Electronic Colloquium on Computational Complexity, Report No. 112 (2009)

Hardness of an Asymmetric 2-player Stackelberg Network Pricing Game

Davide Bilò[1], Luciano Gualà[2], and Guido Proietti[1,3]

[1] Dipartimento di Informatica, Università dell'Aquila, 67010 L'Aquila, Italy
[2] Dipartimento di Matematica, Università di Tor Vergata, 00133 Roma, Italy
[3] Istituto di Analisi dei Sistemi ed Informatica, CNR, 00185 Roma, Italy
E-mail: {davide.bilo,proietti}@di.univaq.it, guala@mat.uniroma2.it

Abstract. Consider a communication network represented by a directed graph $G = (V, E)$ of n nodes and m edges. Assume that edges in E are partitioned into two sets: a set C of edges with a fixed non-negative real cost, and a set P of edges whose *price* should instead be set by a *leader*. This is done with the final intent of *maximizing* the payment she will receive for their use by a *follower*, whose goal is to select for his communication purposes a *minimum-cost* (w.r.t. to a given objective function) subnetwork of G. In this paper, we study the natural setting in which the follower computes a *single-source shortest paths tree* of G, and then returns to the leader a payment equal to the *sum* of the selected priceable edges. Thus, the problem can be modeled as a one-round two-player *Stackelberg Network Pricing Game (SNPG)*, with the additional complication that the objective functions of the two players are *asymmetric*. Indeed, the revenue provided to the leader by any of her selected edges is simply its price, while the cost of such an edge in the minimization function of the follower is given by its price multiplied by the number of paths (emanating from the source) it belongs to. As we will see, this asymmetry makes the problem much harder than other previously studied symmetric SNPGs. More precisely, we show that for any $\epsilon > 0$, unless $\mathsf{P} = \mathsf{NP}$, the problem is not approximable within $n^{1/2-\epsilon}$, while if G is unweighted and the leader can only decide which of her edges enter in the solution, then the problem is not approximable within $n^{1/3-\epsilon}$. On the positive side, when edges in C happen to form the common *unweighted star network* topology, then we show the problem becomes APX-hard, and admits a 92-approximation algorithm. Furthermore, for general instances, we devise a *strongly* polynomial-time $O(n)$-approximation algorithm, which favorably compares against the powerful *single-price* algorithm.

Keywords: Communication Networks, Shortest Paths Tree, Stackelberg Games, Network Pricing Games.

1 Introduction

Leader-follower games were introduced by von Stackelberg in the far 1934 [15], with the aim of modeling heterogeneous markets, namely markets in which one or more players are in a leadership position, and can in practice manipulate the market to their own advantage, by directly influencing the choices of the remaining subjects. In the basic formulation, the game is played by only two players: the leader which moves first, and the follower which observes the leader's move and then makes his[4] own move, after which the game is over. The

[4] Throughout the paper, we adopt the convention of referring to the leader and to the follower with female and male pronouns, respectively.

ISSN 1433-8092

strategic aspect of the game consists of the fact that the follower computes a solution by optimizing an objective (public) function, while the leader has her own objective function which is, by the way, computed over the solution selected by the follower. In this way, to optimize her revenue, the leader has to entail in her move the optimal response which will be given by the follower.

Recently, leader-follower games received a considerable attention from the computer science community, because of the perfect paradigm of heterogenous market expressed by the Internet. In fact, the Internet is a vast, pervasive electronic market mainly composed of millions of independent end-users, whose actions are by the way influenced by the owners of physical/logical portions of the network, like for instance the service providers. Under this perspective, it turns out to be particularly intriguing the problem of analyzing the antagonism emerging between leaders and followers whenever a communication subnetwork must be allocated.

1.1 Stackelberg network pricing games

Network games can be easily regarded as Stackelberg games, as soon as it arises a situation in which a subset of dominant players control a higher-level decision phase in which part of the game instance is set, for example by routing a substantial amount of a network flow [6], or by deciding the cost of a subset of network arcs. In particular, games of this latter type, which are of interest for this paper, are widely known as *Stackelberg Network Pricing Games* (SNPGs).

A SNPG can be formalized as follows: We are given an either directed or undirected graph $G = (V, E = C \cup P, C \cap P = \emptyset)$, with an edge cost function $c : e \in C \mapsto \mathbb{R}_{\geq 0}$, while edges in P need to be priced by the leader. In the following, we assume that $n = |V|$ and $m = |C| + |P|$. Then, the leader moves first and chooses a pricing function $p : e \in P \mapsto \mathbb{R}_{\geq 0}$ for her edges, in an attempt to *maximize* her objective function $f_1(p, H(p))$, where $H(p)$ denotes the decision which will be taken by the follower, consisting in the choice of a subgraph of G. This notation stresses the fact that the leader's problem is implicit in the follower's decision. Once observed the leader's choice, the follower reacts by selecting a subgraph $H = (V', E')$ of G which *minimizes* his objective function $f_2(p, H)$, parameterized in p. Note that the leader's strategy affects both the follower's objective function and the set of feasible decisions, while the follower's choice only affects the leader's objective function. Throughout the paper, we will naturally assume that f_1 is *price-additive*, i.e., $f_1(p, H(p)) = \sum_{e \in P \cap E'} p(e)$. This means, the leader decides edge prices having in mind that her revenue equals the overall price of her selected edges.

