A New Approach to Generalized Network Design Problems

PETRICA POP, CORINA POP SITAR, NICOLAE POP, IOANA ZELINA
Department of Mathematics and Computer Science
North University of Baia Mare
Victoriei No. 76, Baia Mare, RO-4800
ROMANIA

Abstract: - The network design problems can be generalized in a natural way by considering a related problem relative to a given partition of the nodes of the graph into node sets (clusters), while the feasibility constraints are expressed in terms of the clusters. Several applications of the generalized network design problems arise in fields of telecommunications, facility and locations, transportation, energy distribution, etc. In this paper we present a new approach to the generalized network design problems based on distinguishing between global connections (connections between clusters) and local connections (connections between nodes from different clusters).

Key-Words: - generalized network design problems, exact algorithms, dynamic programming.

1 Introduction

Several classical combinatorial optimization problems can be cast as network design problems. Broadly speaking, a network design problem consists on identifying an optimal sub-graph F of an undirected graph subject to feasibility constraints. Examples of network design problems are: the minimum spanning tree problem, the traveling salesman problem, Steiner tree problem, etc.

These problems can be generalized in a natural way by considering a related problem relative to a given partition of the nodes of the graph into node sets, while the feasibility constraints are expressed in terms of the clusters. In this way it is introduced the class of generalized network design problems.

In the literature one finds generalized problems such as the generalized minimum spanning tree problem, the generalized traveling salesman problem, the generalized Steiner tree problem, the generalized (subset) assignment problem, etc. These Generalized problems belong to the class of NP-complete problems, are harder than the classical ones and nowadays are intensively studied due to the interesting properties and applications in the real world.

In this paper we present a new approach to the generalized network design problems based on distinguishing between global connections (connections between clusters) and local connections (connections between nodes from different clusters).

2 Applications of the Generalized Network Design Problems

The generalized network design problem have several applications to telecommunication and location problems [4], [6] and [12], railway optimization [5], etc. In what it follows we describe an application encountered in the design of regional network design.

In regional telecommunications networks, it is often necessary to design cost-effective backbone networks that connect sets of local area networks (LANs). The connection of LANs in the regional network can be established using different network designs. However, it is common to consider tree like network structures in order to lower the number of telecommunications facilities and equipment needed to establish connection between all LANs. Typically, each LAN is connected to the backbone network through one or more telecommunication centers that serve as gateways between its own users and other LANs in the region. Although each LAN may have several candidate sites that could serve as gateways, the cost of telecommunication facilities and equipment also motivates the design requirement that a single gateway is used in each LAN. This problem of designing the regional backbone network, with a tree structure spanning exactly one candidate gateway in each LAN, is known in the literature as the *generalized minimum spanning tree problem* (GMSTP) and was introduced by Myung *et al.* [7]. The minimum spanning tree problem (MST) is a special case of the

GMSTP where each cluster consists of exactly one node.

In certain situations, the impact of different candidate gateways in any given LAN can be taken into account through modification of the cost of transmission facilities needed to connect a specific candidate gateway to other gateways in the region. Candidate gateway sites in a LAN may have an incentive to be selected for the regional backbone network, and therefore they may offer compensation that can be collected if that site is selected for the backbone network. We will refer to this problem as the prize-collecting generalized minimum spanning tree problem (PC-GMSTP).

3 The Prize-Collecting Generalized Minimum Spanning Tree Problem

The minimum spanning tree (MST) problem can be generalized in a natural way by considering instead of nodes node sets (clusters) and asking for a minimum cost tree spanning *exactly* one node from each cluster. This problem is called the generalized minimum spanning tree problem (GMSTP) and it was introduced by Myung *et al.* [7]. The MST is a special case of the GMSTP where each cluster consists of exactly one node.

Meanwhile, the GMSTP have been studied by several authors with respect to heuristics and meta-heuristics, LP-relaxations, model formulations, polyhedral aspects and approximability, cf., e.g. Feremans, Labbe, and Laporte [3], Feremans [2], Pop, Kern and Still [11] and Pop [8-10].

Two variants of the generalized minimum spanning tree problem were considered in the literature: one in which in addition to the cost attached to the edges, we have costs attached also to the nodes, called the prize-collecting generalized minimum spanning tree problem, see [10] and the second one consists in finding a minimum cost tree spanning *at least* one node from each cluster, denoted by L-GMSTP and introduced by Dror [1].

