programming relaxations and the parsimonious property. Mathematical Programming, 60: 145-166, 1993.
- [5] N.K. Karmarkar, A new polynomial-time algorithm for linear programming. Combinatorica, 4: 373-395, 1984.
- [6] Y.S. Myung, C.H. Lee, D.w. Tcha, On the Generalized Minimum Spanning Tree Problem. Networks, Vol 26 (1995) 231-241.
- [7] P.C. Pop, U. Faigle, W. Kern, G. Still, Relaxation Methods for the Generalized Minimum Spanning Tree Problem, submitted.
- [8] P. Slavik, On the Approximation of the Generalized Traveling Salesman Problem, submitted.

Paculty of Mathematical Sciences, University of Twente P.O. Box 217, 7500 AE Enschede, The Netherlands E-mail: pop petrica@yahoo.com

Department of Mathematics and Computer Science North University of Baia Marc, Str. Victoriei nr. 76 4800 Baia Marc ROMANIA E-mail: pop_petrica@yahoo.com

REFERENCES

- C. Peremans, M. Labbe and G. Laporte, A comparative Analysis of Several bormulations for the Generalized Minimum Spanning Tree Problem, Technocal Report 15-MG 22, University Librer de Bruyelles, Relgium, 2009.
- [2] C. Ferenaus, M. Lubbe and G. Laporte, Polyhedral Analysis of the Generalized Vinlamum Spauning Tree Problem, Technical Report IS-MG 1, University Librarde Brazeelles, Relgium, 2000.
- [3] M. Grotschel, L. Lovacz and A. Scrijver, The obpsoid method and its consequences in cond-material optimization. Combinatures, 1, 169 197, 1981.
- 4. M.X. Gormans and D.J. Bertzinnas Survivelile networks, linear programming or lacertions and the parsimons of property. Mathematical Programming, 60, 145–168, 1993.
- [5] M.K. Karmarker, A new polynomial-time algorehm for linear neggi combing, Combinations, 4, 373-395, 1984.
- [6] Y.S. Myung, C.H. Lee, D.w. T. a., On the Coerdinal Minimum Spanning Free Problem. Networks, Vol 26 (1987) 231-241.
- 7, P.C. Pop. U. Fagle, W. Kern, C. Still, Relaxation Methods for the Commitged Maintain Spanning Dec Problem, solunited.
- [8] P. Slavik, On the Approximation of the Constalized Transling Salesman Problems. 2, aboutted.