The most immediate SNPG is that in which we are given two specified nodes in G, say s, t, and the follower wants to travel along a shortest path in G between s and t (see [14] for a survey). This problem has been shown to be APX-hard [10], as well as $O(\log |P|)$-approximable [13]. For the case of multiple followers (each with a specific source-destination pair), Labbé *et al.* [11] derived a bilevel LP formulation of the problem (and proved NP-hardness), while Grigoriev *et al.* [9] presented algorithms for a restricted shortest path problem on parallel edges. Another basic SNPG game is that in which the follower wants to use a *minimum spanning tree* of G (now considered as undirected). For this game, in [5] the authors proved the APX-hardness already when the number of possible edge costs is 2, and gave an $O(\log n)$-approximation algorithm.

All the above examples fall within the class of SNPGs which can be handled by the general model proposed in [4], which encompasses all the cases where each follower optimizes a polynomial-time network optimization problem in which the cost of the network is given by the sum of prices and costs of contained edges, namely

$$f_2(p, H) = \sum_{e \in P \cap E'} p(e) + \sum_{e \in C \cap E'} c(e). \tag{1}$$

Thus, in this model we have that f_1 coincides with f_2 as soon as this latter is restricted to the leader's edges. To this respect, the leader's and follower's maximization and minimization functions are therefore *symmetric*. The authors show that all SNPGs in this class can be tightly approximated within $(1+\epsilon)(\mathcal{H}_k + \mathcal{H}_{|P|})$, for any $\epsilon > 0$, where k denotes the number of followers, while $\mathcal{H}_i$ is the i-th harmonic number. But what about the case in which the symmetry does not hold?

1.2 Asymmetric SNPGs

In this paper, we focus on an emblematic *asymmetric* SNPG, namely that in which the follower aims at building a *single-source shortest paths tree (SPT)* $S = (V, E')$ of G rooted at a given node r. This game is clearly asymmetric, since if we denote by $\pi(r, v)$ the path in S between r and $v \in V$, we have that

$$f_2(p, S) = \sum_{e \in P \cap E'} \sum_{v \in V | e \in \pi(r,v)} p(e) + \sum_{e \in C \cap E'} \sum_{v \in V | e \in \pi(r,v)} c(e).$$

Such a game, that we name *Asymmetric Stackelberg Shortest Paths Tree (*ASSPT*)* game, finds its motivation in the popularity of the *multicast* protocol in the Internet: here, the follower needs to perform a broadcasting from r, but this is implemented *without* a replication for each of the destination nodes of the broadcasted message. Thus, just a single message travels on each SPT edge during a broadcasting, and the revenue for a leader's edge should correspondingly be simply equal to its price.

With the study of the ASSPT game we continue in the effort of analyzing the computational aspect of SPT pricing games, which started in [3] with the symmetric version of the game (i.e., that in which the leader's revenue for each selected edge is given by its price multiplied by the number of paths – emanating from the source – it belongs to). Recall that for this game, we proved that finding an optimal pricing for the leader's edges is NP-hard, as soon as $|P| = \Theta(n)$. In the following, as usual we assume that when multiple optimal solutions are available for the follower, then he selects an optimal solution maximizing the leader's revenue.[5]

1.3 Our results

In this paper, we analyze the ASSPT game under several respects. More precisely, we first study the complexity of the game, and we show that finding an optimal pricing for the

[5] At a first glance, this rule looks in contrast with the antagonistic nature of the game. However, if this rule is relaxed, then it easy to see that an optimal solution for the leader can only be reached within any arbitrary small subtractive term.

leader's edges is an extremely difficult task, even under strongly restrictive assumptions. Indeed, we show that for every $\epsilon > 0$, the ASSPT is not approximable within a factor of $n^{1/2-\epsilon}$, unless P=NP. Then, we turn our attention to the *unweighted* ASSPT game, i.e., that in which $c(e) = 1$ for any $e \in C$, while $p : e \in P \mapsto \{1, +\infty\}$, and we prove that for every $\epsilon > 0$, the game is not approximable within a factor of $n^{1/3-\epsilon}$, unless P = NP. As a further hardness result, we show the APX-hardness of the ASSPT game even for the restrictive but still significant case in which edges in C have unitary cost and forms a star rooted at r. Noticeably, this is obtained by proving its equivalence with the well-studied *Rooted Maximum Leaf Outbranching (RMLO)* problem, which asks for finding a spanning arborescence rooted at some prescribed vertex of a digraph with the maximum number of leaves. This problem is known to be NP-hard already on *acyclic digraphs (DAGs)* [2], and it admits a 92-approximation algorithm [7]. Hence, on one hand we obtain an approximation algorithm for our problem, with the same performance ratio, and on the other hand we improve the inapproximability result for the RMLO problem, by showing that it is APX-hard even on DAGs.