Let $G = (V,E)$ be an n-node undirected graph and $V_1,\dots,V_m$ a partition of V into m subsets called *clusters* (i.e., $V=V_1\cup V_2 \cup\dots\cup V_m$ and $V_l \cap V_k= \emptyset$ for all $l,k \in\{1,\dots,m\}$ with $l \neq k$). We assume that edges are defined between all nodes which belong to different clusters. We denote the cost of an edge $e=(i,j) \in E$ by c_{ij} or by $c(i,j)$, the cost of a node $i\in V$ by d_i and the costs of the edges and nodes are chosen integers.

The *prize-collecting generalized minimum spanning tree problem* asks for finding a minimum-cost tree T spanning a subset of nodes which includes exactly one node from each cluster V_k, $k\in\{1,\dots,m\}$. We will call such a tree a *generalized spanning tree*.

Let G' be the graph obtained from G after replacing all nodes of a cluster V_k with a supernode representing V_k. We will call the graph G' the global graph. For convenience, we identify V_k with the supernode representing it. Edges of the graph G' are defined between each pair of the graph vertices $\{V_1,...,V_m\}$.

The *local-global approach* to the prize-collecting generalized minimum spanning tree problem aims at distinguishing between *global connections* (connections between clusters) and *local connections* (connections between nodes from different clusters). As we will see, having a global tree connection of the clusters it is rather easy to find the corresponding best (with respect to cost minimization) generalized spanning tree.

Based on this new approach we will provide an exact exponential time algorithm, a strong mixed integer formulation and a solution procedure based on this formulation of the PC-GMSTP.

4 An Exact Algorithm for the Prize-Collecting Generalized Minimum Spanning Tree Problem

In this section, we present an algorithm that finds an exact solution to the PC-GMSTP based on dynamic programming.

Given a spanning tree of the global graph G', which we shall refer to as a *global spanning tree*, we use dynamic programming in order to find the corresponding best (with respect to cost minimization) generalized spanning tree.

Fix an arbitrary cluster V_{root} as the root of the global spanning tree and orient all the edges away from vertices of V_{root} according to the global spanning tree. A directed edge $\langle V_k,V_l\rangle$ of G', resulting from the orientation of edges of the global spanning tree defines naturally an orientation $\langle i,j\rangle$ of an edge $(i,j)\in E$ where $i\in V_k$ and $j\in V_l$. Let v be a vertex of cluster V_k for some $1\leq k\leq m$. All such nodes v are potential candidates to be incident to an edge of the global spanning tree. On the graph G, we denote by T(v) denote the subtree rooted at such a vertex v from G; T(v) includes all vertices reachable from v under the above orientation of the edges of G based on the orientation of the edges of the global spanning tree. The *children* of $v\in V_k$, denoted by C(v), are those vertices $u\in V_l$ which are heads of the directed edges $\langle v,u\rangle$ in the orientation. The leaves of the tree are those vertices that have no children.

Let W(T(v)) denote the minimum weight of a generalized subtree rooted at v. We want to compute:

$$\text{Min}_{r\in V_{root}} W(T(r)).$$

We are now ready for giving the dynamic programming recursion to solve the subproblem W(T(v)). The initialization is:

$W(T(v))=0$ if $v\in V_k$ and V_k is a leaf of the global spanning tree.

To compute W(T(v)) for an interior to a cluster vertex $v\in V$, i.e., to find the optimal solution of the subproblem W(T(v)), we have to look at all vertices from the clusters V_l such that $C(v) \cap V_l \neq \emptyset$. If u denotes a child of the interior vertex v, then the recursion for v is as follows:

$$W(T(v)) = \Sigma_{l,\, C(v) \cap Vl \neq \emptyset} \min_{u\in Vl} [c_{vu}+d_u+W(T(u))].$$

Hence, for fixed v we have to check at most n vertices. Consequently, for the given global spanning tree, the overall complexity of this dynamic programming algorithm is $O(n^2)$. Since by Cayley's formula, the number of all distinct global spanning trees is m^{m-2}, we have established the following.

Theorem 1. There exists a dynamic programming algorithm which provides an exact solution to the generalized minimum spanning tree problem in $O(m^{m-2}n^2)$ time, where n is the number of nodes and m is the number of clusters in the input graph.

Clearly, the above is an exponential time algorithm unless the number of clusters m is fixed.

5 A new model formulation of the PC-GMSTP

The PC-GMSTP can be formulated as an integer program in many different ways based on trees properties, or by considering flow variables, cf. [2-3] and [7-9]. In this section we will present a new model formulation of the PC-GMSTP as a mixed integer program based on the new approach which distinguishes between *global* variables, i.e. variables modeling the inter-cluster (global) connections, and *local* ones, i.e. expressing whether an edge is selected between two clusters linked in the global graph G'.

