Exact solution for branch vertices constrained spanning problems

Massinissa Merabet, Sylvain Durand, Miklós Molnár

Exact solution for branch vertices constrained spanning problems

Merabet Massinissa[1] Sylvain Durand[2] Miklos Molnar[3]

LIRMM
Montpellier 2 university
Montpellier, France

Abstract

Given a connected graph G, a vertex v of G is said to be a branch vertex if its degree is strictly greater than 2. The Minimum Branch Vertices Spanning Tree problem (MBVST) consists in finding a spanning tree of G with the minimum number of branch vertices. This problem has been well studied in the literature and has several applications specially for routing in optical networks. However, this kind of applications do not explicitly impose a sub-graph as solution. A more flexible structure called hierarchy is proposed. Hierarchy, which can be seen as a generalization of trees, is defined as a homomorphism of a tree in a graph. Since minimizing the number of branch vertices in a hierarchy does not make sense, we propose to search the minimum cost spanning hierarchy such that the number of branch vertices is less than or equal to an integer R. We introduce the Branch Vertices Constrained Minimum Spanning Hierarchy (BVCMSH) problem which is NP-hard. The Integer Linear Program (ILP) formulation of this new problem is given. To evaluate the difference of cost between trees and hierarchies, we confront the BVCMSH problem to the Branch Vertices Constrained Minimum Spanning Tree (BVCMST) problem by comparing its exact solutions. It appears from this comparison that when $R \leq 2$, the hierarchies improve the average cost from more than 8% when $|V_G| = 20$ and from more than 12% when $|V_G| = 30$.

Keywords: Spanning tree, homomorphism, branch vertices constraint.

[1] Email: merabet@lirmm.fr
[2] Email: sylvain.durand@lirmm.fr
[3] Email: molnar@lirmm.fr

1   Introduction

Multicast routing applied in optical networks provide several research problems on spanning tree. In optical networks, the ability of splitting the light signal is limited. Given a connected graph G, a vertex v of G is said to be a branch vertex if its degree is strictly greater than 2. The Minimum Branch Vertices Spanning Tree problem (MBVST) consists in finding a spanning tree of G with the minimum number of branch vertices [CGI09]. All researches on branch vertices constrained spanning problems are based on spanning trees, Conversely the routing do not explicitly impose a sub-graph as solution. A more flexible structure is proposed in [Mol08]. In contrast with trees, this structure (called hierarchy) is not a sub-graph but a homomorphism of a tree in a graph. Surprisingly, in connected graphs, there always exists a spanning hierarchy without branch vertices.

In this paper we introduce the Branch Vertices Constrained Minimum Spanning Hierarchy (BVCMSH) problem which consists in searching the minimum cost spanning hierarchy such that the number of branch vertices is less than or equal to an integer R. To evaluate the difference of cost between trees and hierarchies, we confront the BVCMSH problem to the Branch Vertices Constrained Minimum Spanning Tree (BVCMST) problem. As we will show, the interest of the hierarchy concept to solve these problems is obvious. Often the branch vertices constraints exclude spanning tree to be solution whereas the coverage is possible with a hierarchy. Moreover, the spanning hierarchy always achieves lower cost than the spanning tree's cost since a spanning tree is a special hierarchy. As proved in [GHSV02], it is NP-hard to find a branch vertices constrained spanning tree. We prove in section 2 that BVCMSH is also NP-hard for any value R. In order to compare the cost of solutions, we compute exact solution for both BVCMST and BVCMSH based on ILP formulation. The improvement provided by the use of hierarchy is significant (more than 12% when the graph has more than 30 vertices among which less than 3 are allowed to be branch vertices).

The rest of the paper is organized as follows. First, the branch BVCMSH problem is formulated in Section 2. In Section 3, ILP formulation is developed to compute the optimal hierarchy. Simulations are done in Section 4 to compare optimal hierarchies and optimal trees. Finally the paper is concluded in Section 5.

2   Problem formulation

A hierarchy is not always a sub-graph. It is a graph related structure obtained by a homomorphism of a tree in a graph. Remember that in graphs, a homo-

morphism can be defined as follows. Let $Q = (W, F)$ and $G = (V, E)$ be two (undirected) graphs. An application $h : W \to V$ mapping a vertex in V to each vertex in W is a homomorphism if the mapping preserves the adjacency: $(u, v) \in F$ implies $(h(u), h(v)) \in E$ [HZ94]. If Q is a connected graph without cycle (a tree) then the triple (Q, h, G) defines a *hierarchy in G*.

