Prize-Collecting Steiner Networks via Iterative Rounding

MohammadTaghi Hajiaghayi[1] and Arefeh A. Nasri[2]

[1] AT&T Labs– Research, Florham Park, NJ 07932, U.S.A.
hajiagha@research.att.com
[2] Rutgers University, New Brunswick, NJ 08901, U.S.A.
aanasri@gmail.com

Abstract. In this paper we design an iterative rounding approach for the classic *prize-collecting Steiner forest problem* and more generally the *prize-collecting survivable Steiner network design* problem. We show as an structural result that in each iteration of our algorithm there is an LP variable in a basic feasible solution which is at least one-third-integral resulting a 3-approximation algorithm for this problem. In addition, we show this factor 3 in our structural result is indeed tight for prize-collecting Steiner forest and thus prize-collecting survivable Steiner network design. This especially answers negatively the previous belief that one might be able to obtain an approximation factor better than 3 for these problems using a natural iterative rounding approach. Our structural result is extending the celebrated iterative rounding approach of Jain [13] by using several new ideas some from more complicated linear algebra. The approach of this paper can be also applied to get a constant factor (bicriteria-)approximation algorithm for degree constrained prize-collecting network design problems.

We emphasize that though in theory we can prove existence of only an LP variable of at least one-third-integral, in practice very often in each iteration there exists a variable of integral or almost integral which results in a much better approximation factor than provable factor 3 in this paper (see patent application [11]). This is indeed the advantage of our algorithm in this paper over previous approximation algorithms for prize-collecting Steiner forest with the same or slightly better provable approximation factors.

1 Introduction

Consider a mailing company that wishes to ship packages overnight between several pairs of cities. To this end, this company can build connecting carriers between cities such that at the end by scheduling the carriers, the company is able to ship the packets overnight between pairs of connected cities. Assume the cost of connecting city i to city j is c_{ij} and the costs are symmetric. In addition, the company has the choice of leasing other companies for some pairs (i, j) of cities with cost π_{ij} so that without any worry the leased company do the shipment between cities i and j overnight. The goal is to build some carriers and lease some other companies such that the company do the shipments overnight with minimum total cost.

The above network design problem which has also several applications in expanding telecommunications and transportation networks (see e.g. [15,20]), and cost sharing and Lagrangian relaxation techniques (see e.g. [14,6]) is called the *prize-collecting*

A. López-Ortiz (Ed.): LATIN 2010, LNCS 6034, pp. 515–526, 2010.

Steiner forest (PCSF) problem[1]. In this problem, given a graph $G = (V, E)$, a set of (commodity) pairs $\mathcal{P} = \{(s_1, t_1), (s_1, t_1), \ldots, (s_\ell, t_\ell)\}$, a non-negative cost function $c : E \rightarrow \mathbf{Q}_+$, and finally a non-negative penalty function $\pi : \mathcal{P} \rightarrow \mathbf{Q}_+$, our goal is a minimum-cost way of buying a set of edges and paying the penalty for those pairs which are not connected via bought edges. When all sinks are identical in the PCSF problem, it is the classic prize-collecting Steiner tree problem. Bienstock, Goemans, Simchi-Levi, and Williamson [5] first considered this problem (based on a problem earlier proposed by Balas [2]) for which they gave a 3-approximation algorithm. The current best approximation algorithm for this problem is a primal-dual $2 - \frac{1}{n-1}$ approximation algorithm (n is the number of vertices of the graph) due to Goemans and Williamson [7]. The general form of the PCSF problem first has been formulated by Hajiaghayi and Jain [12]. They showed how by a primal-dual algorithm to a novel integer programming formulation of the problem with doubly-exponential variables, we can obtain a 3-approximation algorithm for the problem (see also [10]). In addition, they show that the factor 3 in the analysis of their algorithm is tight. However they show how a direct randomized LP rounding algorithm with approximation factor 2.54 can be obtained for this problem. Their approach has been generalized by Sharma, Swamy, and Williamson [21] for network design problems where violated arbitrary 0-1 connectivity constraints are allowed in exchange for a very general penalty function. The work of Hajiaghayi and Jain has also motivated a game-theoretic version of the problem considered by Gupta et al. [8].

In this paper, we also consider a generalized version of prize-collecting Steiner forest, called *prize-collecting survivable Steiner network design*, in which we are also given connectivity requirements r_{uv} for all pairs of vertices u and v and a non-increasing marginal penalty function for u and v in case we cannot satisfy all r_{uv}. Our goal is to find a minimum way of constructing a network (graph) in which we connect u and v with $r'_{uv} \leq r_{uv}$ edge-disjoint paths and paying the marginal penalty for $r_{uv} - r'_{uv}$ violated connectivity between u and v. When all penalties are ∞, the problem is the classic survivable Steiner network design problem. For this problem, Jain [13] using the method of iterative rounding obtains a 2-approximation algorithm improving on a long line of earlier research that applied primal-dual methods to this problem.

