Approximations for Steiner Tree Problem

Chihao Zhang

BASICS, Shanghia Jiao Tong University

January 7, 2013

Chihao Zhang (BASICS, Shanghia Jiao Tong Approximations for Steiner Tree Problem

January 7, 2013 1 / 20

Steiner Tree: Problem Statement

Input:	Given an undirected graph $G = (V, E)$, an edge
	cost $c_e \ge 0$ for each $e \in E$. V is partitioned into
	two sets, terminals and Steiner vertices.
Problem:	Find a minimum cost tree in G that contains all the
	terminals and any subset of the Steiner vertices.

• If G is a complete graph and the costs satisfy *the triangle inequality*, i.e.,

$$cost(u, v) \leq cost(u, w) + cost(w, v),$$

then we call it metric Steiner tree problem.

Theorem

There is an 2-approximation algorithm for metric Steiner tree problem.

MST Based Algorithm (cont'd)

- The algorithm simply returns the minimum spanning tree on terminal vertices.
- To see it is a 2-approximation algorithm, oberseve that we can obtain a MST of terminals from optimal solution of Steiner tree problem by doubling its cost using triangular inequality.

Theorem

There is an approximation factor preserving reduction from the Steiner tree problem to the metric Steiner tree problem.

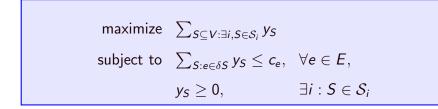
Input:	Given an undirected graph $G = (V, E)$, nonnega- tive costs $c_e \ge 0$ for all edges $e \in E$ and k pairs of vertices $s_i, t_i \in V$
Problem:	Find a minimum cost subset of edges $F \subseteq E$ such that every s_i - t_i pair is connected in the set of selected edges.

Integer ProgramLinear Program Relaxation

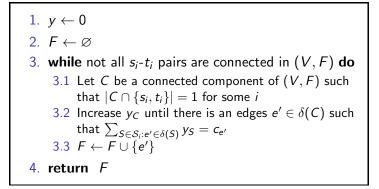
$$\begin{array}{ll} \text{minimize} & \sum_{e \in E} c_e x_e \\ \text{subject to} & \sum_{e \in \delta(S)} x_e \geq 1, \qquad \forall S \subseteq V : S \in \mathcal{S}_i \text{ for some } i, \\ & x_e \in \{0, 1\} x_e \geq 0, \quad e \in E. \end{array}$$

where $S_i := \{S \subseteq V \mid |S \cap \{s_i, t_i\}| = 1\}$ and $\delta(S) := \{e = \{u, v\} \in E \mid u \in S, v \notin S\}$

The Dual Program



Standard Primal-Dual Schema



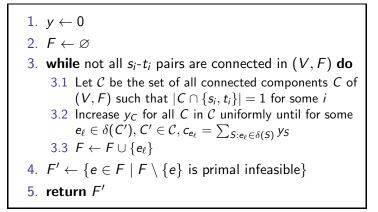
Standard Primal-Dual Schema (cont'd)

Using the standard primal-dual analysis, we have

$$\sum_{e\in F} c_e = \sum_{e\in F} \sum_{S:e\in \delta(S)} y_S = \sum_S |\delta(S) \cap F| y_S.$$

- However, $|\delta(S) \cap F|$ can be as large as k!
- We will use an average case analysis instead of worst case analysis.

Primal-Dual Schema with Synchronization



Analysis

Using standard primal-dual analysis, we have

$$\sum_{e \in F'} c_e = \sum_{e \in F'} \sum_{S: e \in \delta(S)} y_S = \sum_S |F' \cap \delta(S)| y_S.$$

We will show that

$$\sum_{S} |F' \cap \delta(S)| y_{S} \leq 2 \sum_{S} y_{S}$$

This follows from the following lemma:

Lemma

For any ${\mathcal C}$ in any iteration of the algorithm,

$$\sum_{C\in\mathcal{C}}|\delta(C)\cap F'|\leq 2|\mathcal{C}|$$

Proof.

The high-level idea is the following: If in each iteration, we contract every components into a single vertex, then F' is a forest in the contracted graph. Then we apply the fact that the average degree of vertices in a forest is at most 2.

