Approximation Algorithms Workshop June 13-17, 2011, Princeton

Lecture Steiner Tree Approximation- June 15

Fabrizio Grandoni *Scribe: Fabrizio Grandoni*

1 Overview

In this lecture we will summarize our recent work on Steiner tree approximation together with J. Byrka, T. Rothvoß and L. Sanità. Part of the results appeared in "An Improved LP-based Approximation for Steiner Tree" (STOC'10), and some other results are contained in the journal version of the paper, currently under revision. We will not be very formal, and we will abuse notation. We apologize for possible typos and small inconsistencies, and we refer the reader to the mentioned papers for a more careful and detailed description of the results.

2 The Problem and Related Work

In the Steiner tree problem we are given an undirected graph $G = (V, E)$, with edge costs $c : E \rightarrow \mathbb{R}_{\geq 0}$, and a set R of terminal nodes. The goal is to compute the cheapest (i.e., minimum cost $c(S) := \sum_{e \in S} c(e)$) tree S which spans R. Note that S might contain other nodes besides R, which are called Steiner nodes. A *terminal spanning tree* is a Steiner tree without Steiner nodes.

Steiner tree is NP-hard even when edge costs are either 1 or 2 [Bern,Plassmann'89] and there is no (polynomial-time) $96/95-\varepsilon$ approximation algorithm for it unless $P = NP$ [Chlebik,Chlebikova'02]. On the positive side, a very simple algorithm, the so-called *minimum spanning tree heuristic*, provides a 2-approximation. Essentially, one considers the metric closure (G', w') of the graph, and computes a minimum cost terminal spanning tree in G'.

All the following improvements, culminating with the 1.55 approximation in [Robins,Zelikovsky'00], are based on a local search strategy: one starts with the 2-approximate solution obtained with the minimum spanning tree heuristic, and performs local improvements as long as possible.

All the approximation factors better than 2 are w.r.t. the optimal integral solution. A relevant open problem is to find an LP-relaxation with integrality gap smaller than 2.

3 Our Results

We developed a deterministic $\ln(4) + \varepsilon < 1.39$ approximation algorithm. In these notes we will focus on a simpler expected $1.5 + \varepsilon$ approximation. Based on the same approach, we also prove that a natural LP-relaxation has integrality gap at most $1 + \ln(3)/2 < 1.55$. In these notes we will focus on a simpler $1 + \ln(2) < 1.7$ bound.

Our results are based on two main ingredients. First, a directed-component cut relaxation, which is inspired by the classical bidirected cut relaxation and by the notion of k-components. Second, we exploit a novel algorithmic framework, *iterative randomized rounding*, which combines some aspects of the randomized rounding and iterative rounding techniques.

4 Directed-Component Cut Relaxation

As mentioned before, our directed-component cut relaxation is inspired by the classical bidirected cut relaxation (BCR), which is defined as follows. Let us choose an arbitrary terminal r as a root. We replace each undirected edge $e = \{u, v\}$ with two oppositely directed edges (u, v) and (v, u), which inherit the cost of e. Let E' be the new set of edges. The relaxation is as follows.

$$\begin{aligned} \min \sum_{e \in E} c(e) z_e & \qquad \text{(BCR)} \\ \sum_{e \in \delta^+(U)} z_e \geq 1 & \qquad \forall U \subseteq V \setminus \{r\} : U \cap R \neq \emptyset \\ z_e \geq 0 & \qquad \forall e \in E'. \end{aligned}$$

Here $\delta^+(U)$ is the set of directed edges (u, v) with $u \in U$ and $v \notin U$ (the edges *crossing* U). It is easy to see that the optimum integral Steiner tree defines a feasible solution. In a fractional solution, we can interpret z_e as the capacity reserved on edge e. Then a feasible fractional solution supports one unit of flow (non-simultaneously) from each terminal to the root.

In [Edmonds'67] it is show that BCR is integral in the spanning tree case, i.e. when $R = V$. Observe that the integrality gap of the undirected version of BCR is 2 even in this special case. This suggests the idea that using directed edges can help.

