Information Processing Letters 167 (2021) 106067

Contents lists available at ScienceDirect

Information Processing Letters

www.elsevier.com/locate/ipl

An optimal rounding for half-integral weighted minimum strongly connected spanning subgraph

D. Ellis Hershkowitz *, Gregory Kehne, R. Ravi

Carnegie Mellon University, United States of America

ARTICLE INFO

Article history:
Received 15 April 2020
Received in revised form 23 October 2020
Accepted 9 November 2020
Available online 12 November 2020
Communicated by Ranko Lazic

Keywords:
Approximation algorithms
Minimum strongly connected spanning subgraph
Integrality gap
Network design

ABSTRACT

In the weighted minimum strongly connected spanning subgraph (WMSCSS) problem we must purchase a minimum-cost strongly connected spanning subgraph of a digraph. We show that half-integral linear program (LP) solutions for WMSCSS can be efficiently rounded to integral solutions at a multiplicative 1.5 cost. This rounding matches a known 1.5 integrality gap lower bound for a half-integral instance. More generally, we show that LP solutions whose non-zero entries are at least a value $f > 0$ can be rounded at a multiplicative cost of $2 - f$.

1. Introduction

The weighted minimum strongly connected spanning subgraph (WMSCSS) problem is arguably the simplest NP-hard connectivity problem on directed graphs. In WMSCSS we are given a strongly connected[1] digraph $D = (V, A)$ with weight function $w : A \to \mathbb{R}^+$. Our goal is to purchase a strongly connected spanning subgraph $H = (V, A')$ of G of minimum cost, where the cost of H is $w(A') := \sum_{a \in A'} w(a)$. The simplicity of WMSCSS has lent itself to several applications in network design and computational biology [1,19,23,26].

Unfortunately WMSCSS is NP-hard [16]. Even worse, it has been shown to be MaxSNP hard, meaning that it admits no polynomial-time approximation scheme assuming $\mathrm{P} \neq \mathrm{NP}$ [19].

Fortunately WMSCSS admits a simple 2-approximation due to Frederickson and JáJá [11] which employs min-cost arborescences. Given digraph $D = (V, A)$ and root $r \in V$, an r-in-arborescence of D is a spanning subgraph $I = (V, A')$ such that every node $v \neq r$ has exactly one path to r along edges in A'. An r-out-arborescence O is defined analogously with paths from r to v. Minimum-weight r-in- and r-out-arborescences can be computed in polynomial time [6,7]. The mentioned 2-approximation simply fixes an arbitrary root r and then takes the union of a min-cost r-in-arborescence and a min-cost r-out-arborescence. Since every node has a path to and from r in their union, the result is a feasible WMSCSS solution. Moreover, the result is a 2-approximation since the optimal WMSCSS is a strongly connected spanning subgraph and so contains a feasible r-in- and r-out-arborescence as a subgraph for any choice of r. Remarkably, this 2-approximation has remained the best-known polynomial-time approximation for WMSCSS for almost 40 years.

* Corresponding author.
E-mail addresses: dhershko@cs.cmu.edu (D.E. Hershkowitz), gkehne@andrew.cmu.edu (G. Kehne), ravi@andrew.cmu.edu (R. Ravi).

[1] A strongly connected digraph is one in which every node has a directed path to every other node.

https://doi.org/10.1016/j.ipl.2020.106067

Thus, WMSCSS falls into a class of combinatorial optimization problems which admit simple, polynomial-time algorithms whose constant approximation ratios have not been improved in many decades. Notable other examples include the Traveling Salesman Problem (TSP) for which Christofides' simple 1.5-approximation [5] has remained the best polynomial-time approximation since 1976 and the Weighted Tree Augmentation Problem (WTAP) for which the best known polynomial-time approximation ratio is 2 as established by Frederickson and JáJá [11] in 1981 in the same work that gave a 2 approximation for WMSCSS. Given the apparent difficulty in improving these bounds, a great deal of work has focused on improving the approximation ratios of algorithms for special cases of these problems [2,9,13,14,19,20,22,25,27].