Finally, we turn our attention to the development of an approximation algorithm for the ASSPT game, and we devise a pricing strategy that in $O(m + n \log n)$ time returns an $(n-1)$-approximation of the optimal leader's revenue. Although our algorithm leaves an $O(\sqrt{n})$ gap open with the corresponding inapproximability result, remarkably its running time is *strongly polynomial* (i.e., it does not depend on the numerical values of the costs of the edges in C). Therefore, it compares favorably with the powerful but non-strongly polynomial *single-price algorithm* [4], that has a provably logarithmic approximation ratio for *any* symmetric SNPGs in which f_2 is of the form (1), while in constrast we show that for the ASSPT game it can only guarantee an $\Omega(n)$ factor.

The rest of the paper is organized as follows: Sections 2 contains inapproximability results, while Section 3 deals with the unweighted star case. Finally, Section 4 presents our approximation algorithm, along with a bad instance example for the single-price algorithm.

2 Non-approximability results

In this section we prove that the ASSPT game along with some basic variants of it are very hard to approximate. For the sake of simplifying the presentation, we first analyze a *binary pricing* version of the game, denoted as ASSPT($\{\rho_1, \rho_2\}$), in which the leader is constrained to price her edges either ρ_1 or ρ_2. Then, we start by proving the following:

Theorem 1. *For every $\rho_2 > \rho_1 \geq 0$ and for every $\epsilon > 0$, the* ASSPT($\{\rho_1, \rho_2\}$) *game is not approximable within a factor of $n^{1/2-\epsilon}$, unless* P = NP*, even on DAGs.*

Proof. The reduction is from the *Maximum Independent Set (MIS)* problem, i.e., that of finding a maximum cardinality set I of vertices of a given undirected graph G such that no pair of vertices in I is linked by an edge of G. For every $\epsilon > 0$, the MIS problem is not approximable within a factor of $n^{1-\epsilon}$, unless P = NP, where n is the number of vertices of G [16]. The non-approximability result still holds for the class of connected graphs because finding an independent set of a graph is equivalent to finding an independent set for each of its connected components.

The reduction works as follows. From a given undirected connected graph $G = (V, E)$ of n vertices, we build a 3-layered DAG G'. The first layer of G' contains the root vertex

r (w.l.o.g., we can assume that $r \notin V$), the second layer of G' contains a copy of all the vertices of G, while the third layer of G' contains n copies of all the vertices of G. The set of fixed-cost edges is the following: there is an edge of cost ρ_2 from r to every vertex in the second layer, and there is an edge of cost $\rho = \max\{0, 2\rho_1 - \rho_2\}$ from a vertex u in the second layer to all the copies of vertex v in the third layer iff $(u, v) \in E$. The set P of leader's edges is the following: there is an edge from r to every vertex in the second layer, and every vertex u in the second layer has an outgoing edge towards each of its copies in the third layer. An example of the reduction is shown in Figure 1.

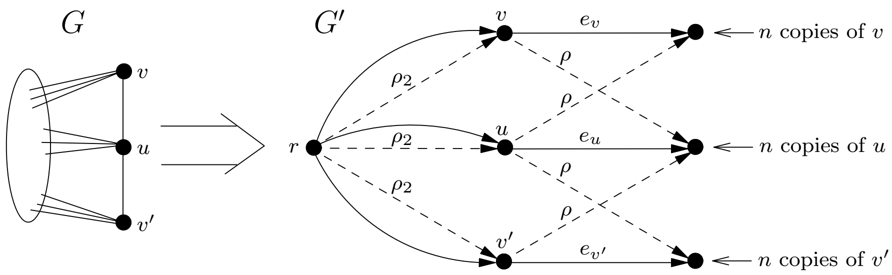

Fig. 1. An example of the reduction defined in Theorem 1 for the graph G on the left side. The DAG G' restricted to the vertices u, v, v' of G is shown on the right side. Dashed edges are fixed-cost edges while the other edges of G' are owned by the leader. Observe that the distance from r to v' does not depend on the price set on the edge (r, v) but depends on the price set on the edge (r, u). This is because $\{v, v'\}$ is an independent set of G while $\{u, v\}$ is not.

By the connectivity property of G, there is a fixed-cost path of length ρ_2 from r to every vertex in the second layer, and a fixed-cost path of length $\rho_2 + \rho$ from r to every vertex in the third layer. Let ρ' be equal to ρ_1 if $\rho_1 > 0$, ρ_2 otherwise. In what follows, we prove that every pricing $p : P \to \{\rho_1, \rho_2\}$ defines an independent set I_p of G, and yields a revenue of at most $\rho_2 n + \rho' n|I_p|$. Moreover, we also prove that for every independent set I of G, there exists a pricing p yielding a revenue of at least $\rho' n|I|$, and such that $I_p = I$. The claim then follows from the non-approximability result of the MIS problem and by observing that G' has $\Theta(n^2)$ vertices. Indeed, for the hardest instances of MIS problem, the maximum independent set I^* has size at least $n^{1-\epsilon'}$, for every $\epsilon' > 0$. This implies that $\frac{\rho' n|I^*|}{\rho_2 n} > n^{1-\epsilon''}$, for some ϵ''.