Figure 1 shows an example of hierarchy. Each vertex of the tree Q is associated with a unique vertex of the graph G. But, some vertices of G are mapped to several vertices in Q. A vertex in Q can be labelled by the vertex in G with which it is associated. To distinguish the occurrences related to a same vertex v in G, we will use the labels $v^1, v^2, ..., v^k$ if needed. If the application h is injective, then the hierarchy corresponds to a tree in G.
Let u be a vertex in Q such that the mapping associates $v \in V_G$ ($h(u) = v$) to this vertex in the hierarchy. The degree $d_H(v^i)$ of the vertex occurrence v^i in H is : $d_H(v^i) = d_Q(v^i)$ where $d_Q(v^i)$ gives the degree of v^i in Q.

Since the spanning structure for routing must be connected but is not necessary a sub-graph, the branch vertices constrained minimum spanning hierarchy problem is defined as follows. Let $G = (V, E, c)$ be the weighted graph of the network with positive edge costs $c(e), \forall e \in E$.

Definition 2.1 Given a weighted graph $G = (V, E, c)$ and an integer R, the BVCMSH problem consists in finding a minimum cost spanning hierarchy $H = (Q, h, G)$ of G such that the number of branch vertices in H is less than or equal to R.
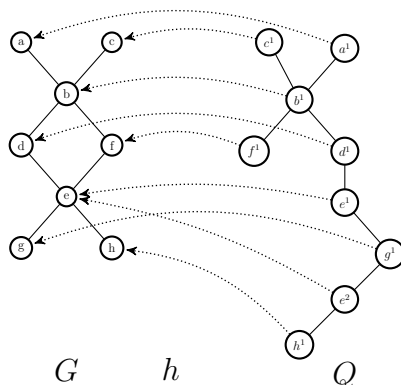

Figure 1. Mapping of vertices for a hierarchy

Figure 1 illustrates how a hierarchy can satisfy the branch vertices constraint. In all connected spanning trees of G both b and e are branch vertices

(degree strictly greater than 2). However the number of branch vertices in H is equal to 1 because of duplication of the vertex e. So there is no feasible spanning tree of G for branch vertices constrained spanning problem with $R = 1$ but there is at least one feasible hierarchy in G. There is even a hierarchy without branch vertex since any traversal of G is a hierarchy.

For both tree and hierarchy structures, finding an optimal branch vertices constraint spanning structure is NP-hard for any value of R. When $R = 0$ and all edge costs are equals, solving the problem allows to find a Hamiltonian path if it exists. When $R \geq 1$, let $G = (V, E)$ be a graph and $v \in V$ be a vertex of G (randomly chosen). Make $2 * R$ copies of G, all edges having cost 1 in these copies. Build a chain of R vertices connected by edges of cost 1. Each of these vertices is connected to 2 copies of G by adding an edge of cost K (with $K > R * |E|$) to v. Let G' be the graph obtained. In any optimal solution of BVCMSH with a bound of R, all the vertices of the chain are branch vertices (otherwise, the edges of cost K would be used more than once). Thus, there is a solution of cost $(R - 1) + 2 * (|V| - 1)$ if and only if there is an Hamiltonian path starting from v in G.

3 ILP formulation of the BVCMSH problem

In this section, an Integer Linear Program is solved to find the branch vertices constrained optimal hierarchy. In our model the connectivity is preserved by flow formulation. Thus, the initial graph has to be transformed in a symmetric oriented graph. Furthermore, a flow can transit more than once in each direction of an arc. Then each arc has to be duplicated $|V_G| - 1$ times (Figure 2) since $|V_G| - 1$ is the maximal number of times that an edge can be used in a hierarchy. This model allows to respect the branch vertices constraint. We introduce for each vertex $m \in V_G$ a binary variable Y_m equal to 1 if m is a branch vertex and 0 otherwise. Obviously, solving our linear program does not give a solution for the BVCMSH problem since the relation must be done between edges duplication and vertices duplication. A polynomial step is added to construct the optimal hierarchy.
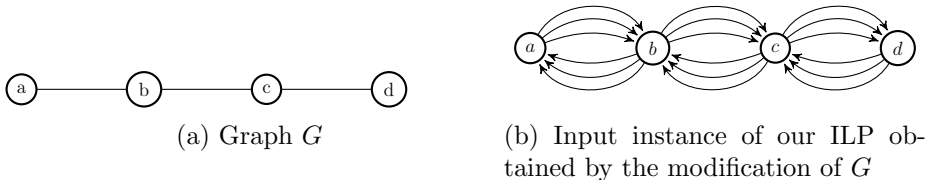

(a) Graph G

(b) Input instance of our ILP obtained by the modification of G

Figure 2. Our transformation of graph to feasible input of the ILP

In the following the linear program is presented.

Network parameters:

$In(m)$: The set of vertices which have an outgoing arc leading to node m.

