European Journal of Operational Research 254 (2016) 755–768

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Discrete Optimization

New formulations for the elementary shortest-path problem visiting a given set of nodes

Rafael Castro de Andrade*

Universidade Federal do Ceará, Departamento de Estatística e Matemática Aplicada, Campus do Pici, bloco 910. CEP 60.440–900 Fortaleza, CE, Brazil

ARTICLE INFO

Article history:
Received 17 October 2013
Accepted 5 May 2016
Available online 13 May 2016

Keywords:
Combinatorial optimization
Shortest-path visiting given nodes
Compact extended formulation
Primal-dual formulation

ABSTRACT

Consider a directed graph $G = (V, A)$ with a set of nodes V and a set of arcs A, and let c_{uv} denote the length of an arc $uv \in A$. Given two nodes s and t of V, we are interested in the problem of determining a shortest-path from s to t in G that must visit only once all nodes of a given set $P \subseteq V - \{s, t\}$. This problem is NP-hard for $P = V - \{s, t\}$. In this paper, we develop three new compact formulations for this problem. The first one is based on the spanning tree polytope. The second model is a primal-dual mixed integer model presenting a small number of variables and constraints; and the last one is obtained from a flow-based compact model for the Steiner traveling salesman problem (TSP). Numerical experiments indicate that the second compact model allows the efficient solution of randomly generated and benchmark (from the TSPLIB) instances of the problem with hundreds of nodes.

1. Introduction

The elementary $(s - P - t)$-shortest-path problem in a directed graph $G = (V, A)$ with set of nodes V and set of weighted arcs A consists in finding a path of minimum length between an origin node $s \in V$ and a destination node $t \in V$ that visits only once all nodes of a given set $P \subseteq V - \{s, t\}$. We know that for $P = V - \{s, t\}$ the problem is equivalent to finding a Hamiltonian path of minimum length in G, which is NP-hard. In a brief literature review, we find only a few publications on this problem with such a structure, e.g. Dreyfus (1969), Ibaraki (1973), Saksena and Kumar (1966). However, we can find related NP-hard variants of this problem. Indeed, the Steiner Traveling Salesman Problem (STSP) in Letchford, Nasiri, and Theis (2013) asks for non-elementary tours visiting a given set of nodes as feasible solutions (i.e. tours where nodes can be visited more than once in a feasible solution).

Fig. 1 shows an example of a digraph G and its arc costs. Suppose we want to determine a $(1 - \{2\} - 4)$-shortest-path for this digraph considering two situations: (i) a path based on the STSP idea of Letchford et al. (2013); and (ii) an elementary path as in Dreyfus (1969). In this figure, we see the difference between the optimal solution structures for both situations. While the optimal STSP-based $(1 - \{2\} - 4)$-shortest-path is $\{(1, 2), (2, 1), (1, 4)\}$ and has cost equal to 12 (observe here that node 1 is visited twice), the elementary $(1 - \{2\} - 4)$-shortest-path is $\{(1, 2), (2, 3), (3, 4)\}$ and has cost equal to 24. It is not difficult to show that an STSP-based $(s - P - t)$-shortest-path gives a lower bound for the solution of the elementary $(s - P - t)$-shortest-path problem.

It seems that the first (and erroneous, as showed by Dreyfus (1969)) algorithm for the $(s - P - t)$-shortest-path problem is due to Saksena and Kumar (1966). Dreyfus (1969) proposes to solve the problem by reducing it to an instance of the traveling salesman problem. Ibaraki (1973) introduces an exponential dynamic programming algorithm and a branch-and-bound (B&B) method. Ibaraki's model used in the B&B algorithm is defined only with continuous variables. His model is equivalent to the well-known flow formulation of the classic shortest-path problem. Thus, relaxed node solutions in the B&B tree are integer and present at least one cycle when the node solution is not an elementary path (see Ibaraki, 1973). The idea behind Ibaraki's B&B algorithm is to fix at zero (one at a time) an arc of a given cycle C of a B&B node solution as branching rule to create $|C|$ new B&B subproblems. This means possibly enumerating all cycles in G in a B&B tree, because relaxed solutions of the classical flow-based model in Ibaraki (1973) are known to be very weak for this problem.