Let e_u denote any leader's edge from a vertex u in the second layer to one of its copies in the third layer. For a pricing p, let I_p denote the set of vertices u having some e_u with $p(e_u) > 0$ being part of an SPT of G'. The edge e_u is contained in an SPT of G' iff

$$\forall v \in V \text{s.t. } (u, v) \in E, \quad p(r, u) + p(e_u) \leq p(r, v) + \rho.$$

From the above inequalities, we have that e_u is contained in an SPT of G' and $p(e_u) > 0$ iff

$$\begin{cases} p(r, u) = \rho_1, \\ p(e_u) = \rho' \\ p(r, v) = \rho_2 \quad \forall v \in V \text{ s.t. } (u, v) \in E. \end{cases} \tag{2}$$

This implies that if edge e_u is in some SPT of G' and $p(e_u) > 0$, then no SPT of G' contains the edge e_v, for every vertex v such that $(u, v) \in E$. As a consequence, the set I_p always defines an independent set of G. Moreover, observe that the revenue yielded by p is at most $\rho_2 n$ for the leader's edges outgoing from r, and at most $\rho' n|I_p|$ for the other edges. Therefore, the total revenue yielded by p is at most $\rho_2 n + \rho' n|I_p|$.

To complete the proof, let I be an independent set of G. The pricing p that satisfies the equalities in (2) for every e_u and for every $u \in I$, yields a revenue of at least $\rho' n|I|$, and defines an independent set I_p such that $I_p = I$. □

From the above theorem we can easily derive the following general result:

Theorem 2. *For every $\epsilon > 0$, the* ASSPT *game is not approximable within a factor of $n^{1/2-\epsilon}$, unless* $\mathsf{P} = \mathsf{NP}$*, even on DAGs where fixed-cost edges have all cost 1.*

Proof. Consider the DAG G' built in the reduction of Theorem 1 for the case $\rho_1 = \frac{n+1}{2}$ and $\rho_2 = n$, which implies $\rho = 1$. Build a graph G'' from a copy of G' by replacing every fixed-cost edge of cost k with k edges all having cost 1. Figure 2 shows an example of the reduction. Observe that G'' still contains $\Theta(n^2)$ vertices.

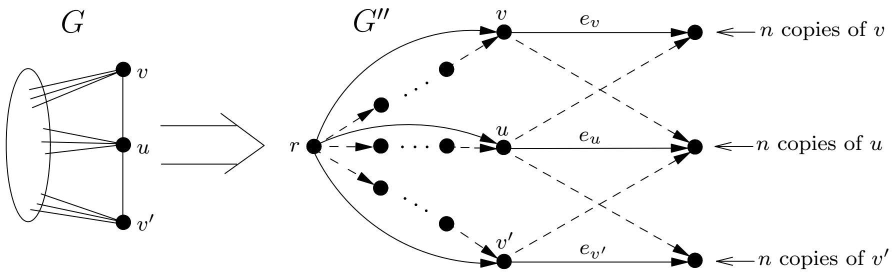

Fig. 2. An example of the reduction defined in Theorem 2 for the graph G on the left side. The fixed-cost paths going from r to u, v, v' contains n edges, respectively.

For a feasible pricing p, let I_p denote the set of vertices u having some e_u with $p(e_u) > 1$ being part of an SPT of G'. The edge e_u is contained in an SPT of G' iff

$$\forall v \in V \text{s.t. } (u, v) \in E, \quad \min\{\rho_2, p(r, u)\} + p(e_u) \leq \min\{\rho_2, p(r, u)\} + 1.$$

This implies that if edge e_u is in some SPT of G' and $p(e_u) > 1$, then no SPT of G' contains the edge e_v, for every vertex v such that $(u, v) \in E$. Therefore, I_p is an independent set of G. Moreover, p yields a revenue of at most $O(n^2) + n^2|I_p|$. In Theorem 1 we have shown a pricing p in G' yielding a revenue R of at least $\rho_2 n|I| = n^2|I|$, where I is a maximum independent set of G. It is easy to see that pricing G'' with p yields a revenue of at least R. This completes the proof. □

We now turn our attention to the *unweighted ASSPT* game, i.e., that in which $c(e) = 1$ for any $e \in C$, while the leader's price function is restricted to $p : e \in P \mapsto \{1, +\infty\}$. Even in this simplified binary version, the game is very hard:

Theorem 3. *For every $\epsilon > 0$, the unweighted* ASSPT *game is not approximable within a factor of $n^{1/3-\epsilon}$, unless $\mathsf{P} = \mathsf{NP}$, even on DAGs.*

Proof. Consider the DAG G'' defined in the proof of Theorem 2. G'' has been built from the DAG G' defined in Theorem 1 for the case $\rho_1 = \frac{n+1}{2}$ and $\rho_2 = n$, which implies $\rho = 1$.