The second ingredient is the notion of k-components. A component of a Steiner tree is a maximal subtree whose terminals coincide with its leaves. A k-component is a component with at most k terminals, and a Steiner tree made of k-components is k-restricted. The following theorem in [Borchers,Du'97] shows that, for k large enough, a minimum k-restricted Steiner trees is a very good approximation of the minimum (unrestricted) Steiner tree.

Theorem 1. *If opt_k and opt are the costs of an optimal k-restricted Steiner tree and an optimal Steiner tree, respectively, then $opt_k \leq \left(1 + \frac{1}{\lfloor \log_2 k \rfloor}\right) opt$.*

Let us choose a root terminal r, and direct all the the edges of the optimum Steiner tree S towards r. Each component is turned into a *directed component*, with a unique sink terminal. We call sources the other terminals of the component. The following directed-component cut relaxation (DCR) is inspired by the above construction:

$$\begin{aligned} \min \sum_{C \in \mathcal{C}} c(C) x_C & \qquad \text{(DCR)} \\ \sum_{C \in \delta^+_{\mathcal{C}}(U)} x_C \geq 1 & \qquad \forall U \subseteq R - r, U \neq \emptyset \\ x_C \geq 0 & \qquad \forall C \in \mathcal{C} \end{aligned}$$

Here $\mathcal{C}$ denotes the set of possible directed components, and $\delta^+_{\mathcal{C}}(U)$ is the set of directed components with at least one source in the cut U and its sink outside U. We remark that DCR is strictly stronger than BCR.

We will need the following two lemmas.

Lemma 2. *A* $(1+\varepsilon)$ *approximation of the optimal fractional solution* opt^f *to DCR can be computed in polynomial time*

The idea in the lemma above is to focus on directed components on at most k-terminals. Then the separation problem can be solved in polynomial time.

Lemma 3. *The cost of a minimum terminal spanning tree is at most* $2\,opt^f$*.*

5 The Algorithm

We exploit DCR within a (seemingly) novel *iterative randomized rounding* framework. The basic idea is to solve an LP-relaxation for the considered problem, and round one of the variables which is chosen with probability proportional to its fractional value. The problem is then simplified, and the process is iterated until a proper halting condition is satisfied. This approach combines features of *randomized rounding* (where variables are rounded randomly and, typically, simultaneously) and *iterative rounding* (where variables are rounded deterministically and, typically, one at a time).

For the considered problem, at each iteration t, we compute a $1+\varepsilon$ approximate solution x^t for DCR. Then we sample a component $C = C^t$ with probability $p^t_C := x^t_C / \sum_{D \in \mathcal{C}} x^t_D$. We contract C^t and update DCR consequently. We iterate the process until all the terminals are contracted into the root. At that point, we output the sampled components.

By adding a dummy component in the root, we can assume w.l.o.g. that $M := \sum_{D \in \mathcal{C}} x^t_D$ is fixed for all iterations t. We can also ideally let the algorithm run foreover (at some point it will always sample the dummy component, at zero extra cost).

6 The Bridge Lemma

Let us focus on a lemma (*Bridge Lemma*), which plays a crucial role in our analysis.

Given a Steiner tree S and a subset $R' \subseteq R$ of terminals, the *bridges* $br_{S,c}(R')$ of S w.r.t. R' (and to edge costs c) are the edges of S which do not belong to the minimum spanning tree of the nodes of S (including Steiner nodes) after the contraction of R'. Alternatively, one might connect terminals R' via dummy edges of cost zero, and compute a minimum spanning tree S' of the resulting graph: the edges in S but not in S' define $br_{S,c}(R')$. We remark that the most expensive edge on a path between two terminals in R' is a bridge. We will abuse notation by omitting c when it is clear from the context, and letting $br_S(R') := c(br_S(R'))$ and $br_S(C) := br_S(R \cap C)$ for a component C. We we call *candidate bridge set* a minimal set B of edges such that S remains connected after the contraction of R' and the removal of B. In particular, $br_{S,c}(R')$ is the most expensive candidate bridge set.