In this paper, for the first time, we are using the iterative rounding approach for prize-collecting versions of Steiner forest and more generally survivable Steiner network design. To the best of our knowledge, so far this method of iterative rounding has not been used for any prize-collecting problem. After several years since Jain's work, the method of iterative rounding has been revived recently to obtain the best possible bicriteria $(1, B_v + 1)$-approximation algorithm for minimum bounded-degree spanning trees [23] (B_v is the degree bound on vertex v) and minimum-bounded degree variants of other problems such as arborescence, Steiner forest and survivable Steiner network design [18,3,19]. The approach of iterative rounding in this paper can be extended further for other prize-collecting problems such as prize-collecting survivable network design with degree constraints B_v on each vertex (i.e., in our solution we should buy at most B_v edges attached to each vertex v) to get factor 3 (bicriteria-)approximation algorithms.

[1] In the literature, they also called this problem *prize-collecting generalized Steiner tree (PCGST).*

1.1 Our Results

In this paper, we are extending our current knowledge of iterative rounding approaches to prize-collecting Steiner forest and more generally survivable Steiner network design. For the sake of presentation, after introducing the novelty of our approach by stating it precisely for prize-collecting Steiner forest, then we show how it can be extended for prize-collecting survivable Steiner network design. Note that as mentioned in the introduction, so far the only approach to obtain a constant factor approximation algorithm for the survivable Steiner network design, a special case of the prize-collecting survivable Steiner network design problem in which all penalties are ∞, is the method of iterative rounding. Other approaches such as primal-dual methods do not consider the global structure of the network enough to be used for this problem.

We first show as an structural result that in a natural LP for prize-collecting Steiner forest, either a variable corresponding to an edge or a variable corresponding to a penalty for a pair is at least one-third-integral in any basic feasible solution (see Section 3). Indeed we also show this variable of one-third-integral is best that one can hope in a basic feasible solution (see Section 5). This one-third-integral bound obtains a 3-approximation for this problem via much stronger structural results (see Section 2).

There are several novelties in our approach of iterative rounding for the PCSF problem mostly coming from linear algebra. First, so far in all iterative rounding approaches the main constraint is that the fractional value of a cut corresponding to a set S is at least a submodular function of S. This has been relaxed in our setting where the fractional value of a cut is also a (not necessarily submodular) function of a penalty associated with a commodity pair separated by this cut. Second, in all previous iterative rounding approaches (in which indeed the heart is obtaining a laminar family using linear algebra, first introduced by Jain [13]) the linear dependence between constraints is a simple addition with all coefficients having absolute values ones (see Theorem 3, Part 5). We show a more complicated fractional dependence between constraints which is crucial to our results. Third, our approach of constructing a laminar family is more complicated than previous approaches when we replace a constraint with one of five (instead of two in previous approaches) constraints (see Theorem 4). Last but not least, obtaining a variable of at least one-third-integral in previous approaches (see e.g. Jain [13]) is relatively easy, however in our case it is much more complicated and needs new ideas from linear algebra (see Theorem 5). Subsequent and separate to our work Konemann et al. [22] obtain the same iterative algorithm as ours for PCSF with some proofs simplified.

After presenting our one-third-integral result for the PCSF problem (which results in a 3-approximation), we show how we can generalize this approach to obtain a variable of at least one-third-integral (and thus a 3-approximation algorithm) for the minimum prize-collecting survivable Steiner network design problem. We briefly discuss the case in which we also have degree constraints on bought edges.

Finally we should emphasize that though in theory we can prove existence of only an LP variable of at least one-third-integral, in practice very often in each iteration there exists a variable of integral or almost integral which results in a much better approximation factor than provable factor 3 in this paper (see AT&T patent application [11] on this regard). This is indeed the advantage of our algorithm in this paper over previous approximation algorithms for prize-collecting Steiner forest with the same or slightly better provable approximation factors.

2 Iterative Rounding Approximation Algorithm

The traditional LP relaxation for the PCSF problem which can be solved using Ellipsoid algorithm[2] is as follows:

$$\text{OPT} = \quad \text{minimize} \sum_{e\in E} c_e x_e \quad + \sum_{(i,j)\in\mathcal{P}} \pi_{ij} z_{ij} \tag{1}$$

$$\text{subject to} \quad \sum_{e\in\delta(S)} x_e + z_{ij} \geq 1 \quad \forall S \subset V, (i,j) \in \mathcal{P}, S \odot (i,j) \tag{2}$$

$$x_e \geq 0 \qquad \forall e \in E \tag{3}$$

$$z_{ij} \geq 0 \qquad \forall (i,j) \in \mathcal{P} \tag{4}$$

Here for a set $S \subset V$, we denote $|\{i,j\} \cap S| = 1$ by $S \odot (i,j)$.

Let x^*, z^* be an optimal basic feasible solution for LP 1. For $0 < \alpha \leq 1$, let E_α be the set of edges whose value in x^* is at least α and let $\mathcal{P}_\alpha$ be the set of edges whose value in z^* is at least α. We define $G_{res} = E - E_\alpha$ and $\mathcal{P}_{res} = \mathcal{P} - \mathcal{P}_\alpha$. Now we consider the following LP, called the *residual LP*, in which we fix all values in edges in E_α and pairs in $\mathcal{P}_\alpha$ to be 1.

$$\text{OPT}_{res} = \quad \text{minimize} \sum_{e\in E} c_e x_e \quad + \sum_{(i,j)\in\mathcal{P}_{res}} \pi_{ij} z_{ij} \tag{5}$$

$$\text{subject to} \quad \sum_{e\in\delta(S)} x_e + z_{ij} \geq 1 \quad \forall S \subset V, (i,j) \in \mathcal{P}_{res}, \tag{6}$$

$$S \odot (i,j), \delta(S) \cap E_\alpha = \emptyset$$

$$x_e \geq 0 \qquad \forall e \in E_{res} \tag{7}$$

$$z_{ij} \geq 0 \qquad \forall (i,j) \in \mathcal{P}_{res} \tag{8}$$

Note that in the above LP by contracting edges in E_α and ignoring pairs in $\mathcal{P}_\alpha$, indeed we can always work with an LP similar to that for OPT. Our approximation algorithm for the PCSF problem based on this LP is as follows.