Here, we adapt the cycle elimination constraints of the compact extended formulation for the spanning tree polytope of an undirected graph (see Conforti, Cornuéjols, and Zambelli, 2010; Martin, 1991; Yannakakis, 1991) to deal with the oriented arcs of the $(s - P - t)$-shortest-path problem. In fact, we show that the use of Martin's formulation is not straightforward for oriented graphs. Moreover, we explore a nice property of elementary paths to obtain a primal-dual based mixed integer compact extended

* Tel.: +55 85 3366 9840; fax: +55 85 3366 9836.
E-mail address: rca@lia.ufc.br, rca_fortal@yahoo.com.br

http://dx.doi.org/10.1016/j.ejor.2016.05.008
0377-2217/

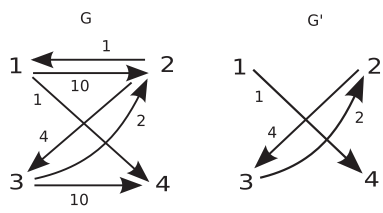

Fig. 1. A digraph $G = (\{1, 2, 3, 4\}, \{(1, 2), (1, 4), (2, 1), (2, 3), (3, 2), (3, 4)\})$ and a sub-digraph G' of G. The arc lengths are reported near each arc.

formulation. The idea is to characterize feasible solutions by linking primal and dual variables in a unique set of constraints exploring that property. We do this without implementing the well-known complementary slackness optimality condition. Such a model has a dual flow-based structure more interesting than the one we derive from the STSP model of Letchford et al. (2013). The new primal-dual model seems to present the smallest number of variables and constraints we can propose to solve this problem. We show the efficiency of our proposed models by performing numerical experiments on randomly generated and on asymmetric TSPLIB instances adapted accordingly. To the best of our knowledge, this is the first work exploring these techniques for solving the elementary $(s - P - t)$-shortest-path problem.

2. Problem formulation

Consider $G = (V, A)$ a directed graph with a set of nodes V and a set of weighted arcs A. Let $c_{ij} \in \mathbb{R}_+$ represent the length of arc $ij \in A$. The problem is to determine an elementary path in G of minimum length between nodes s and t of V that visits a given set of nodes $P \subseteq V - \{s, t\}$.

We represent an $(s - P - t)$-path in G by a vector $x \in \{0, 1\}^{|A|}$, where $x_{ij} = 1$ if ij belongs to the $(s - P - t)$-path, and $x_{ij} = 0$, otherwise. Thus, a mathematical model for this problem is

$$(Q) \quad \min_{x \in \{0,1\}^{|A|}} \sum_{ij \in A} c_{ij} x_{ij} \tag{1}$$

$$s.t. \sum_{i \,|\, iv \in A} x_{iv} - \sum_{j \,|\, vj \in A} x_{vj} = \begin{cases} -1, & \text{if } v = s \\ 1, & \text{if } v = t \\ 0, & \text{otherwise} \end{cases} \quad \forall v \in V \tag{2}$$

$$\sum_{i \,|\, iv \in A} x_{iv} = 1, \ \forall \, v \in P \tag{3}$$

$$\sum_{ij \in A(S)} x_{ij} \leq |S| - 1, \ \forall \, S \subset V \ : |S| \geq 2 \tag{4}$$

where $A(S)$ represents the set of arcs with both extremities in S. Constraints (2) define the existence of an $(s - t)$-path in G and constraints (3) impose that each node $v \in P$ must be visited by imposing that one arc enters v. Constraints (4) avoid the existence of cycles in any solution.

Now let $(Q)_{LR}$ represent hereafter the linear relaxation of model (Q), defined by (1)–(4), where we replace $x \in \{0, 1\}^{|A|}$ in (1) by $\mathbf{0} \leq x \leq \mathbf{1}$.

Observation 1. If arc costs are positive, then dropping constraints (4) in model (Q) can lead to the occurrence of cycles containing at least one visiting node. Indeed, in Fig. 1, the sub-digraph G' is the solution to the model defined by (1), (2), and (3) with respect to the digraph G.