A recently fruitful such special case has been the assumption that solutions to the relevant linear program (LP) are half-integral—that is, each coordinate of an optimal solution is assumed to lie in $\{0, \frac{1}{2}, 1\}$. Notably, Cheriyan et al. [4] showed that WTAP admits a 4/3-approximation if the relevant LP is half-integral and Iglesias and Ravi [15] generalized this by showing that a $2/(1+f)$-approximation is possible for WTAP if non-zero LP values are assumed to be at least $f > 0$. Similarly, a recent breakthrough of Karlin et al. [17] showed that a ≈ 1.49993 approximation is possible for TSP if the LP solution is assumed to be half-integral. Studying such special cases offers the opportunity to develop tools and to help delineate lower and upper bounds for the general case. For example, TSP instances with half-integral optimal LP solutions are conjectured to be the hardest TSP instances to approximate [24] and so the work of Karlin et al. [17] was taken as evidence that Christofides' algorithm does not, in fact, attain the best constant approximation among all polynomial-time algorithms—a suspicion recently confirmed by Karlin et al. [18].

1.1. Our contributions

In this work we take this approach to WMSCSS. Specifically, we adapt a deterministic algorithm of Laekhanukit et al. [21] to show that half-integral solutions for the WMSCSS LP can be deterministically rounded in polynomial time at a multiplicative cost of 1.5. In this LP, we enforce that every non-trivial cut has at least one purchased edge leaving. We use $\delta^+(S)$ to denote the set of arcs leaving $S \subset V$.

$$\begin{aligned} \min\ & w(x) \stackrel{\text{def}}{=} w^T \cdot x & & \text{(WMSCSS LP)} \\ \text{s.t.}\ & x(\delta^+(S)) \geq 1 & \forall \emptyset \subset S \subset V & \\ & x_a \geq 0 & \forall a \in A & \end{aligned}$$

Laekhanukit et al. [21] gave a family of half-integral instances of WMSCSS for which the integrality gap of the WMSCSS LP is bounded below by $1.5 - \epsilon$ for any $\epsilon > 0$, and so our upper bound of 1.5 is tight.

More generally, we show how to round any LP solution with non-zero entries bounded below by $f > 0$ at a multiplicative cost of $2 - f$.

Our result for half-integral solutions may be seen as adding to a growing body of evidence that a polynomial-time 1.5-approximation is both achievable and the best possible for WMSCSS. Prior evidence includes a series of works that culminated in a 1.5-approximation for the unit-cost case of WMSCSS [2,19,25] and the aforementioned 1.5 integrality gap lower bound, which is the best known integrality gap lower bound for WMSCSS LP. Since the best known integrality gap lower bound is attained by a half-integral solution, it seems possible that for WMSCSS, like for TSP, the hardest-to-approximate instances may be half-integral.

Our rounding algorithm will be a degenerate form of one proposed by Laekhanukit et al. [21]. In particular, Laekhanukit et al. [21] proposed an algorithm to study the k-arc connected subgraph problem for $k \geq 2$ on *unit-cost* graphs. This is the k-arc connected generalization of the unit-cost WMSCSS. We make use of the same algorithm but use it with $k = 1$ *on arbitrary cost graphs*; thus, the setting in which we apply this algorithm is somewhat different from that of prior work.

2. Rounding

We review the algorithm of Laekhanukit et al. [21] for $k = 1$ before presenting our new analysis.

2.1. Algorithm of Laekhanukit et al. [21]

The algorithm makes use of the r-in- and r-out-arborescence LPs, defined as follows.

$$\begin{aligned} & \min\ w(x) & & \text{(In-Arborescence LP)} \\ \text{s.t.}\ & x(\delta^+(S)) \geq 1 & \forall \emptyset \subset S \subset V \text{ s.t. } r \notin S & \\ & x_a \geq 0 & \forall a \in A & \end{aligned}$$

The r-out-arborescence LP is defined symmetrically; in particular "$\delta^+(S)$" is replaced with "$\delta^-(S)$" (the set of edges entering the set S).

Given an x which is feasible for the arborescence LP, it is known how to efficiently sample arborescences in a manner consistent the marginals defined by x, as summarized by the following claim. In the following, $\mathcal{D}_{\text{in}}$ is a distribution over subgraphs of D each of which contains an in-arborescence and $\mathcal{D}_{\text{out}}$ is a distribution over subgraphs of D each of which contains an out-arborescence.[2]

Lemma 1 *([3,8,10,12]). Given an x feasible for the r-in- and r-out-arborescence LP's, we can, in polynomial time, independently sample sub-graphs $I \sim \mathcal{D}_{\text{in}}$ and $O \sim \mathcal{D}_{\text{out}}$ such that $\Pr(a \in I) = \Pr(a \in O) = x_a$ for every $a \in A$. Moreover, the size of the support of $\mathcal{D}_{\text{in}}$ and $\mathcal{D}_{\text{out}}$ is polynomial in the size of the graph and is computable in polynomial-time.*

[2] In the algorithm originally presented by $\mathcal{D}_{\text{in}}$ Laekhanukit et al. [21] the lemma has I *equal* to an in-arborescence and O *equal* to an out-arborescence and $\Pr(a \in I) \leq x_a$ and $\Pr(a \in O) \leq x_a$. It is easy to see that the claimed lemma immediately follows by adding extra arcs to I and O to increase the probability arc a is sampled to be exactly x_a whenever that is not the case.