We will exploit the fact that the bridge set of a Steiner tree with respect to all its terminals is quite expensive.

Lemma 4. *For any Steiner tree* S *on* R*,* $br_S(R) \geq \frac{1}{2}c(S)$*.*

The idea behind the lemma above is to turn S into a binary tree, rooted at some node s, whose leaves coincide with its terminals. Then we take the most expensive of the two edges from each internal node to its children. This defines a candidate bridge set.

We are now ready to state the Bridge Lemma.

Lemma 5 (Bridge Lemma). *For any terminal spanning tree T and any feasible fractional solution x to DCR, $\sum_{C\in\mathcal{C}} x_C \cdot br_T(C) \geq c(T)$.*

Intuitively, this lemma relates the cost of a terminal spanning tree to the solutions of DCR via the notion of bridges. Let us remark that our lemma applies to terminal spanning trees, not to arbitrary Steiner trees.

In the proof of this lemma we crucially exploit the following construction. Consider any directed component C. We construct a directed terminal spanning tree Y_C on $R' := C \cap R$, with edge weights and capacities, as follows. Let us remove $br_C(T)$ from T. We obtain a forest where each tree contains exactly one terminal in R'. We put back one bridge $b \in br_C(T)$ at a time. This way, we connect two trees T' and T'' of the mentioned forest. We add to Y_C one edge between the two terminals $r' \in R \cap T'$ and $r'' \in R \cap T''$, and we set $w(r'r'') := c(b)$. Finally, we direct all the edges of Y_C towards the sink of C, and assign capacity x_C to the edges of Y_C. Let us remark that Y_C supports the same flow to the root as C, with respect to the terminals.

Now, we replace each component C in the fractional solution with the corresponding Y_C. Note that parallel edges might appear: in that case they have the same w-cost; we merge them and sum together the corresponding capacities. Altogether we obtain a directed weighted graph G', containing (all the) terminals only. Furthermore, edge capacities support one unit of flow from each terminal to the root. Hence, edge capacities define a feasible fractional directed terminal spanning tree. In other terms, they give a feasible fractional solution to BCR. The mentioned theorem by Edmonds then implies that there exists an integral directed (terminal) spanning tree F in G' of w-cost not larger than the w-cost of the mentioned fractional terminal spanning tree (since BCR is integral when all nodes are terminals). Finally, we observe that the w-cost of F is lower bounded by the c-cost of T via the cycle rule. Intuitively, if we consider any edge $f = (u, v)$ of F, and the corresponding path in T (considered as an undirected tree), then in the resulting cycle the larger edge cost is $w(f)$ by construction.

7 A First Bound

Let us show that our algorithm computes a solution of cost at most $(1 + \ln 2 + \varepsilon)\, opt^f$. This also implies that the integrality gap of DCR is at most $1 + \ln 2 < 1.7$.

Let apx be the cost of the approximate solution. Then $E[apx] = \sum_{t\geq 1} E[c(C^t)]$. In turn $E[c(C^t)] = E[\sum_C \frac{x^t_C}{M} c(C)] = \frac{1+\varepsilon}{M} E[opt^{f,t}]$. Here $opt^{f,t}$ denotes the cost of the optimum fractional solution at the beginning of iteration t. It remains to upper bound $E[opt^{f,t}]$. For small enough values of t (according to a proper threshold value), we use the trivial upper bound opt^f (the cost of the optimal fractional solution cannot increase due to contractions). For large enough values of t, we use the upper bound $opt^{f,t} \leq c(T^t)$, where T^t is the minimum terminal spanning tree at the beginning of iteration t (which is a feasible integral Steiner tree). Altogether

$$\begin{aligned}
E[apx] &= \sum_{t\geq 1} E[c(C^t)] \leq \sum_{t\geq 1} E[\sum_{C} \frac{x_C^t}{M} c(C)] \leq \frac{1+\varepsilon}{M} \sum_{t\geq 1} E[opt^{f,t}] \\
&\leq \frac{1+\varepsilon}{M} \sum_{t=1}^{M\ln 2} opt^f + \frac{1+\varepsilon}{M} \sum_{t>M\ln 2} E[c(T^t)].
\end{aligned}$$

