Prize Collecting Steiner Tree

20 March 2018 14:17

The prize collecting Steiner Tree problem is defined as follows:

I/P : Undirected graph $G=(V,E)$

$C: E \to \mathbb{R}_{\geq 0}$ [cost of edges]

$\pi : V \to \mathbb{R}_{\geq 0}$ [penalty of a vertex]

$r \in V$: root

O/P : Steiner Tree on a subset $(T \subseteq V) \cup \{r\}$.

Obj : Minimize $c(T) + \pi(V - T \cup \{r\})$.

We shall prove a randomized 2.54-approximation for PCST.

$$LP(PCST): \quad \min \underbrace{\sum_{e \in E} c_e \cdot x_e}_{\text{Tree Cost}} + \underbrace{\sum_{v \in V - \{r\}} \pi_v (1 - y_v)}_{\text{Penalty}}.$$

s.t. :

$$\boxed{\sum_{e \in \delta(S)} x_e \geq y_v} \quad \forall v, S : S \text{ is a } \{v,r\}\text{-cut in } G \quad \text{— (1)}$$

Same as the standard cut constraints for St. Tree, when $y_v = 1$

$$x \geq 0 \quad \text{— (2)}$$

The above LP can be solved using Ellipsoid since the separation oracle consists of running $O(|V|)$ min-cut algorithms.

Assume we are given the optimal solution: x^*, y^*.

Pick a value α uniformly at random in the interval $[\gamma, 1]$ (we will decide γ later).

Now form set $V(T) = \{v : y_v > \alpha\}$. Build a 2-approximate Steiner Tree T on $V(T) \cup \{r\}$ using primal-dual.
Pay penalty for $V - V(T)$.

Lemma 1. $C(T) \leq \frac{2}{\alpha} \sum_e c_e \cdot x_e^*$

Proof: Consider the St. Tree LP on the set $V(T) \cup \{r\}$. The constraints are

$$\sum_{e \in \delta(S)} x_e' \geq 1 \quad \forall S : S \text{ is cut between some } v \in V(T) \text{ and } r$$

Now, consider the solution x_e^* of PCST and scale up by $1/\alpha$. : $\tilde{x}_e = \frac{x_e^*}{\alpha}$.

Then, for any $v \in V(T)$ and any cut S,

$$\sum_{e \in \delta(S)} \tilde{x}_e = \underbrace{\frac{1}{\alpha} \sum_{e \in \delta(S)} x_e^* \geq \frac{y_v^*}{\alpha}}_{\text{Constraint (1)}} \geq 1 \left[\because y_v^* \geq \alpha \text{ if } v \in V(T)\right]$$

The above implies that $\tilde{x}_e$ is a feasible solution to x_e'.

$$\Rightarrow \quad C(T) \leq 2 \sum_e x_e' \cdot c_e \leq 2 \sum \tilde{x}_e \cdot c_e = \frac{2}{\alpha} \sum x_e^* \cdot c_e . \quad \blacksquare$$

Lemma 2. $\pi(V - V(T)) \leq \frac{1}{1-\alpha} \sum_v \pi_v (1 - y_v^*)$

Proof: $y_v^* \leq \alpha$ for $v \in V - V(T)$

$\Rightarrow 1 - y_v^* \geq 1 - \alpha$ "

$\Rightarrow \pi_v \leq \frac{\pi_v}{1-\alpha} (1 - y_v^*)$ $\blacksquare$

However, since α is a continuous r.v. in $[\gamma, 1]$, we need to find the expected costs.

$$E[C(T)] = E\left[\frac{2}{\alpha}\right] \cdot \sum_e x_e^* \cdot c_e$$

[Note that T is a random tree and $C(T)$ is an r.v.]

The probability density function for α in the interval $[\gamma, 1]$, is $\frac{1}{1-\gamma}$. [fact]. Hence

$$\mathbb{E}[2/\alpha] = \frac{1}{1-\gamma}\int_{\gamma}^{1} 2/\alpha \cdot d\alpha = \frac{2}{1-\gamma}[\ln\alpha]_{\gamma}^{1}$$

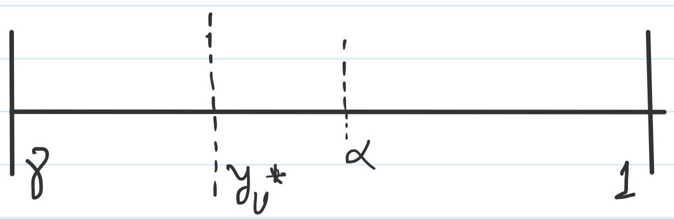
$$= \frac{2}{1-\gamma}\ln(1/\gamma) \quad \text{— } A$$

$$\mathbb{E}[\pi(V - V(T))] = \pi_v \cdot \Pr[v \notin V(T)]$$

We prove that $\Pr[v \notin V(T)] \leq \frac{1-y_v^*}{1-\gamma}$

Case I: if $y_v^* < \gamma$: $\Pr[v \notin V(T)] = 1 < \frac{1-y_v^*}{1-\gamma}$ $[\because y_v^* < \gamma]$

Case II: $\gamma \leq y_v^*$:



The event that $v \in V - V(T)$ is only possible if α is picked from the interval $(y_v^*, 1]$. The probability of this is $\leq \frac{1-y_v^*}{1-\gamma}$. $\boxed{\Xi}$

Hence $\mathbb{E}[\pi(V - V(T))] \leq \frac{1}{1-\gamma}\sum \pi_v \cdot (1-y_v^*)$

The total cost of the solution is

$$\frac{2}{1-\gamma}\ln(1/\gamma)\cdot\sum_e x_e^* c_e + \frac{1}{1-\gamma}\sum_v \pi_v(1-y_v^*)$$

The factor is optimized at $\gamma = e^{-1/2}$ which gives a value of 2.54

The factor is optimized at $\gamma = e^{1/2}$ which gives a value of 2.54.

<u>Derandomization</u>: The above algorithm can be derandomized easily as follows. Note that it is enough to consider α such that $\alpha = y_v^*$ for some $v \in V$. We run for all such α's (polynomially many) and output the cheapest solution.

Since expectation is bounded as above, the cheapest is bounded as well.

Note that here randomization helps in the analysis. We do not know any other way to analyze the deterministic variant.