Algorithm PCSF-ALG which is based on the the following theorem is as follows: First we find an optimal basic feasible solution x^*, z^* to LP 1. Then we pay all the penalties of pairs (i,j) whose $z^*_{ij} \geq \alpha$ and remove them from further consideration. We include all edges e whose $x^*_e \geq \alpha$ in the solution and contract them and remove multiple edges by keeping only an edge e with minimum c_e among them. We solve the residual problem recursively.

[2] Indeed we can also write the corresponding standard flow-based LP rather than the cut-based LP here, and then use other LP-solver algorithms for a polynomial number of variables and constraints.

Theorem 1. *In any basic feasible solution for LP 1, for at least one edge* $e \in E$, $x_e >= \frac{1}{3}$, *or for at least one pair* $(i,j) \in \mathcal{P}$, $z_{ij} >= \frac{1}{3}$.

We prove Theorem 1 in Section 3.

Theorem 2. *If* x^I, z^I *is an integral solution to the LP 5 with value at most* $\frac{1}{\alpha}\text{OPT}_{res}$, *then* $E_{x_e=1} \cup E_{\alpha}$, $P_{z_{ij}=1} \cup \mathcal{P}_{\alpha}$ *is feasible solution for LP 1 with value at most* $\frac{1}{\alpha}\text{OPT}$.

The proof of Theorem 2 is standard and hence omitted. By combining Theorems 1 and 2 we obtain the following conclusion:

Corollary 1. *There is an iterative rounding* 3*-approximation algorithm for PCSF.*

3 One-Third-Integrality Result

In this section, we prove Theorem 1. Let x, z be a basic feasible solution. If for an edge e, $x_e = 1$ or for a pair (i', j'), $z_{ij} = 1$, then the theorem follows. Also, if for an edge e, $x_e = 0$, then we can assume that the edge was never there before. This assumption does not increase the cost of the optimum fractional solution x_e . Thus we can assume that $0 < x_e < 1$ and $0 \leq z_e < 1$ for all $e \in E$ and $(i,j) \in \mathcal{P}$.

Let $\mathcal{M}(S, ii')$ be the row of the constraint matrix corresponding to a set $S \subset V$ and pair $(i,i') \in \mathcal{P}$. Let $x(A,B)$ be the sum of all x_e's, where e has one end in A and the other end in B. We represent $x(A, \overline{A})$ by $x(A)$, for ease of notation. We say a set A is tight with pair (i,i') if $A \odot (i,i')$ and $x(A) + z_{ii'} = 1$.

Theorem 3. *If* A *is tight with* (i,i') *and* B *is tight with* (j,j') *then at least one of the following holds:*

1. $A-B$ *is tight with* (i,i'), $B-A$ *is tight with* (j,j') *and* $\mathcal{M}(A,ii')+\mathcal{M}(B,jj') = \mathcal{M}(A-B,ii')+\mathcal{M}(B-A,jj')$.
2. $A-B$ *is tight with* (j,j'), $B-A$ *is tight with* (i,i') *and* $\mathcal{M}(A,ii')+\mathcal{M}(B,jj') = \mathcal{M}(A-B,jj')+\mathcal{M}(B-A,ii')$.
3. $A\cap B$ *is tight with* (i,i'), $A\cup B$ *is tight with* (j,j') *and* $\mathcal{M}(A,ii')+\mathcal{M}(B,jj') = \mathcal{M}(A\cap B,ii')+\mathcal{M}(A\cup B,jj')$.
4. $A\cap B$ *is tight with* (j,j'), $A\cup B$ *is tight with* (i,i') *and* $\mathcal{M}(A,ii')+\mathcal{M}(B,jj') = \mathcal{M}(A\cap B,jj')+\mathcal{M}(A\cup B,ii')$.
5. $A-B$ *is tight with* (i,i'), $B-A$ *is tight with* (i',i), $A\cap B$ *is tight with* (j,j'), $A\cup B$ *is tight with* (j,j') *and* $2\mathcal{M}(A,ii')+2\mathcal{M}(B,jj') = \mathcal{M}(A-B,ii')+\mathcal{M}(B-A,ii')+\mathcal{M}(A\cap B,jj')+\mathcal{M}(A\cup B,jj')$.

Proof. The proof is by case analysis. For the ease of notation, if a set A is tight with pair (i,i'), we assume $i \in A$ (and thus $i' \notin A$).