$Out(m)$: The set of vertices which can be reached from m.

$C(m,n)$: The cost of the arc linking node m to node n. All duplications of this link have the same cost.

β : Global upper bound on the number of duplications of each vertex (equals to $|V_G| - 1$).

ILP variables:

$L_i(m,n)$: Binary variable. Equals to 1 if the occurrence i of the arc (m,n) is in the output graph, 0 otherwise.

$F_i(m,n)$: Commodity flow variable. Denotes the quantity of flow transiting on the occurrence i of the arc (m,n).

$Y(m)$: Binary variable. Equals to 1 if the vertex m is a branch vertex, 0 otherwise.

The objective of our problem is to minimize the total cost of edges belonging the hierarchy structure:

$$Minimize : \sum_{m \in V_G} \sum_{n \in Out(m)} \sum_{i=1}^{\beta} L_i(m,n) * C(m,n) \qquad (1)$$

This objective function is subject to a set of constraints.

Degree constraints:

$$\sum_{n \in Out(s)} \sum_{i=1}^{\beta} L_i(s,n) - \sum_{n \in In(s)} \sum_{i=1}^{\beta} L_i(n,s) - 2 \leq (|V_G| - 1) * Y(s) \qquad (2)$$

$$\sum_{n \in Out(m)} \sum_{i=1}^{\beta} L_i(m,n) - \sum_{n \in In(m)} \sum_{i=1}^{\beta} L_i(n,m) \leq (|V_G| - 1) * Y(m) \qquad \forall m,n \in V \backslash \{s\}$$
$$(3)$$

Constraints (2) and (3) ensure that, if $Y(m)$ is equal to 0 then for each vertex except the source, the number of successors is at most equal to the number of predecessors. Otherwise, the number of successors is as large as necessary (bounded by $|V_G| - 1$) .

$$\sum_{m \in V_G} Y(m) \leq R \qquad\qquad (4)$$

Constraint (4) ensures the respect of the branch vertices bound R.

$$L_i(m, n) \leq 1 - Y(m) \qquad\qquad ,\forall m \in V, \quad \forall n \in Out(m), \quad \forall i > 1 \qquad (5)$$

Constraint (5) ensures that if a vertex m is a branch vertex then this vertex is duplicated only once.

Connectivity constraints:
In order to guarantee the connectivity of the output graph, we have introduced some flow constraints adapted to the specificities of degree constrained hierarchies.

$$\sum_{i=1}^{\beta} \sum_{n \in Out(s)} F_i(s, n) = \sum_{i=1}^{\beta} \sum_{n \in In(s)} F_i(n, s) + |V| - 1 \qquad\qquad (6)$$

$$\sum_{i=1}^{\beta} \sum_{n \in In(m)} F_i(n, m) = \sum_{i=1}^{\beta} \sum_{n \in Out(m)} F_i(m, n) + 1 \qquad\qquad \forall m \in V - \{s\} \quad (7)$$

Constraints (6) and (7) ensures that each vertex except the source "consumes" one and only one unit of flow. This constraint also guarantees that each vertex is reachable from the source s. the flow emitted by the source is equal to $|V| - 1$.

$$F_i(n, m) \geq L_i(m, n) \qquad\qquad \forall e = (m, n) \in E \qquad\qquad (8)$$

$$F_i(m, n) \leq (|V| - 1) * L_i(m, n) \qquad\qquad \forall e = (m, n) \in E \qquad\qquad (9)$$

Constraints (8) and (9) allow each arc to carry non-zero flow if and only if it is used in the output graph. The value of this flow should not exceed the total flow emitted by the source.

4 Experimentation

In order to demonstrate the advantage of the proposed hierarchy structure, simulation is conducted to compare it with the spanning tree structure. ILP formulations are implemented in C with GLPK package [Mak09] by using random graphs generated with NetGen [KNS74].
We consider two values for number of vertices of random graph: $|V| \in \{20, 30\}$. The density value $d = |E|/|V|$ is fixed to 2. Graphs with this density are considered as sparse graphs. Sparse graphs are more accurately modelling real applications in networks. Thirty feasible instances for the BVCMST with random edge costs are generated for each value of $|V|$. Then ILP are

solved to search the optimal branch vertices constrained spanning tree and the optimal branch vertices constrained spanning spanning hierarchy with $R \in \{1, 2, 3, 4, 5\}$. Since it is the first exact method for solving the BVCMSH problem, we do not focus on the time complexity of our method. Our goal is mainly to evaluate the improvement obtained by using a hierarchy instead of a tree. Nevertheless, the average running time to solve the BVCMSH problem is reasonable (about 14 minutes for $|V_G| = 20$ using a 2.20GHz Intel(R) Core(TM) i3 with 4GB of RAM running the Linux operating system).