In order to formulate the PC-GMSTP as a mixed integer program we introduce the variables $x_e\in \{0,1\}$, $e \in E$ and $y_i \in\{0,1\}$, $i \in V$, to indicate whether an edge e, respectively a node i is contained in the spanning tree and in addition the variables z_{ij} $i,j\in\{1,...,m\}$ to describe the inter-cluster (global) connections. Hence $z_{ij} =1$ if cluster V_i is connected to cluster V_j and $z_{ij} =0$ otherwise.

We assume that z represents a spanning tree. The convex hull of all these z-vectors is generally known as the spanning tree polytope on the global graph G'. Following Yannakakis [13] this polytope, denoted by
P_{MST} , can be represented by the following polynomial number of constraints:

$$\sum_{\{i,j\}} z_{ij} = m-1$$
$$z_{ij} = \lambda_{kij} + \lambda_{kji}, \text{ for } 1 \leq k,i,j \leq m \text{ and } i \neq j \quad (1)$$
$$\sum_j \lambda_{kij} = 1, \text{ for } 1 \leq k,i,j \leq m \text{ and } i \neq k \quad (2)$$
$$\lambda_{kkj} = 0, \text{ for } 1 \leq k,j \leq m \quad (3)$$
$$z_{ij}, \lambda_{kij} \geq 0, \text{ for } 1 \leq k,i,j \leq m.$$

where the variables λ_{kij} are defined for every triple of nodes k,i,j, with $i \neq j \neq k$ and their value for a spanning tree is: $\lambda_{kij} =1$ if j is the parent of i when we root the tree at k and 0 otherwise.

The constraints (1) mean that an edge (i,j) is in the spanning tree if and only if either i is the parent of j or j is the parent of i; the constraints (2) mean that if we root a spanning tree at k then every node other than node k has a parent and finally constraints (3) mean that the root k has no parent.

If the vector z describes a spanning tree on the global graph G', the corresponding best (with respect to cost minimization) *generalized spanning tree* $(x,y)\in\{0,1\}^{|E| \times |V|}$ can be obtained either by using the dynamic programming algorithm presented in Section 4, or by solving the following 0-1 programming problem:

$$\begin{aligned}
&\text{Min} \sum_{e\in E} c_e x_e + \sum_{i\in V} d_i y_i \\
&\text{s.t. } y(V_k) = 1,\ \forall\, k\in\{1,...,m\} \\
&\quad x(V_l,V_r) = z_{lr},\ \forall\, l,r\in\{1,...,m\},\ l \neq r \\
&\quad x(i,V_r) \leq y_i,\ \forall\, r\in\{1,...,m\},\ \forall\, i\in V \setminus V_r \\
&\quad x_e, y_i \in\{0,1\},\ \forall\, e=(i,j)\in E,\ \forall\, i\in V,
\end{aligned}$$

where $x(V_l,V_r) = \sum_{l\in Vl, j\in Vr} x_{ij}$ and $x(i,V_r)=\sum_{j\in Vr} x_{ij}$.

For given z, we denote the feasible set of the linear programming relaxation of this program by $P_{local}(z)$. The following result holds:

Proposition 2. If z is the 0-1 incidence vector of a spanning tree of the contracted graph then the polyhedron $P_{local}(z)$ is integral.

The proof of this result can be found in [11]. The observations presented so far lead to our formulation, called *local-global formulation* of the PC-GMSTP as a 0-1 mixed integer programming problem, where only the global variables z are forced to be integral:

$$\begin{aligned} &\text{Min } \textstyle\sum_{e\in E} c_e x_e + \sum_{i\in V} d_i y_i \\ &\text{s.t.} \quad z \in P_{MST} \\ &\qquad (x,y) \in P_{local}(z) \\ &\qquad z_{lr} \in \{0,1\}, \ \ 1 \le l,r \le m . \end{aligned}$$

This new formulation of the PC-GMST problem was obtained by incorporating the constraints characterizing P_{MST}, with $z \in \{0,1\}$, into $P_{local}(z)$.

6 Solution Procedure and Numerical Results

We consider the relaxation of the mixed integer programming formulation described in the previous Section obtained by choosing randomly one cluster V_k and rooting the global tree only at the root k.

If an optimal solution of this relaxation produces a generalized spanning tree, then we have given an optimal solution of the PC-GMSTP. Otherwise, we either choose another root, or add a second root and proceed in this way till we get the optimal solution of the PC-GMST problem. We call this procedure the *rooting procedure*.

Our algorithms have been coded in C and compiled with a HP-UX cc compiler. For solving the linear and mixed integer programming problems we used CPLEX 6.5. The computational experiments were performed on a HP 9000/735 computer with a 125 Mhz processor and 144 Mb memory.