We consider two cases $i \in A-B$ and $i \in A\cap B$. Without loss of generality, we assume in the latter case $j \in A\cap B$ also (otherwise we consider j instead of i in our arguments). Because of tightness we have:

$$x(A)=x(A-B,B-A)+x(A-B,\overline{A\cup B})+x(A\cap B,B-A)+x(A\cap B,\overline{A\cup B})=1-z_{ii'}$$

$$x(B)=x(B-A,A-B)+x(B-A,\overline{A\cup B})+x(A\cap B,A-B)+x(A\cap B,\overline{A\cup B})=1-z_{jj'}$$

Let's first start with the case in which $i \in A \cap B$ (and thus $j \in A \cap B$). In this case $i' \in \overline{A \cup B}$ and $j' \in \overline{A \cup B}$. Because of the feasibility:

$$x(A \cap B) = x(A \cap B, A - B) + x(A \cap B, B - A) + x(A \cap B, \overline{A \cup B}) \geq 1 - z_{ii'}$$

$$x(A \cup B) = x(A - B, \overline{A \cup B}) + x(A \cap B, \overline{A \cup B}) + x(B - A, \overline{A \cup B}) \geq 1 - z_{jj'}$$

Since $x(.,.) \geq 0$, by summing up the two inequalities above and using the equalities for $x(A)$ and $x(B)$, we conclude that the inequalities should be tight, i.e., $x(A \cap B) = 1 - z_{ii'}$ and $x(A \cup B) = 1 - z_{jj'}$ and in addition $x(A - B, B - A) = 0$, i.e., $\mathcal{M}(A, ii') + \mathcal{M}(B, jj') = \mathcal{M}(A \cap B, ii') + \mathcal{M}(A \cup B, jj')$. Thus we are in the case 3 of the statement of the theorem.

Now assume that $i \in A - B$ and $j \in B - A$. Then independent of the place of i', j', by the feasibility of the solution we have:

$$x(A - B) = x(A - B, A \cap B) + x(A - B, B - A) + x(A - B, \overline{A \cup B}) \geq 1 - z_{ii'}$$

$$x(B - A) = x(B - A, A - B) + x(B - A, A \cap B) + x(B - A, \overline{A \cup B}) \geq 1 - z_{jj'}$$

Since $x(.,.) \geq 0$, by summing up the two inequalities above and using the equalities for $x(A)$ and $x(B)$, we conclude that the inequalities should be tight, i.e., $x(A - B) = 1 - z_{ii'}$ and $x(B - A) = 1 - z_{jj'}$ and in addition $x(A \cap B, \overline{A \cup B}) = 0$, i.e., $\mathcal{M}(A, ii') + \mathcal{M}(B, jj') = \mathcal{M}(A - B, ii') + \mathcal{M}(B - A, jj')$. Thus we are in the case 1 of the statement of the theorem.

Finally we consider the case in which $i \in A - B$ and $j \in A \cap B$ (and thus $j' \in \overline{A \cup B}$).

Now if $i' \in \overline{A \cup B}$, then by the feasibility of the solution we have:

$$x(A \cap B) = x(A \cap B, A - B) + x(A \cap B, B - A) + x(A \cap B, \overline{A \cup B}) \geq 1 - z_{jj'}$$

$$x(A \cup B) = x(A - B, \overline{A \cup B}) + x(A \cap B, \overline{A \cup B}) + x(B - A, \overline{A \cup B}) \geq 1 - z_{ii'}$$

Since $x(.,.) \geq 0$, by summing up the two inequalities above and using the equalities for $x(A)$ and $x(B)$, we conclude that the inequalities should be tight, i.e., $x(A \cap B) = 1 - z_{jj'}$ and $x(A \cup B) = 1 - z_{ii'}$ and in addition $x(A - B, B - A) = 0$, i.e., $\mathcal{M}(A, ii') + \mathcal{M}(B, jj') = \mathcal{M}(A \cap B, jj') + \mathcal{M}(A \cup B, ii')$. Thus we are in the case 4 of the statement of the theorem.

Finally if $i' \in B - A$ then, because of feasibility we have

$$x(A - B) = x(A - B, A \cap B) + x(A - B, B - A) + x(A - B, \overline{A \cup B}) \geq 1 - z_{ii'}$$

$$x(A \cap B) = x(A \cap B, A - B) + x(A \cap B, B - A) + x(A \cap B, \overline{A \cup B}) \geq 1 - z_{jj'}$$

$$x(B - A) = x(B - A, A - B) + x(B - A, A \cap B) + x(B - A, \overline{A \cup B}) \geq 1 - z_{jj'}$$

$$x(A \cup B) = x(A - B, \overline{A \cup B}) + x(A \cap B, \overline{A \cup B}) + x(B - A, \overline{A \cup B}) \geq 1 - z_{ii'}$$

Since $x(.,.) \geq 0$, by summing up the four inequalities above and and use the equalities for $2x(A)$ and $2x(B)$, we conclude that all inequalities should be tight, and in addition $x(A - B, B - A) = 0$ and $x(A \cap B, \overline{A \cup B}) = 0$, i.e., and $2\mathcal{M}(A, ii') + 2\mathcal{M}(B, jj') = \mathcal{M}(A - B, ii') + \mathcal{M}(B - A, jj') + \mathcal{M}(A \cap B, ii') + \mathcal{M}(A \cup B, jj')$. So the case 5 of the statement of the theorem holds. □

Note that especially Case 5 in Theorem 3 is novel to our extension of iterative rounding methods.