From G'', we built the DAG G''', which is obtained by replacing each edge e_u by a path π_u of n leader's edges. Moreover, for every vertex $v' \in \pi_u$ and for every vertex v such that (u, v) is a unit-cost edge of G'', we add the unit-cost edge (v, v') to G'''. As a consequence, there is a fixed-cost path of length $n + 1$ from r to every v'. Figure 3 shows an example of the reduction. Observe that G''' has $\Theta(n^3)$ vertices.

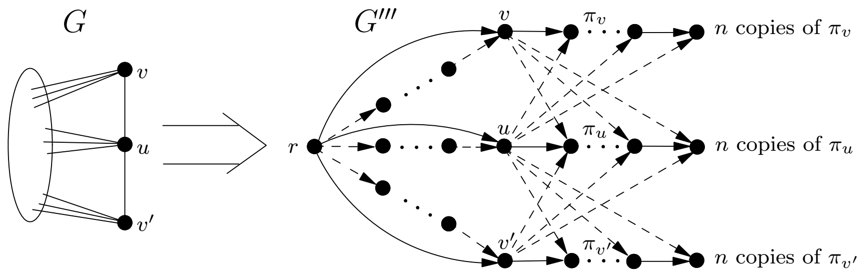

Fig. 3. An example of the reduction defined in Theorem 3 for the graph G on the left side. The reduction is an extension of the reduction of Theorem 2. The paths π_u, π_v, and $\pi_{v'}$ contains n priceable edges each.

Let p be a pricing in G''' that prices edges either 1 or $+\infty$. Let p' be a pricing in G'' defined as follows. For an edge e incident to the root vertex, $p'(e) = 1$ if $p(e) = 1$. An edge e_u has a price of i if the first $i \geq 1$ edges of the corresponding path π_u in G''' are contained in an SPT of G''', and no SPT of G''' contains the first $i+1$ edges of the corresponding path π_u in G'''. All the remaining edges are priced with an arbitrarily large value.[6] Observe that if p yields a revenue of R, than so does p'. Indeed, an edge (r, u) is contained in an SPT of G'' iff it is contained in an SPT of G'''. Furthermore, an edge e_u with price i is contained in an SPT of G'' iff the first i edges of the corresponding path π_u are contained in an SPT of G'''.

In the proof of Theorem 2, we have shown that p' defines an independent set $I_{p'}$ of G, and yields a revenue R' of at most $O(n^2) + n^2|I_{p'}|$. As a consequence, p defines an independent set $I_p = I_{p'}$ of G, and yields a revenue R of at most R'. Let I be a maximum independent set of G. Let p be a pricing in G''' which prices an edge e with 1 iff $e = (r, u)$, or $e \in E(\pi_u)$ and $u \in I$. Observe that p yields a revenue of $|I_p| + n^2|I_p|$. This completes the proof. □

3 Dealing with unweigthed stars

In this section we focus on instances where edges in C have cost 1 and form a star emanating from the source node r. We show that the problem of finding a pricing maximizing the

[6] Observe that pricing an edge with $+\infty$ is equivalent to pricing it with an arbitrarily large value.

revenue is equivalent to another problem known in literature as the *Rooted Maximum Leaf Outbranching (RMLO) problem*, which asks for finding a spanning arborescence rooted at some prescribed vertex of a digraph with the maximum number of leaves. This problem is known to be NP-hard even when restricted to DAGs [2], and MaxSNP-hard, even on undirected graphs [8]. Moreover, it admits a 92-approximation algorithm [7]. Hence, the equivalence result immediately implies the existence of a constant-factor approximation algorithm for our problem, with the same approximation ratio. Moreover, we improve the inapproximability result for the RMLO problem by showing that it is APX-hard even for DAGs.

First of all, we observe that we can restrict ourselves to consider instances of pricing where all nodes can be reached by a path of priceable edges, since any other node would be reached through a single fixed-cost edge in any SPT computed by the follower, and thus it cannot influence the revenue. We can prove the following:

Lemma 1. *Let $G = (V, C \cup P)$ be an instance of the ASSPT game where edges in C have unitary cost and form a star rooted at r. Then, G admits a pricing with revenue greater than or equal to $k \in \mathbb{N}$ if and only if $G' = (V, P)$ has a spanning arborescence rooted at r with at least k leaves.*

Proof. Let T be a spanning arborescence of G' rooted at r with at least k leaves. Then, we can define the following pricing for G: $p(e) = 1$, if e enters into a leaf of T, 0 otherwise. It is easy to see that such a pricing yields a revenue of k, since T is a SPT of G.

On the other hand, let us consider a pricing p with revenue $R \geq k$. Let T be the SPT computed by the follower after the leader prices her edges according to p, and let (r, v) be an edge of T. Then, all the edges of the subtree of T rooted at v, say $T(v)$, must be priceable edges. Moreover, if (r, v) is a fixed-cost edge, then all the edges in $T(v)$ must have price 0. Now, let T' be the subtree of T obtained by removing every fixed-cost edge (r, v) and the corresponding subtree $T(v)$. From the arguments given above, we have that $R = \sum_{e \in E(T')} p(e)$. Now, we show that the number of leaves of T', say ℓ, is at least R. Indeed, we have that

$$R = \sum_{e \in E(T')} p(e) \leq \sum_{v \mid v \text{ is a leaf in } T'} d_{T'}(r, v) \leq \ell,$$

where the last inequality holds since T' is a subtree of a SPT of G, which implies that for each node v, $d_{T'}(r, v) \leq 1$.