The algorithm of Laekhanukit et al. [21] which we adapt for our case is as follows: given an x feasible for WMSCSS LP for an arbitrary root r, enumerate the support of $\mathcal{D}_{\text{in}}$ and $\mathcal{D}_{\text{out}}$ as in Lemma 1 to get digraphs $I_1, \dots I_\alpha$ and $O_1, \dots O_\beta$. Let χ_{ij} be $I_i \cup O_j$. Return the χ_{ij} of minimum weight.[3]

2.2. Rounding LP solutions

We now apply this algorithm to LP solutions whose non-zero entries are at least f. Let P_{MSCS} be the polytope corresponding to WMSCSS LP and let $P_f := \{x \in P_{\text{MSCS}} : x_a \notin (0, f) \text{ for all } a \in A\}$ be all feasible points whose non-zero entries are at least f.

Theorem 1. *Given an $x \in P_f$, the above deterministic algorithm outputs an integral* WMSCSS *solution $\hat{x}$ such that $w(\hat{x}) \le (2 - f) \cdot w(x)$ in polynomial time.*

Proof. Let χ be the digraph returned by the algorithm. Since each O_i and I_j contains an r-out and r-in-arborescence, the characteristic vector corresponding to χ is a feasible integral solution to WMSCSS LP. Moreover, a polynomial runtime follows immediately from Lemma 1 and the fact that there are only polynomially many χ_{ij}'s.

Thus, let us bound the cost of χ. By an averaging argument, it suffices to upper bound the expected cost of $\hat{\chi} = I \cup O$ where $I \sim \mathcal{D}_{\text{in}}$ and $O \sim \mathcal{D}_{\text{out}}$ since χ is in the support of this distribution, and the minimum cost member in the collection has cost at most the average.

Consider a fixed arc a. By the inclusion-exclusion principle, the fact that $a \in I$ is independent of whether $a \in O$, and Lemma 1, the probability that a is in $\hat{\chi}$ is

$$
\begin{aligned}
\Pr(a \in \hat{\chi}) &= \Pr(a \in I) + \Pr(a \in O) - \Pr(a \in I)\Pr(a \in O) \\
&= 2x_a - x_a^2.
\end{aligned}
$$

Thus, by linearity of expectation we have

$$
\begin{aligned}
\mathbb{E}[w(\hat{\chi})] &= \sum_a w(a) \cdot (2x_a - x_a^2) \\
&\le \sum_a w(a) \cdot (2x_a - x_a \cdot f) \\
&\qquad (\text{Since } x_a > 0 \implies x_a \ge f \quad \forall x_a \in P_f) \\
&= (2 - f) \cdot w(x). \quad \square
\end{aligned}
$$

As a corollary of Theorem 1 we recover our 1.5-cost rounding for half-integral solutions:

Corollary 1. *There is a deterministic polynomial-time algorithm which, given an $x \in P_{1/2}$, outputs a feasible integral $\hat{x}$ such that $w(\hat{x}) \le 1.5 \cdot w(x)$.*

[3] Here Laekhanukit et al. [21] actually used the optimal x. Additionally, Laekhanukit et al. [21] stated their algorithm as choosing uniformly at random from among the χ_{ij} but then remarked that it can be derandomized; we give the derandomized version.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

D Ellis Hershkowitz supported in part by NSF grants CCF-1527110, CCF-1618280, CCF-1814603, CCF-1910588, NSF CAREER award CCF-1750808 and a Sloan Research Fellowship. R Ravi supported in part by the U.S. Office of Naval Research award N00014-18-1-2099.