Let $\mathcal{T}$ be the set of all tight constraints. For any set of tight constraints $\mathcal{F}$, we denote the vector space spanned by the vectors $\mathcal{M}(S, ii')$, where $S \subset V$ and $(i, i') \in \mathcal{P}$, by $Span(\mathcal{F})$. We say two sets A and B *cross* if none of the sets $A - B$, $B - A$ and $A \cap B$ is empty. We say a family of tight constraints is *laminar* if no two sets corresponding to two constraints in it cross.

The proof of the following theorem is similar to that of Jain [13] and hence omitted.

Theorem 4. *For any maximal laminar family $\mathcal{L}$ of tight constraints, $Span(\mathcal{L}) = Span(\mathcal{T})$.*

Since x, z is a basic feasible solution, the dimension of $Span(\mathcal{T})$ is $|E(G)|+|\mathcal{P}|$. Since $Span(\mathcal{L}) = Span(\mathcal{T})$, it is possible to choose a basis for $Span(T)$ from the vectors in $\{M(S, ii)\} \in \mathcal{L}$. Let $\mathcal{B} \subseteq \mathcal{L}$ forms a basis for $Span(\mathcal{T})$. Hence we have the following theorem.

Corollary 2. *There exists a laminar family, $\mathcal{B}$, of tight constraints satisfying 1) $|\mathcal{B}| = |E(G)|+|\mathcal{P}|$; 2) The vectors in $\mathcal{B}$ are independent; and 3) All constraints in $\mathcal{B}$ are tight.*

Note that in our laminar family if a set S is tight with both (i, i') and (j, j') in two different constraints, since $z_{ii'} = z_{jj'}$, we can remove variable $z_{jj'}$ and just use $z_{ii'}$ instead. Since we removed one variable and one constraint, still we have a basic feasible solution which is laminar. By this reduction, we always can make sure that each set is tight with only one pair. Thus a tight set uniquely determines the tight pair and we use a tight constraint and a tight set interchangeably in our discussion below.

Now we are ready to prove Theorem 1.

Theorem 5. *In any basic feasible solution for LP 1, for at least one edge $e \in E$, $x_e \geq \frac{1}{3}$, or for at least one pair $(i, j) \in \mathcal{P}$, $z_{ij} \geq \frac{1}{3}$.*

Proof. We are giving a token to each end-point of an edge (and thus two tokens for an edge) and two tokens to all z variables (notice that some z variables are used for more than one commodity pairs as discussed above). Now, we will distribute the tokens such that for every set in the laminar family gets at least two tokens and every root at least four tokens unless the corresponding cut has exactly three edges. (note that each cut has at least three edges since the value of each variable is less than $\frac{1}{3}$) in which the root gets at least three token. This contradict the equality $|V(F)| = |E(G)| + |\mathcal{P}|$ where F is the rooted forest of laminar sets in the laminar family. The subtree of F rooted at R consists of R and all its descendants. We will prove this result by the induction on every rooted subtree of F.

Consider a subtree rooted at R. Since all x_e and z_{ij} are at most $\frac{1}{3}$, if R is a leaf node, it has at least three edges crossing it and thus gets at least three tokens (and more than 3 tokens if the degree is more than 3). This means the induction is correct for a leaf node, as the basis of the induction.

If R has four or more children, by the induction hypothesis each child has at least three tokens and each of their descendants gets at least two tokens. We re-assign one extra token from each child to the node R. Thus R has at least four tokens and the induction hypothesis is correct in this case.

If R has three children, if there is a *private vertex* u to R, i.e., a vertex which is in R but not in any of its children, then we are done (since all x_e values are fractional, the degree of u is at least two and thus can contributes at least two extra tokens toward R). Also if one of the children has at least four edges in its corresponding cut, by the induction hypothesis it has at least two extra tokens to contribute toward those of R and we are done.

Next, if R has exactly three children each with exactly three edges in its corresponding cut, then by parity R has an odd number of edges in its corresponding cut. If R has edges in the its cut then the three extra token by its children suffices. If R has seven or nine edges in the cut, then at least one of its children has all three edges in the cut and the corresponding pair is not satisfied. But this means all other edges than those of this cut should be zero which is contradiction to fractional value assumption. Now if R tight with $z_{pp'}$ has exactly five edges in the cut, it should be the case that two children C_1 tight with $z_{ii'}$ and C_3 tight with $z_{kk'}$ have two edges in the cut and C_2 tight with $z_{jj'}$ has one edges in the cut. Note that in this case $z_{pp'} > \min\{z_{ii'}, z_{jj'}, z_{kk'}\}$ then at least for one of $z_{ii'}$, $z_{jj'}$, and $z_{kk'}$ all pairs should be inside R for the first time and thus we have at least two extra tokens towards the requirement of R and we are done. It also means that p should be inside the child C with $\min\{z_{ii'}, z_{jj'}, z_{kk'}\}$ and it should be equal to its corresponding z value (otherwise child C violates the condition for $z_{pp'}$). Assume that $z_{pp'} = z_{jj'}$. In this case, it is easy to see that since C_2 is tight with three edges and with five edges, the sum of x variables of C_1 and C_3 in the cut R is equal to the sum of x variables of C_1 and C_3 to C_2. But it means at least for one of C_1 and C_3, the edge e to C_2 has $x_e \geq x_{e'} + x_{e''}$ where e' and e'' are the edges in the cut R. But since $x_e + x_{e'} + x_{e''} > \frac{2}{3}$ (due to the fact that all z variables are less than $\frac{1}{3}$), $x_e \geq \frac{1}{3}$ which is a contradiction. If $z_{pp} = z_{ii'} \leq z_{kk'}$ where $z_{ii'} < z_{jj'}$. In this case the edge from C_1 to C_2 should has an x value equal to that those edges of C_2 and C_3 in the cut R. It means the total x value of two edge of C_3 in the cut is less than $\frac{1}{3}$ which is a contradiction, since the third edge has x value at least $\frac{1}{3}$.