Now, notice that T' is an arborescence rooted at r of G' as well, but it may not span V. However, we can add priceable edges to T' in order to make it spanning. This does not decrease the number of leaves. □

Theorem 4. *The ASSPT game where the edges of C have unitary cost and form a star rooted at r is* APX*-hard, unless* P = NP.

Proof. The reduction is from *Set Cover problem (SCP)*. An instance $I = \langle O, \mathcal{S} \rangle$ of SCP consists of a set of $O = \{o_1, \ldots, o_n\}$ objects and a set $\mathcal{S} = \{S_1, \ldots, S_m\}$ of m subset of O. The objective is to find a minimum-size collection of subsets in $\mathcal{S}$ whose union is O. In [12], it is shown that SCP is NP-hard.

Given an instance $I = \langle O, \mathcal{S} \rangle$ of SCP we build the following instance of the ASSPT game. We have the root vertex r, a node u_j for each S_j, and a node v_i for each o_i. There is a fixed-cost edge (of cost 1) from r to each other node. Moreover, we have a priceable edge (r, u_j) for each u_j, and a priceable edge (u_j, v_i) if and only if $o_i \in S_j$ (see Figure 4). We have the following:

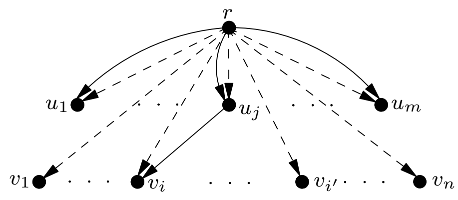

Fig. 4. The reduction of Theorem 4. Dashed edges are fixed-cost edges, while the other edges are owned by the leader. In this example, o_i is in S_j, but $o_{i'}$ is not.

Lemma 2. *The instance of the ASSPT game has a pricing yielding a revenue of at least $n + m - k$ if and only if I admits a cover of size at most k.*

Proof. Let $\mathcal{C}$ be a cover of size k for I. Then, define the following pricing: set $p(r, u_j) = 0$ for each $S_j \in \mathcal{C}$, while all the other priceable edges are set to 1. It is easy to see that we obtain a unit of revenue for each v_i, and a unit of revenue for each u_j with $S_j \notin \mathcal{C}$, which provides a total revenue of $n + m - k$.

Conversely, let p be a pricing yielding a revenue $R \geq n + m - k$, and let S be the SPT computed by the follower according to p. We define two sets of nodes. Let X be the set containing every u_j which has in S at least one outgoing priceable edge, and let Y be the set containing every v_i which has in S an ingoing priceable edge. Then, since every node has a distance from r of at most 1, an upper bound for R is:

$$R = \sum_{e \in E(S) \cap P} p(e) \leq \sum_{v_i \in Y} d_S(r, v_i) + \sum_{u_j \notin X} \max\{1, p(r, u_j)\} \leq |Y| + m - |X|.$$

We now define a pricing p' as follows: $p'(r, u_j) = 0$ for each $u_j \in X$, while all the other edges are priced at 1. Hence, it is easy to see that p' yields a revenue of at least $|Y| + m - |X|$, since now each $v_i \in Y$ is at distance 1 from r (and there is a path of priceable edges with length 1), and since every $u_j \notin X$ has an ingoing leader's edge with price 1.

Now, we modifies p' in such a way that (i) the revenue does not change, and (ii) in the corresponding SPT S' computed by the follower, every v_i has an ingoing priceable edge with price 1. We repeatedly do as follows. Consider any v_i that is not reached in S' by a priceable edge with price 1, and consider a node u_j such that $o_i \in S_j$. We change $p'(r, u_j)$ from 1 to 0. This does not change the revenue, since we loose a unit of revenue from the edge (r, u_j), while we obtain an additional unit of revenue from the edge (u_j, v_i) which is now selected by the follower.

Finally, we define a cover for I by selecting every S_j corresponding to a node u_j having distance 0 in p'. Since the revenue of p' is at least $n + m - k$, and since in S' every v_i is reached by a priceable edge with price 1, it follows that the size of the cover is at most k. This concludes the proof. ☐

In order to prove the APX-hardness, we restrict ourselves to instances of the SCP in which each subset has cardinality at most 3, and $m \geq n$. Even in this case, the SCP is APX-hard [1]. We show that a $(1+\epsilon)$-approximate algorithm for the ASSPT game would imply a $(1+\epsilon')$-approximate algorithm for the SCP, for a suitable ϵ'. Assume that we have a $(1+\epsilon)$-approximate algorithm for the ASSPT, and let k^* be the size of an optimum set cover for I. We have that the algorithm returns a pricing p with revenue $R \geq (m+n-k^*)/(1+\epsilon)$. In the proof of Lemma 2, we have shown that p can be modified in order to yield a revenue of at least $\lceil R \rceil$. Let k be the integer such that $\lceil R \rceil = m+n-k$. Hence, from Lemma 2, we have that p induces a set cover of size k. Then, since $m \leq 3k^*$, we have that:

$$k \leq k^* + \frac{\epsilon}{1+\epsilon}(n+m-k^*) \leq k^* + \frac{5\epsilon}{1+\epsilon}k^* \leq (1+\epsilon')k^*.$$