Proposition 1. *The following constraints are valid for* (Q)

$$\sum_{i \,|\, is \in A} x_{is} = 0, \quad \sum_{j \,|\, tj \in A} x_{tj} = 0; \tag{5}$$

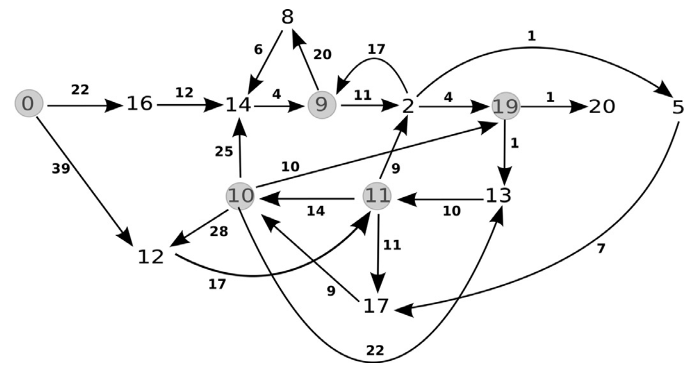

Fig. 2. A digraph with its arc lengths. An elementary $(0 - \{9, 10, 11\} - 19)$-shortest-path is $\{(0, 12), (12, 11), (11, 10), (10, 14), (14, 9), (9, 2), (2, 19)\}$ of cost 114.

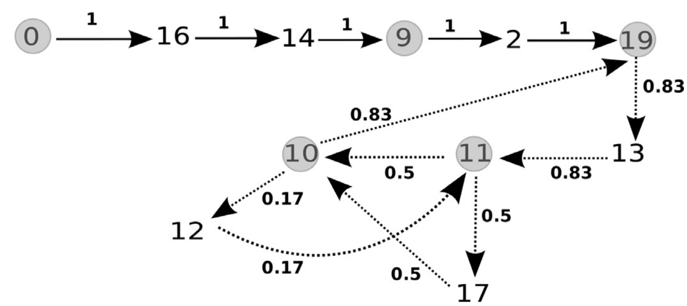

Fig. 3. The continuous optimal solution of value 95 for $(Q)_{LR}$ with respect to a $(0 - \{9, 10, 11\} - 19)$-shortest-path for the digraph in Fig. 2. Values near each arc correspond to the decision variables x associated with the arcs in the solution.

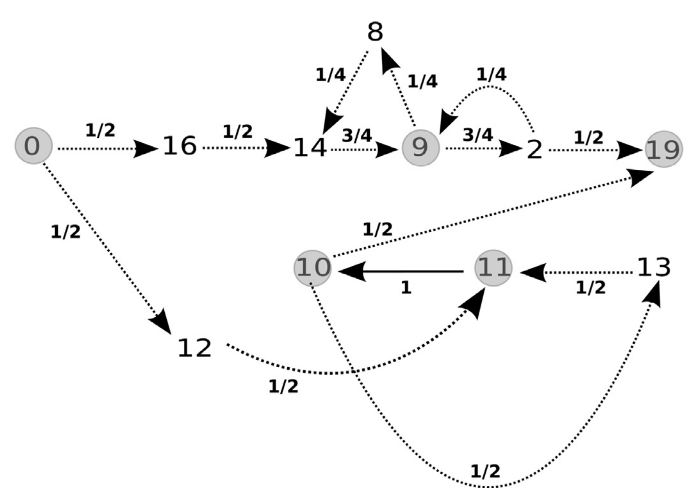

Fig. 4. New optimal solution of value 104 with respect to a $(0 - \{9, 10, 11\} - 19)$-shortest-path of the digraph in Fig. 2 for $((Q) \bigcap (5))_{LR}$.

Proof. In any elementary $(s - P - t)$-path, no arc can enter s or leave t. □

As an example of Proposition 1, consider the digraph in Fig. 2. The optimal solution of $(Q)_{LR}$ for a $(0 - \{9, 10, 11\} - 19)$-shortest-path in this digraph is given in Fig. 3 and has a value equal to 95. When we add constraints (5) to $(Q)_{LR}$, we obtain a new relaxed solution reported in Fig. 4 of value 104.

Proposition 2. *The following constraints strengthen* $((Q) \bigcap (5))_{LR}$.

$$\sum_{i \,|\, iv \in A} x_{iv} \leq 1, \quad \forall \, v \in V - P - \{s, t\} \tag{6}$$