Now we consider the case in which R is tight with $z_{pp'}$ and has two children . If there is a private vertex u to R we have at least four tokens to satisfy R (two from u and one from each of its children). If both of these children have degree at least four, then we have four extra tokens for R (two from each child). Then at least one of two children, namely C_1 tight with $z_{ii'}$, has exactly three edges in its corresponding cut. The other child C_2 tight with $z_{jj'}$ has at least three edges in the cut. Note that in this case $z_{pp'} > \min\{z_{ii'}, z_{jj'}\}$ then at least for one of $z_{ii'}$ and $z_{jj'}$ all pairs should be inside R for the first time and thus we have at least two extra tokens towards the requirement of R and we are done. It also means that p should be inside the child C with $\min\{z_{ii'}, z_{jj'}, z_{kk'}\}$ and it should be equal to its corresponding z value (otherwise child C violates the condition for $z_{pp'}$). First assume that $z_{pp'} = z_{ii'} \leq z_{jj'}$. In this case it is not possible that all three edges of C_1 are in the cut R, since then all edges of C_2 are in the cut R and they are zero (since the cut R is already tight with $z_{pp'} = z_{ii'}$). If C_1 has two edges in the cut, since $z_{ii'} \leq z_{jj'}$, it means sum of the x values of the edges in the cut corresponding to R, which has one edges from C_1 and the rest are the edges of C_2 in the R cut, should be at least the sum of x values of the cut corresponding to C_2. But these means the x value of the edge of C_1 in the R cut is at least the sum of x values of the two edges from C_1 to C_2. Since value of all three edges in C_1 is at least

$\frac{2}{3}$, it means the x value of the edge of C_1 in the R cut is at least $\frac{1}{3}$, a contradiction. In case C_1 does not have any edges in the cut R, then all edges should go C_2 which means $z_{ii'}$ should be tight with a proper subset of edges of C_2 though we know that x values of all edges of C_2 is at most $1 - z_{jj}$, a contradiction. we know even all edges of minus those edges should be tight in R with the same $z_{jj'}$ which means all edges between C_1 and C_2 should be zero which is a contradiction.

Next assume that $z_{pp'} = z_{jj'} \leq z_{ii'}$. In this case it is not possible that all edges of C_2 and thus C_1 are in the cut since all edges of C_1 should have zero x value. In case if one edge of C_1 or two edges of C_1 are in the cut R, then x value of one edge of C_1 is equal to the x value of two edges of C_1 which means that edge should have x value at least $\frac{1}{3}$ which is a contradiction. In case C_1 does not have any edges in the cut R, then C_2 minus those edges should be tight in R with the same $z_{jj'}$ which means all edges between C_1 and C_2 should be zero which is a contradiction.

It only remains the case in which R has only one child C. In this case if R and C are both tight with respect to $z_{ii'}$ then since R and C are independent there is a vertex $u \in R - C$. However, if R is tight with $z_{ii'}$ and C is tight with $z_{jj'}$ since $z_{ii'} \neq z_{jj'}$, these two cuts should be different and thus again there is vertex $u \in R - C$. Since all x_e values are fractional the degree of u is at least two and thus u gets at least two tokens. Without loss of generality assume u is the node with maximum degree. If u has degree at least three then we can assign at least these three private tokens of u and at least one extra token of C to R to have the induction hypothesis satisfied. In case u has degree two and C has at least four edges in the cut, then we have at least two tokens from u and two extra tokens from C to assign at least four tokens to R and satisfy the hypothesis.

The only remaining case when R has only one child is when u has degree two and C has an odd number of edges in its cut. However in this case because of parity, R should have an odd number of edges in its cut (note that in this case, we may have some other vertices than u of degree two in $R - C$.) If this odd number is three then two tokens of u and one extra token of C satisfies the required number of tokens for R. If there is a vertex other than u in R_C it has also two extra tokens and we are done. The only case is that u has degree two, C has three edges and all these five edges are in the cut corresponding to R. It means in this case R should be tight with $z_{ii'}$ and C should be tight with $z_{jj'}$ where $z_{ii'} < z_{jj'}$ (otherwise the edges from u in the cut should zero which is a contradiction to the fractional values for x_es). Here $i \neq u$ otherwise, u has degree three and thus three extra tokens and we have at least four tokens for R. It means $i \in C$ which is again a contradiction since the current cut for C violates the cut condition for i in the LP. □

Finally, it is worth mentioning though we guarantee that during the course of the algorithm, we can get a variable which is only one-third-integral, in the first iteration always we can find an integral z variable. Below there is a more general proposition regarding this issue.