Exploiting the Bridge Lemma, it is not hard to show that the cost of T^t decreases by a factor $(1-1/M)$ at each iteration in expectation. In fact, given T^t, a feasible terminal spanning tree at iteration $t+1$ is obtained by removing from T^t the bridges of T^t w.r.t. the t-th sampled component C^t:

$$E[c(T^{t+1})] \leq c(T^t) - E[br_{T^t}(C^t)] = c(T^t) - \sum_C \frac{x_C^t}{M} br_{T^t}(C) \overset{\text{Bridge Lem}}{\leq} c(T^t) - \frac{1}{M} c(T^t).$$

As a consequence $E[c(T^t)] \leq (1-\frac{1}{M})^{t-1} c(T^1) \leq (1-\frac{1}{M})^{t-1} 2\, opt^f$. We can conclude that

$$\begin{aligned}
E[apx] &\leq \frac{1+\varepsilon}{M} \sum_{t=1}^{M\ln 2} opt^f + \frac{1+\varepsilon}{M} \sum_{t>M\ln 2} E[c(T^t)] \\
&\leq opt^f (1+\varepsilon)\ln 2 + 2\, opt^f (1+\varepsilon) \sum_{t>M\ln 2} \frac{1}{M}\left(1-\frac{1}{M}\right)^{t-1} \\
&\leq (1+\varepsilon)(\ln 2 + 2e^{-\ln 2}) \cdot opt^f.
\end{aligned}$$

8 A Better Bound

We next show a refined $1.5+\varepsilon$ approximation bound for the same algorithm. However, the bound this time is w.r.t. the optimal integral solution *opt*, and does not imply the same bound on the integrality gap of DCR.

The analysis starts in the same way as before. However, this time we replace the trivial upper bound $opt^{f,t}$ for small values of t with $c(S^t)$, where S^t is the optimum Steiner tree at the beginning of iteration t. We next show that the expected cost of S^t decreases in expectation by at least a factor $(1-\frac{1}{2M})$ at each iteration. The idea is to compute a weighted terminal spanning tree (Y^t, w) with respect to S^t and all its terminals R^t (the ones surviving till iteration t) with the same construction as in the bridge lemma. In particular, we will remove the bridges of S^t with respect to all its terminals etc. Recall that in this construction each edge e of Y^t is in one to one correspondence with some edge $b(e) \in br_{S^t}(R^t)$. We also recall that the c-cost of such bridges, and hence the w-cost of Y^t, is at least $c(S^t)/2$. A feasible Steiner tree at iteration $t+1$ is obtained by removing from S^t the edges $b(e)$ such that e is a bridge of Y^t with respect to the t-th sampled component C^t (and edge costs w). In other words, a feasible Steiner tree at iteration $t+1$ is given by $S' := S^t - \{b(e) \in S^t : e \in br_{Y^t,w}(C^t)\}$. Then

$$
\begin{aligned}
E[c(S^{t+1})] \leq E[c(S')] = c(S^t) - E[c(\{b(e) \in S^t \,|\, e \in br_{Y^t,w}(C^t)\})] = c(S^t) - E[br_{Y^t,w}(C^t)] \\
= c(S^t) - \frac{1}{M}\sum_C x_C^t br_{Y^t,w}(C) \overset{\text{Bridge Lem}}{\leq} c(S^t) - \frac{1}{M}w(Y^t) \\
= c(S^t) - \frac{1}{M} br_{S^t,c}(R^t) \leq c(S^t) - \frac{1}{M}\frac{c(S^t)}{2}.
\end{aligned}
$$