□

Therefore, from Lemma 1 and Theorem 4, we obtain the following interesting result, which improves over the NP-hardness on DAGs of the RMLO problem given in [2]:

Corollary 1. *The RMLO problem is* APX-*hard even on DAGs.*

4 A strongly polynomial $O(n)$-approximation algorithm

For symmetric SNPGs, Briest *et. al* [4] proved the existence of an algorithm, called the *single-price algorithm*, that guarantees an approximation of $(1+\epsilon)(\mathcal{H}_k + \mathcal{H}_{|P|})$, where k denotes the number of followers. Notice that this result implies an $O(\log n)$-approximation for the symmetric version of the Stackelberg SPT game studied in [3]. Therefore, a simple question is whether the single-price algorithm provides a good approximation also for the ASSPT game. Not surprisingly, this is not the case, as illustrated in Figure 5, where we give an instance for which the algorithm returns an $(n-3)$-approximate solution. Thus, the power of the single-price algorithm seems to rely on the alignment of the leader's and follower's objective functions.

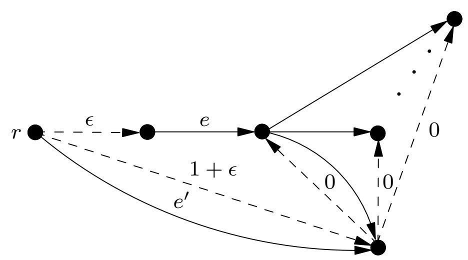

Fig. 5. An example where the single-price algorithm, i.e., the algorithm that eventually will price all the edges with a same value $\sigma \geq 0$, does not return a better than $(n-3)$-approximate solution. Dashed edges are fixed-cost edges, while the other edges are owned by the leader. Pricing e with 0, e' with $1+\epsilon$, and all the other edges with 1 yields a revenue of $n-3$. On the other hand, pricing all edges with any value $\sigma \geq 0$ yields a revenue of only σ.

Besides being only $\Omega(n)$-approximating, the single-price algorithm has also another drawback, namely it is not *strongly polynomial*, since it is polynomial in the input size, but its running time depends on the given edge costs. Indeed, for our problem, it requires the testing of $1 + \frac{\log\lceil c_0 \rceil}{\log(1+\epsilon)}$ different weights for the priceable edges, where c_0 is the cost of a cheapest feasible solution not containing leader's edges. Therefore, in an effort of improving on that, we developed the following simple strategy the leader can play in order to get in strongly polynomial-time an $(n-1)$-approximation of the maximum achievable revenue. Actually, the goal of closing the $O(\sqrt{n})$ gap left open w.r.t. the corresponding inapproximability result remains a challenging open problem.

Theorem 5. *For the* ASSPT *problem, the pricing function p that prices each edge $e = (u,v) \in P$ with $p(e) = \min\{0, d_{G_C}(r,v) - d_{G_C}(r,u)\}$ is computable in $O(m+n\log n)$ time and yields a revenue of at least $\frac{1}{n-1}R^*$, where d_{G_C} denotes distances computed in $G_C = (V,C)$, and R^* denotes the revenue yielded by an optimal pricing.*

Proof. Let p^* be a pricing such that $f_1(p^*, H(p^*)) = R^*$. For a path π, we denote by $\texttt{revenue}(\pi)$ the value $\sum_{e \in E(\pi)\cap P} p^*(e)$. Let π be a path in $H(p^*)$ outgoing from r such that $\texttt{revenue}(\pi) \geq \texttt{revenue}(\pi')$, for every other path π' of $H(p^*)$ outgoing from r. Clearly, $\texttt{revenue}(\pi) \geq \frac{1}{n-1}R^*$. In what follows we show that $f_1(p, H(p)) \geq \texttt{revenue}(\pi)$.

Let $e_1, \ldots, e_k$ denote the edges of $E(\pi) \cap P$ in the order in which they appear if we traverse π starting from r. Let $e_i = (u_i, v_i)$, for every $i = 1, \ldots, k$. Moreover, for every $i = 1, ..., k$, let us denote by ℓ_i the length of the (fixed-cost) subpath of π going from v_{i-1} to u_i, where $v_0 = r$. We have that

$$\forall i = 1, \ldots, k-1,\ d_{G_C}(r, v_i) + \ell_{i+1} \geq d_{G_C}(r, u_{i+1}).$$

Summing over all i's, adding up the two equalities $d_{G_C}(r, u_1) = \ell_1$ and $d_{G_C}(r, v_k) = d_{G_C}(r, v_k)$, and rearranging the terms, we obtain

$$\sum_{i=1}^{k} \big(d_{G_C}(r, v_i) - d_{G_C}(r, u_i)\big) \geq d_{G_C}(r, v_k) - \sum_{i=1}^{k} \ell_i. \tag{3}$$