Proposition 1. *If there is a set S of fractional variables which contains exactly one variable from each tight constraint in our laminar family, our solution cannot be a basic optimum solution. In particular, there is no basic optimum solution in which all constraints are tight with fractional z variables.*

Proof. The second statement follows immediately by taking set S in the first statement to be the set of all fractional z variables. The first statement follows from the fact that we can always increase (decrease) each variable w in set S by $\varepsilon(1-w)$, for a very small $\varepsilon > 0$, and decrease (increase) each other variable u by εu (increase/decrease is depending on which option does not increase the objective function). It is easy to see in this way we can always get another feasible solution which makes all our current constraints in the laminar family tight and whose value is not larger than that of optimum. □

4 Prize-Collecting Survivable Steiner Network Design

In this section, we show how we can generalize our approach of iterative rounding to obtain a 3-approximation algorithm for the prize-collecting survivable Steiner network design problem. In this problem, we are given connectivity requirements r_{uv} for all pairs of vertices u and v and a non-increasing marginal penalty function $\pi_{uv}(.)$ for u and v. Our goal is to find a minimum way of constructing a graph in which we connect u and v with $r'_{uv} \leq r_{uv}$ edge-disjoint paths and paying the marginal penalty $\pi_{uv}(r'_{uv}+1) + \pi_{uv}(r'_{uv}+2) + \cdots + \pi_{uv}(r_{uv})$ for violating the connectivity between u and v to the amount of $r_{uv} - r'_{uv}$.

Let us first start with the following natural LP.

$$\text{OPT} = \quad \text{minimize} \sum_{e \in E} c_e x_e \quad + \sum_{i,j \in \mathcal{P}} \sum_{k=1}^{r_{ij}} \pi_{ij}(k) z_{ij}^k \tag{9}$$

$$\text{subject to} \quad \sum_{e \in \delta(S)} x_e + \sum_{k=1}^{r_{ij}} z_{ij}^k \geq r_{ij} \quad \forall S \subset V, (i,j) \in \mathcal{P}, S \odot (i,j) \tag{10}$$

$$x_e \geq 0 \quad \forall e \in E \tag{11}$$

$$z_{ij}^k \geq 0 \quad \forall (i,j) \in \mathcal{P}, 0 \leq k \leq r_{ij} \tag{12}$$

First, it is easy to see that since π_{uv}'s are non-increasing without loss of generality we can assume $0 < z_{uv}^k$ only if $z_{uv}^{k+1} = 1$ for $1 \leq k < r_{uv}$. Now the algorithm indeed is very similar to PCSF-ALG in Figure 1, except for an edge e with $x_e^* \geq \frac{1}{3}$, we do not contract that edge (indeed the contraction was only due to simplicity in PCSF-ALG). Instead we choose edge e to be in our solution and consider it like an edge of $x_e^* = 1$ value in the rest of the rounding. We repeat this process until we satisfy all the commodity pairs either by connecting or paying enough penalty. The argument follows almost the same as the argument for PCSF with the change of connectivity r_{ij} instead of 1 in our arguments in Theorem 3. Note that since $0 < z_{uv}^k$ only if $z_{uv}^{k+1} = 1$ for $1 \leq k < r_{uv}$, we can assume that each constraint is tight with only one variable z_{uv}^k, $1 \leq k \leq r_{uv}$ (all $z_{uv}^k = 1$ can be rounded to one and removed from further consideration in the LP without costing any extra penalties with respect to the optimum solution of the LP in Theorem 2). Thus as a result we have the following theorem.

Theorem 6. *There is an iterative rounding 3-approximation algorithm for the prize-collecting survivable Steiner network design problem.*

Finally, it is worth mentioning that by combining the technique of this paper in obtaining a one-third-integral variable and that of Lau et. al [18] (which essentially use the work of Jain for survivable network design as a block-box), it is not hard to get $(3, 3B_v + 3)$-approximation algorithm for the *prize-collecting survivable network design with bounded-degree constraints* B_v, where the cost of the returned solution is at most three times the cost of an optimum solution satisfying the degree bounds and the degree of each vertex is at most $3B_v + 3$.

5 Conclusions and Tight Example

In this paper, we presented a new approach of iterative rounding for prize-collecting problems which generalizes the use of iterative rounding when we do not have necessarily submodular functions. In addition, we used more linear dependence between constraints instead of just some simple additions with all coefficients one. The replacement of one of four sets instead of two sets in our laminar family is another extension to previous iterative rounding approaches (e.g. see [13]). In addition, next we show that indeed our approach of iterative rounding for getting a 3-approximation algorithm is tight even for prize-collecting Steiner forest, i.e., there is an instance with a basic feasible solution in which all x and z variables, except one zero z variable[3], are $\frac{1}{3}$.

Tight example: Consider a complete bipartite graph $K_{3,2} = (\{v_1, v_2, v_3\} \cup \{v_4, v_5\}, E)$ with (penalty) pairs $\mathcal{P} = \{(v_1, v_3), (v_2, v_3), (v_4, v_5)\}$. Assume all edges in E and penalties in $\mathcal{P}$ are ones. Consider a basic feasible solution in which all x and z variables are $\frac{1}{3}$, except $z_{4,5} = 0$. The cost of this fractional solution is $\frac{8}{3}$ which is less than the optimum integral solution 3 for this example. Also, it is easy to check that sets $\{v_1\}$ with (v_1, v_3), $\{v_2\}$ with (v_2, v_3), $\{v_3\}$ with (v_1, v_3), $\{v_3\}$ with (v_2, v_3), $\{v_4\}$ with (v_4, v_5), $\{v_5\}$ with (v_4, v_5), $\{v_1, v_4\}$ with (v_4, v_5), $\{v_2, v_5\}$ with (v_4, v_5) form a laminar family of tight constraints. These eight tight constraints in addition of tight constraint $z_{4,5} = 0$ form nine tight independent constraints of the aforementioned basic feasible solution. In this case, by fixing $z_{4,5} = 0$ and omitting variable $z_{4,5}$, we end up with exactly the same instance in which all variables are $\frac{1}{3}$. This shows that $\frac{1}{3}$ in our Theorem 5 is indeed tight.