$|V_G| = 20$					
R	1	2	3	4	5
Average cost of BVCMST	7711.31	7003.42	6544.86	6280.93	6084.01
Average cost of BVCMSH	6787.38	6423.11	6215.03	6120.51	6084.01
Improvement	11.98%	8.28%	5.03%	2.55%	1.26%
$|V_G| = 30$					
R	1	2	3	4	5
Average cost of BVCMST	13075.17	11041.47	10045.07	9500.05	9214.98
Average cost of BVCMSH	10537.10	9645.13	9005.65	8900.74	8850.68
Improvement	19.41%	12.64%	10.35%	6.31%	3.95%

Table 1

Hierarchies average cost versus trees average cost
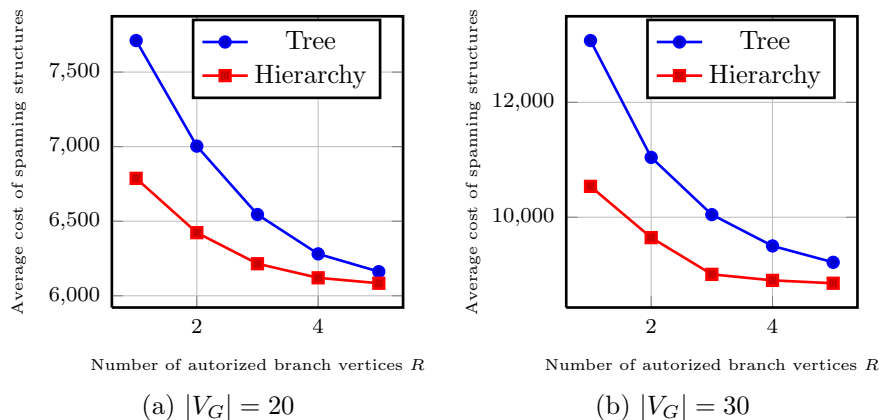

(a) $|V_G| = 20$

(b) $|V_G| = 30$

Figure 3. Optimal spanning tree versus optimal spanning hierarchy

The numerical results of the cost are presented in Table 1. Since an optimal tree for BVCMST is a feasible solution for BVCMSH, the optimal hierarchy cost is always lower than the tree cost.

The cost improvement with hierarchies is clearer in Figures 3a and 3b which represent the average costs of trees and hierarchies for $|V_G| = 20$ and $|V_G| = 30$ respectively. As shown in these figures, the average hierarchy cost is lower than tree cost in any situations. When $R \leq 2$, it can be observed that the hierarchies improve the average cost from more than 8% when $|V_G| = 20$ and from more than 12% when $|V_G| = 30$. As expected, the improvement

grows with the size of the graph and decreases when the allowed number of branch vertices grows.

5 Conclusion

Our experiments show that when the the spanning structure should not necessarily be a sub-graph, hierarchies can be a good alternative to trees. Besides the fact that for any instance and any positive R there always exists a feasible solution for the MVCMSH, the cost of the optimal solution (hierarchy) of the BVCMSH problem is lower than the cost of the optimal solution (tree) of the BVCMST problem. This observation is obvious since trees are a special case of hierarchies, but these experiments give a more accurate assessment of the cost improvement. We can particularly note that when $R \leq 2$, the hierarchies improve the average cost from more than 8% when $|V_G| = 20$ and from more than 12% when $|V_G| = 30$. For future work, it could be interesting to study more this structure in order to explore the approximation of the MBCMSH problem.

References

[CGI09] R Cerulli, M Gentili, and A Iossa. Bounded-degree spanning tree problems: models and new algorithms. *Comput. Optim. Appl*, 42:353–370, April 2009.

[GHSV02] L Gargano, P Hell, L Stacho, and U Vaccaro. Spanning Trees with Bounded Number of Branch Vertices. In *Proceedings of the 29th International Colloquium on Automata, Languages and Programming*, ICALP '02, pages 355–365, London, UK, UK, 2002. Springer-Verlag.

[GJ79] M R Garey and D S Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman & Co, New York, NY, USA, 1979.

[HZ94] P Hell and X Zhu. Homomorphisms to oriented paths. *Discrete Mathematics*, 132(13):107 – 114, 1994.

[KNS74] D Klingman, A Napier, and J Stutz. A program for generating large scale capacitated assignment, transportation, and minimum cost flow network problems. *Management Science*, 20(5):814–821, 1974.

[Mak09] O Makhorin. *GNU Linear Programming Kit (GLPK) v 4.38*, gnu project edition, May 2009.

[Mol08] M Molnár. Hierarchies for Constrained Partial Spanning Problems in Graphs. Technical report PI-1900, 2008.