Prize-collecting Steiner Tree Problem

Input:	An undirected graph $G = (V, E)$, an edge cost
	$c_e \geq 0$ for each $e \in E$, a selected root vertex
	$r \in V$, and a penalty $\pi_i \ge 0$ for each $i \in V$.
Problem:	Find a tree T that contains the root vertex r so as
	to minimize the cost of the edges in the tree plus
	the penalties of all vertices not in the tree

• A generalization of Steiner Tree Problem, where $\pi_i = \infty$ if *i* is terminal and $\pi_i = 0$ otherwise.

Integer ProgramLinear Program Relaxation

$$\begin{array}{ll} \text{minimize} & \sum_{e \in E} c_e x_e + \sum_{i \in V} \pi_i (1 - y_i) \\ \text{subject to} & \sum_{e \in \delta S} x_e \geq y_i, & \forall S \subseteq V - r, S \neq \emptyset, \forall i \in S, \\ & y_r = 1, & \\ & y_i \in \{0, 1\} y_i \geq 0, & \forall i \in V, \\ & x_e \in \{0, 1\} x_e \geq 0, & \forall e \in E. \end{array}$$

Definition (Separation Oracle)

A separation oracle takes as input a solution x and either verifies that x is a feasible solution or produces a constraint that is violated by x.

• Provided a polynomial-time separation oracle, the ellipsoid method can solve a linear program in polynomial time.

Polynomial-time Separation Oracle

The following algorithm is a polynomial-time separation oracle for our linear program relaxation of prize-collecting Steiner tree problem:

```
Given a solution (x, y),
1. Construct a network folow problem on G in which the capacity of each edge e is x<sub>e</sub>.
2. For each i ∈ V
2.1 If the maximum flow from from i to r is less than y<sub>i</sub>
2.1.1 Return the constraint (S, i) where S is the minimum cut from i to r
3. Return "(x, y) is a feasible solution".
```

Deterministic Rounding

- 1. Let $\alpha \in [0,1)$ be a parameter to be fixed.
- 2. $U := \{i \in V \mid y_i \ge \alpha\}.$
- 3. Find a Steiner tree T on G with terminals U.
- 4. Return T.

Lemma

The tree T returned by the primal-dual algorithm for Steiner tree has cost

$$\sum_{e\in T} c_e \leq \frac{2}{\alpha} \sum_{e\in E} c_e x_e^*.$$

Proof.

Oberseve that if $\{x_e^*, y_i^*\}$ is a solution of LP for prize-collecting Steiner tree problem, then $\{\frac{x_e^*}{\alpha}\}$ is a solution of LP for Steiner tree problem.

Deterministic Rounding (cont'd)

Lemma

$$\sum_{i\in V\setminus V(T)}\pi_i\leq \frac{1}{1-\alpha}\sum_{i\in V}\pi_i(1-y_i^*)$$

Theorem

The cost of the solution produced by the deterministic rounding algorithm is

$$\sum_{e\in T} c_e + \sum_{i\in V\setminus V(T)} \pi_i \leq \frac{2}{\alpha} \sum_{e\in E} c_e x_e^* + \frac{1}{1-\alpha} \sum_{i\in V} \pi_i (1-y_i^*).$$

Taking $\alpha = \frac{2}{3}$, the algorithm is a 3-approximation algorithm for the prize-collecting Steiner tree problem.

Randomized Rounding

• Let 0 $<\gamma \leq$ 1 be a constant to be fixed. Choose α uniformly from $[\gamma,1].$

Lemma

$$E\left[\sum_{e \in T} c_e\right] \leq \left(\frac{2}{1-\gamma} \ln \frac{1}{\gamma}\right) \sum_{e \in E} c_e x_e^*$$

Lemma

$$E\left[\sum_{i\in V\setminus V(\mathcal{T})}\pi_i
ight]\leq rac{1}{1-\gamma}\sum_{i\in V}\pi_i(1-y_i^*)$$

Randomized Rounding (cont'd)

Theorem

The expected cost of the solution produced by the randomized algorithms is

$$E\left[\sum_{e\in T} c_e + \sum_{i\in V\setminus V(T)} \pi_i\right] \leq \left(\frac{2}{1-\gamma}\ln\frac{1}{\gamma}\right)\sum_{e\in E} c_e x_e^* + \frac{1}{1-\gamma}\sum_{i\in V} \pi_i(1-y_i^*).$$

- Choosing $\gamma = e^{-1/2}$ gives a $(1 e^{-1/2})^{-1}$ -approximation algorithm for the prize-collecting Steiner tree problem, where $(1 e^{-1/2})^{-1} \approx 2.54$.
- This algorithm can be easily derandomized.