Finally, we do believe that our iterative rounding approach might be applicable for other problems such as *multicommodity connected facility location (MCFL)* and *multicommodity rent-or-buy (MRoB)* (see e.g. [1,4,9,16,17]) to obtain simpler approximation algorithms with better factors than those currently exist.

Acknowledgement. The first author would like to thank Philip Klein and Mohammad-Hossein Bateni for several fruitful discussions and reading an early draft of this paper. Thanks especially goes to Howard Karloff whose program generated an example whose simplified version is the tight example in Section 5.

[3] Note that always there exists a z variable with an integral value when we solve the original LP according to Proposition 1.

References

1. Awerbuch, B., Azar, Y.: Buy-at-bulk network design. In: FOCS 1997, p. 542 (1997)
2. Balas, E.: The prize collecting traveling salesman problem. Networks 19, 621–636 (1989)
3. Bansal, N., Khandekar, R., Nagarajan, V.: Additive guarantees for degree bounded directed network design. In: STOC 2008, pp. 769–778 (2008)
4. Becchetti, L., Konemann, J., Leonardi, S., Pal, M.: Sharing the cost more efficiently: Improved approximation for multicommodity rent-or-buy. In: SODA 2005, pp. 375–384 (2005)
5. Bienstock, D., Goemans, M.X., Simchi-Levi, D., Williamson, D.: A note on the prize collecting traveling salesman problem. Math. Prog. 59, 413–420 (1993)
6. Chudak, F.A., Roughgarden, T., Williamson, D.P.: Approximate k-MSTs and k-Steiner trees via the primal-dual method and lagrangean relaxation. In: Aardal, K., Gerards, B. (eds.) IPCO 2001. LNCS, vol. 2081, pp. 60–70. Springer, Heidelberg (2001)
7. Goemans, M.X., Williamson, D.P.: A general approximation technique for constrained forest problems. SIAM J. Comput. 24, 296–317 (1995)
8. Gupta, A., Könemann, J., Leonardi, S., Ravi, R., Schäfer, G.: An efficient cost-sharing mechanism for the prize-collecting steiner forest problem. In: SODA 2007, pp. 1153–1162 (2007)
9. Gupta, A., Kumar, A., Pal, M., Roughgarden, T.: Approximation via cost-sharing: a simple approximation algorithm for the multicommodity rent-or-buy problem. In: FOCS 2003, p. 606 (2003)
10. Gutner, S.: Elementary approximation algorithms for prize collecting Steiner tree problems. Inf. Process. Lett. 107, 39–44 (2008)
11. M. Hajiaghayi, Designing minimum total cost networks using iterative rounding approximation methods. Pending patent with USPTO of application number 12/315,657 (July 2008)
12. Hajiaghayi, M.T., Jain, K.: The prize-collecting generalized Steiner tree problem via a new approach of primal-dual schema. In: SODA 2006, pp. 631–640 (2006)
13. Jain, K.: A factor 2 approximation algorithm for the generalized Steiner network problem. Combinatorica 21, 39–60 (2001)
14. Jain, K., Vazirani, V.V.: Approximation algorithms for metric facility location and k-median problems using the primal-dual schema and Lagrangian relaxation. J. ACM 48, 274–296 (2001)
15. Johnson, D.S., Minkoff, M., Phillips, S.: The prize collecting Steiner tree problem: theory and practice. In: SODA 2000, pp. 760–769 (2000)
16. Karger, D.R., Minkoff, M.: Building steiner trees with incomplete global knowledge. In: FOCS 2001, p. 613 (2001)
17. Kumar, A., Gupta, A., Roughgarden, T.: A constant-factor approximation algorithm for the multicommodity. In: FOCS 2002, p. 333 (2002)
18. Lau, L.C., Naor, J.S., Salavatipour, M.R., Singh, M.: Survivable network design with degree or order constraints. In: STOC 2007, pp. 651–660 (2007)
19. Lau, L.C., Singh, M.: Additive approximation for bounded degree survivable network design. In: STOC 2008, pp. 759–768 (2008)
20. Salman, F.S., Cheriyan, J., Ravi, R., Subramanian, S.: Approximating the single-sink link-installation problem in network design. SIAM J. Optim. 11, 595–610 (2000)
21. Sharma, Y., Swamy, C., Williamson, D.P.: Approximation algorithms for prize collecting forest problems with submodular penalty functions. In: SODA 2007, pp. 1275–1284 (2007)
22. Konemann, J., Grandoni, F., Rothvoss, T., Qian, J., Schaefer, G., Swamy, C., Williamson, D.P.: An iterated rounding 3-approximation algorithm for prize-collecting Steiner forest (2009) (unpublished manuscript)
23. Singh, M., Lau, L.C.: Approximating minimum bounded degree spanning trees to within one of optimal. In: STOC 2007, pp. 661–670 (2007)