Hence $E[c(S^t)] \leq (1 - \frac{1}{2M})^{t-1}c(S^1) = (1 - \frac{1}{2M})^{t-1}opt$. Altogether

$$
\begin{aligned}
E[apx] &= \sum_{t\geq 1} E[c(C^t)] \leq \sum_{t \geq 1} E[\sum_j \frac{x_j^t}{M} c(C_j^t)] \leq \frac{1+\varepsilon}{M}\sum_{t\geq 1} E[opt^{f,t}] \\
&\leq \frac{1+\varepsilon}{M}\sum_{t=1}^{M\ln 4} E[c(S^t)] + \frac{1+\varepsilon}{M}\sum_{t > M\ln 4} E[c(T^t)] \\
&\leq (\frac{1+\varepsilon}{M} opt) \cdot (\sum_{t=1}^{M\ln 4}(1 - \frac{1}{2M})^{t-1} + \sum_{t > M\ln 4} 2(1 - \frac{1}{M})^{t-1}) \\
&\leq (1+\varepsilon)(2 - 2e^{-\ln(4)/2} + 2e^{-\ln(4)}) \cdot opt.
\end{aligned}
$$

9 An Even Better Bound

The claimed $\ln 4 + \varepsilon$ approximation is based on the following idea. We want to obtain a better bound on $c(S^t)$, and use it to upper bound $opt^{f,t}$ at any iteration t. Let S be the initial optimum Steiner tree. First of all, we sample a (proper) random candidate bridge set B of S w.r.t. all the terminals R. Based on B, we create a terminal spanning tree W (*witness tree*) with the same construction as in the Bridge Lemma. We associate to each $e \in S$ a subset $W(e)$ of edges of W (*witness set*) as follows: edge $\{u, v\} \in W$ belongs to $W(e)$ iff e lies along the path from u to v in S. The random choice of B is such that $|W(e)|$ is equal to 1, 2,... with probability $\frac{1}{2}$, $\frac{1}{4}$, ... Whenever we sample a component C^t, we will remove a (proper) random candidate bridge set from W w.r.t. the terminals of C^t. The latter set is chosen according to a probability distribution which guarantees that, at each iteration, each edge of W which is not sampled yet, is sampled in the considered iteration with probability at least $1/M$. Intuitively, we wish to *uniformly* delete edges from W. The existence of such probability distribution is guaranteed by the Bridge Lemma in combination with Farkas' Lemma.

When all the edges of $W(e)$ are removed, we delete e from S. At any time, the undeleted edges in S plus the sampled components define a feasible Steiner tree, whose cost upper bounds $c(S^t)$. The cost of the approximate solution is at most $\frac{1+\varepsilon}{M}\sum_{t\geq 1} E[c(S^t)]$. This can be rewritten as $\frac{1+\varepsilon}{M} \cdot opt$ times the expected number of iterations until a given edge of S is deleted. It is not hard to show that this number is at most $M \cdot \ln 4$.

10 Derandomization

It is possible to derandomize our algorithm (and achieve the same approximation factor). The idea is to run a constant number $1/\varepsilon^2$ of phases. In each phase we sample several components from the same LP solution, and contract them. This sampling is performed in a dependent way, so that $O(\log n)$ random bits are sufficient (and hence the algorithm is easy to derandomize). At the same time we guarantee that, in each phase, each component is sampled with probability at most $O(\varepsilon)$ times its fractional value, and each edge of the witness tree W is deleted with probability at least $\Omega(\varepsilon)$. At the end of the last phase, one can simply output a minimum terminal spanning on the residual instance (whose expected cost is very small) plus the sampled components.

11 Open Problems

A first interesting question is whether the $\ln 4 + \varepsilon$ (or even $1.5 + \varepsilon$) approximation factor can be proved w.r.t. the optimal fractional solution opt^f.

A second natural goal is to find other applications of the iterative randomized rounding framework. Natural (still challenging) candidates are problems related to Steiner tree, such as prize-collecting Steiner tree, k-MST and Single-Sink Rent-or-Buy.