References

[1] Satabdi Aditya, Bhaskar DasGupta, Marek Karpinski, Algorithmic perspectives of network transitive reduction problems and their applications to synthesis and analysis of biological networks, Biology 3 (1) (2013) 1–21.
[2] Piotr Berman, Bhaskar DasGupta, Marek Karpinski, Approximating transitive reductions for directed networks, in: Workshop on Algorithms and Data Structures, Springer, 2009, pp. 74–85.
[3] Robert Carr, Santosh Vempala, Randomized metarounding, Random Struct. Algorithms 20 (3) (2002) 343–352.
[4] Joseph Cheriyan, Tibor Jordán, R. Ravi, On 2-coverings and 2-packings of laminar families, in: Annual European Symposium on Algorithms (ESA), Springer, 1999, pp. 510–520.
[5] Nicos Christofides, Worst-case analysis of a new heuristic for the travelling salesman problem, Technical report, Carnegie Mellon University, 1976.
[6] Yoeng-Jin Chu, On the shortest arborescence of a directed graph, Sci. Sin. 14 (1965) 1396–1400.
[7] Jack Edmonds, Optimum branchings, J. Res. Natl. Bur. Stand. B 71 (4) (1967) 233–240.
[8] Jack Edmonds, Edge-disjoint branchings, Comb.l Algorithms (1973) 91–96.
[9] Samuel Fiorini, M. Groß, J. Könemann, L. Sanità, A 3/2-approximation algorithm for tree augmentation via chvátal-gomory cuts, in: 15th Cologne-Twente Workshop on Graphs and Combinatorial Optimization, 2017, p. 77.
[10] Andras Frank, Covering branchings, Acta Sci. Math. 41 (1979) 77–81.
[11] Greg N. Frederickson, Joseph JáJá, Approximation algorithms for several graph augmentation problems, SIAM J. Comput. 10 (2) (1981) 270–283.
[12] Harold N. Gabow, A matroid approach to finding edge connectivity and packing arborescences, J. Comput. Syst. Sci. 50 (2) (1995) 259–273.
[13] Shayan Oveis Gharan, Amin Saberi, Mohit Singh, A randomized rounding approach to the traveling salesman problem, in: Symposium on Foundations of Computer Science (FOCS), IEEE, 2011, pp. 550–559.
[14] Benjamin Grimmer, Dual-based approximation algorithms for cut-based network connectivity problems, Algorithmica 80 (10) (2018) 2849–2873.
[15] Jennifer Iglesias, R. Ravi, Coloring down: 3/2-approximation for special cases of the weighted tree augmentation problem, preprint, arXiv:1707.05240, 2017.
[16] David S. Johnson, Michael R. Garey, Computers and Intractability: A Guide to the Theory of NP-Completeness, vol. 1, WH Freeman, San Francisco, 1979.
[17] Anna Karlin, Nathan Klein, Shayan Oveis Gharan, An improved approximation algorithm for tsp in the half integral case, in: Annual ACM Symposium on Theory of Computing (STOC), 2020.
[18] R. Anna Karlin, Nathan Klein, Shayan Oveis Gharan, A (slightly) improved approximation algorithm for metric tsp, preprint, arXiv:2007.01409, 2020.
[19] Samir Khuller, Balaji Raghavachari, Neal Young, Approximating the minimum equivalent digraph, SIAM J. Comput. 24 (4) (1995) 859–872.

[20] Guy Kortsarz, Zeev Nutov, A simplified 1.5-approximation algorithm for augmenting edge-connectivity of a graph from 1 to 2, ACM Trans. Algorithms 12 (2) (2015) 1–20.

[21] Bundit Laekhanukit, Shayan Oveis Gharan, Mohit Singh, A rounding by sampling approach to the minimum size k-arc connected subgraph problem, in: International Colloquium on Automata, Languages, and Programming, Springer, 2012, pp. 606–616.

[22] Tobias Mömke, Ola Svensson, Approximating graphic tsp by matchings, in: Symposium on Foundations of Computer Science (FOCS), IEEE, 2011, pp. 560–569.

[23] Dennis M. Moyles, Gerald L. Thompson, An algorithm for finding a minimum equivalent graph of a digraph, Technical report, 1967.

[24] Frans Schalekamp, David P. Williamson, Anke van Zuylen, 2-matchings, the traveling salesman problem, and the subtour lp: a proof of the Boyd-Carr conjecture, Math. Oper. Res. 39 (2) (2014) 403–417.

[25] Adrian Vetta, Approximating the minimum strongly connected subgraph via a matching lower bound, in: Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), 2001, pp. 417–426.

[26] Vincent Dubois, Bothorel Cecile, Transitive reduction for social network analysis and visualization, in: Proceedings of the 2005 IEEE/WIC/ACM International Conference on Web Intelligence, IEEE Computer Society, 2005, pp. 128–131.

[27] Liang Zhao, Hiroshi Nagamochi, Toshihide Ibaraki, A linear time 5/3-approximation for the minimum strongly-connected spanning subgraph problem, Inf. Process. Lett. 86 (2) (2003) 63–70.