We considered an undirected graph G = (V,E) having n nodes which are partitioned into m clusters such that each cluster contains the same number of nodes. Edges are defined between all the nodes from different clusters. The costs of the edges and nodes were generated randomly in the interval [0,100]. For each type of instance we considered five trials.

In the Table 1 we present the computational results for solving the problem using our rooting procedure.

m	n	LB/OPT	CPU	No. of roots
8	32	100	0.0	1
8	80	100	0.7	1
10	40	100	0.8	1
10	100	100	3.7	2
12	48	100	1.9	1
12	120	100	14.9	2
15	60	100	2.9	1
15	150	100	42.3	2
18	54	100	29.8	2
18	180	100	150.7	2
20	60	100	40.3	2
20	200	100	223.1	2
25	75	100	24.6	2
25	200	100	312.5	2
30	90	100	44.2	2
30	240	100	350.1	3
40	120	100	78.9	2
40	160	100	1841.3	3

Table 1. Computational results

The first two columns in the table give the size of the problem: the number of clusters m and the number of nodes n. The next three columns describe the rooting procedure and contain: the lower bounds obtained as a percentage of the optimal value of the PC-GMSTP (LB/OPT), the computational times (CPU) in seconds for solving the PC-GMSTP to optimality and the minimum number of clusters chosen as roots, in order to get the optimal solution of the problem.

As it can be seen, in all the instances that we considered, for graphs with nodes up to 240, the optimal solution of the PC-GMSTP has been found by using our rooting procedure. These numerical experiences with the local-global formulation of the PC-GMSTP are very promising.

7 Conclusion

We described a new approach to tackle the generalized network design problems based on distinguishing between global connections (connections between clusters) and local connections (connections between nodes from different clusters). In the case of the prize-collecting generalized minimum spanning tree problem based on this approach we provided an exact exponential time algorithm, a strong mixed integer programming formulation and a solution procedure. In a similar way using the new approach it is possible to provide exact algorithms and strong mixed integer programming formulations for other generalized network design problems, such as the generalized traveling salesman problem, generalized assignment problem, etc.

References:

[1] M. Dror, M. Haouari and J. Chaouachi, Generalized Spanning Trees, *European Journal of Operational Research*, Vol. 120, 2000, pp. 583-592.

[2] C. Feremans, Generalized Spanning Trees and Extensions, PhD thesis, *Universite Libre de Bruxelles*, Belgium, 2001.

[3] C. Feremans, M. Labbe and G. Laporte, A Comparative Analysis of Several Formulations of the Generalized Minimum Spanning Tree Problem, *Networks*, Vol. 39, Issue 1, 2002, pp. 29-34.

[4] M. Gerla and L. Frata, Tree structured fiber optics MAN's}, *IEEE J.Select. Areas Comm.* SAC-6, 1988, pp. 934-943.

[5] G. Hadjicharalambous, P. Pop, E. Pyrga, G, Tsaggouris, C. Zaroliagis, The Railway Traveling Salesman Problem, in Algorithmic Methods for Railway Optimization, Proc. ATMOS 2004, to appear in *Lecture Notes in Computer Science*.

[6] G. Laporte, A.A.Vaziri and C. Sriskandarajah, Some applications of the generalized traveling salesman problem, *J. Oper. Res. Soc.*, Vol. 47, 1996, pp. 1461-1467.

[7] Y.S. Myung, C.H. Lee, D.w. Tcha, On the Generalized Minimum Spanning Tree Problem. *Networks*, Vol. 26, 1995, pp. 231-241.

[8] P.C. Pop, *The Generalized Minimum Spanning Tree Problem*, PhD thesis, Twente University Press, the Netherlands, 2002.

[9] P.C. Pop, New Models of the Generalized Minimum Spanning Tree Problem, *Journal of Mathematical Modelling and Algorithms*, Vol. 3, Issue 2, 2004, pp. 153-166.

[10] P.C. Pop, On the Prize-Collecting Generalized Minimum Spanning Tree Problem, to appear in *Annals of Operations Research*.

[11] P.C. Pop, W. Kern and G. Still, A New Relaxation Method for the Generalized Minimum Spanning Tree Problem, *European Journal of Operational Research*, 170, 2006, pp. 900-908.

[12] J.J. Prisco, Fiber optic regional area networks in New York and Dallas, *IEEE J. Select. Areas Comm.*, SAC-4, 1986, pp 750-757.

[13] M. Yannakakis, Expressing combinatorial optimization problems by linear programs, *Journal of Computer and System Sciences,* 43, 1991, pp. 441-466.