Since π is a shortest path in G w.r.t. p^*, we have that $\texttt{revenue}(\pi) + \sum_{i=1}^{k} \ell_i \leq d_{G_C}(r, v_k)$, from which we get

$$\texttt{revenue}(\pi) \leq d_{G_C}(r, v_k) - \sum_{i=1}^{k} \ell_i. \tag{4}$$

Furthermore, notice that $d_{G_C}(r, v) = d_{H(p)}(r, v)$ for every vertex v. Next, observe that every edge e with $p(e) > 0$ is contained in a SPT of G w.r.t. p. As a consequence, there is a SPT of G w.r.t. p that contains all the edges e_i's for which $p(e_i) > 0$. Therefore,

$$f_1(p, H(p)) \geq \sum_{i=1}^{k} \big(d_{G_C}(r, v_i) - d_{G_C}(r, u_i)\big). \tag{5}$$

Combining inequalities (3), (4), and (5), we obtain $f_1(p, H(p)) \geq \texttt{revenue}(\pi)$. From this and from the fact that distances from r in G_C can be easily computed in $O(|C| + n\log n)$, the claim follows. $\square$

We conclude by observing that the upper-bound of the approximation ratio is asymptotically tight. Indeed, the digraph in Figure 5 without edge e', shows an example where pricing all edges according to the formula given in the above theorem does not return a better than $(n-3)$-approximate solution.

References

1. P. Alimonti and V. Kann, Some APX-completeness results for cubic graphs, *Theoretical Computer Science*, 237(1-2):123–134 (2000).
2. N Alon, F.V. Fomin, G. Gutin, M. Krivelevich, and S. Saurabh, Spanning directed trees with many leaves, *SIAM J. Discrete Mathematics*, 23(1): 466–476 (2009).
3. D. Bilò, L. Gualà, G. Proietti, and P. Widmayer, Computational aspects of a 2-player Stackelberg shortest paths tree game, *Proc. of the 4th Int. Workshop on Internet & Network Economics (WINE)*, Shanghai, China, December 17-20, 2008, Vol. 5385 of Lecture Notes in Computer Science, Springer-Verlag, 251–262.
4. P. Briest, M. Hoefer, and P. Krysta, Stackelberg network pricing games, *Proc. of the 25th Ann. Symp. on Theoretical Aspects of Computer Science (STACS)*, `http://drops.dagstuhl.de/opus/volltexte/2008/1340`, 133–142, 2008.
5. J. Cardinal, E.D. Demaine, S. Fiorini, G. Joret, S. Langerman, I. Newman, and O. Weimann, The Stackelberg minimum spanning tree game, *Proc. of the 10th Int. Workshop on Algorithms and Data Structures (WADS)*, Vol. 4619 of Lecture Notes in Computer Science, Springer, 64–76, 2007.
6. R. Cole, Y. Dodis, and T. Roughgarden, Pricing network edges for heterogeneous selfish users, *Proc. 35th Ann. ACM Symp. on Theory of Computing (STOC)*, 521-530, 2003.
7. J. Daligault and S. Thomasse, On finding directed trees with many leaves, *Computing Research Repository (CoRR)*, `http://arxiv.org/abs/0904.2658`, 2009.
8. G. Galbiati, F. Maffioli, and A. Morzenti, A short note on the approximability of the maximum leaves spanning tree problem, *Information Processing Letters*, 52(1): 45-49 (1994).
9. A. Grigoriev, S. van Hoesel, A. van der Kraaij, M. Uetz, and M. Bouhtou, Pricing network edges to cross a river, *Proc. of the 3rd Workshop on Approximation and Online Algorithms (WAOA)*, Vol. 3351 of Lecture Notes in Computer Science, Springer, 140–153, 2005.
10. G. Joret, Stackelberg network pricing is hard to approximate, manuscript, 2009.
11. M. Labbé, P. Marcotte, and G. Savard, A bilevel model of taxation and its application to optimal highway pricing, *Management Science*, 44(12):608–622, 1998.
12. R. Raz and S. Safra, A sub-constant error-probability low-degree test, and sub-constant error-probability PCP characterization of NP, *Proc. 29th Ann. ACM Symp. on Theory of Computing (STOC)*, 475–484, 1997.
13. S. Roch, G. Savard, and P. Marcotte, An approximation algorithm for Stackelberg network pricing, *Networks*, 46(1):57–67, 2005.
14. S. van Hoesel, An overview of Stackelberg pricing in networks, Research Memoranda 042, Maastricht: METEOR, Maastricht Research School of Economics of Technology and Organization, 2006.
15. H. von Stackelberg, *Marktform und Gleichgewicht (Market and Equilibrium)*, Verlag von Julius Springer, Vienna, Austria, 1934.
16. D. Zuckerman, Linear degree extractors and the inapproximability of max clique and chromatic number, *Theory of Computing* 3(1):103–128, 2007.

ECCC ISSN 1433-8092

http://eccc.hpi-web.de