

Lecture 25

Lecturer: Debmalya Panigrahi

Scribe: Ios Kotsogiannis

1 Overview

Last week we saw an application of semi-definite programs, as we used this technique to solve the max cut problem. In this lecture we will have an overview of the Group Steiner Tree problem when the input is a tree, and we will see how to use dependent rounding in order to reach our results.

2 Group Steiner Tree on Trees

Let $G = (V, E)$ a rooted tree with root $r \in V$, with leaves $L \subseteq V$, and of height H . Also define terminal groups R_1, \dots, R_k where $\forall i \in [k] : R_i \subseteq L$. Lastly, let a weight function on the edges: $c : E \rightarrow \mathbb{R}_+$. The goal is to find the minimum cost subtree of G that connects at least one vertex from each R_i .

We can formulate the group steiner tree problem in terms of a linear program (1):

$$\begin{aligned} \min \quad & \sum_{e \in E} c_e x_e \\ \text{s.t.} \quad & \sum_{e \in (S, \bar{S})} x_e \geq 1 \quad \forall S \text{ s.t. } \exists i : R_i \subseteq S, r \notin S \\ & x_e \geq 0 \end{aligned}$$

Despite the problem having exponentially many constraints, we will use the ellipsoid method to solve it in polynomial time. The ellipsoid method uses a separation oracle to see if any constraints were violated, and if so, how to satisfy them. More specifically, the oracle picks one group R_i , connects all vertices of R_i to some sink s and then pushes flow of 1 from r (the root of G) to s . The oracle needs to run k times (the total number of groups).

Hence, the crux of this problem is the rounding procedure, and in the following sections we will deal with that.

The work we present in this lecture is from [GKR00]. More work on the group steiner tree problem can be found in [DHK09]. As for prior work in the simple steiner tree problem the reader can refer to [KR95].

2.1 Rounding Algorithm

In this subsection, we present the rounding algorithm that we will use on the fractional solution of problem (1), and then prove our main results that give us the approximation ratio as well as the feasibility bounds.

Rounding Algorithm. Given the solution from problem (1) we do the following. Let $p(e)$ be the parent edge of edge e (since G is a tree, each edge in G has exactly one parent edge). Mark every edge independently with probability $x_e/x_{p(e)}$. For every leaf, add the path to the root the the integer solution *if and only if* all edges on the path are marked.

Lemma 1. *The expected cost of the integer solution is at most $\sum_{e \in E} c_e x_e$.*

Proof. Let \mathcal{E}_e be the event that “edge e is in the integer solution”. Then for any edge e we have:

$$\begin{aligned} \Pr[\mathcal{E}_e] &\leq \frac{x_e}{x_{p(e)}} \cdot \frac{x_{p(e)}}{x_{p(p(e))}} \cdots x_{p(p \dots)} \\ &= x_e \end{aligned}$$

Then let INT be the cost of the rounded solution.

$$\begin{aligned} \mathbb{E}[INT] &= \sum_{e \in E} \Pr[\mathcal{E}_e] \cdot c_e \Rightarrow \\ \mathbb{E}[INT] &\leq \sum_{e \in E} x_e \cdot c_e \end{aligned}$$

This concludes the proof of Lemma 1. □

Lemma 2. *For any R_i : $\Pr[\text{some vertex in } R_i \text{ is connected to } r \text{ in the integer solution}] \geq \frac{1}{H}$*

Proof. Before we continue with the proof, we first present **Janson’s inequality** that we will use later: Let a ground set X and some randomized process of independently selecting subsets $S \subset X$. Also let the following subsets of X : B_1, \dots, B_k and the event \mathcal{E}_i : $B_i \subseteq S$. Then define the following parameters:

$$\begin{aligned} \sum_i \Pr[\mathcal{E}_i] &= \mu \\ \sum_{i \sim j} \Pr[\mathcal{E}_i \cap \mathcal{E}_j] &= \Delta \\ \text{Where: } i \sim j &\implies B_i \cap B_j \neq \emptyset \end{aligned}$$

Then Janson’s inequality (JI) states:

$$\text{If } \mu \leq (1 - \varepsilon) \cdot \Delta, \text{ for some } \varepsilon > 0, \text{ then } \Pr[\cap_i \bar{\mathcal{E}}_i] \leq \exp\left(\frac{-\mu^2(1-\varepsilon)}{2\Delta}\right)$$

We now apply JI to our analysis:

- The ground set X is the set of edges E .
- S is the selected (by the rounding) edges of E .
- B_i is the path to some leaf x_i of the fixed group R_i .
- \mathcal{E}_i is the event that we picked the path to x_i in our solution.
- $\cap_i \bar{\mathcal{E}}_i$ is the event that we picked no paths to any terminals x_i of the group.

Now we compute the parameters μ and Δ .

$$\mu = \sum_i \Pr[\mathcal{E}_i] = \sum_e x_e \geq 1$$

This means we can set $\mu = 1$.

$$\begin{aligned}
\Delta &= \sum_{i \sim j} Pr[\mathcal{E}_i \cap \mathcal{E}_j] \\
&= \sum_j \sum_e \sum_{i: LCA(x_j, x_i)=e} Pr[\mathcal{E}_i \cap \mathcal{E}_j] \\
&= \sum_j \sum_e \frac{x_j}{x_e} \sum_{i: LCA(x_j, x_i)=e} x_i \\
&\leq \sum_j \sum_e \frac{x_j}{x_e} x_e && \text{(this holds because of the flow constraint: } \sum x_i \leq x_e \text{)} \\
&= \sum_j \sum_e x_j \\
&\leq H
\end{aligned}$$

Then applying JI we get:

$$\begin{aligned}
Pr[\cap \bar{\mathcal{E}}_i] &\leq \exp\left(-\frac{\mu^2(1-\epsilon)}{2\Delta}\right) \leq^{\epsilon=0.5} \exp\left(-\frac{1}{4H}\right) \\
&\sim 1 - \frac{1}{H}
\end{aligned}$$

This concludes the proof of Lemma 2. □

Theorem 3. *The cost of the rounded solution is no more than $2 \cdot \log kH + 1$ times the optimal solution.*

Proof. To prove the statement we use Lemma 2. We know that running one time the algorithm gives us probability of success for any group of $\geq 1/H$. For that reason we repeat the procedure $H \cdot 2\log k$ times and we have: $Pr[\text{Failure for group } R_i] \leq (1 - 1/H)^{H \cdot 2\log k} = ((1 - 1/H)^H)^{2\log k} \leq (\frac{1}{e})^{2\log k} = \frac{1}{k^2}$.

We apply union bound and we have: $Pr[\text{Failure for any group}] = 1/k$. Then the remaining unconnected groups we connect them using the cheapest edges. □

Fact 4. *The Group Steiner Tree can be used to encode set cover, and hence it's a NP-hard to compute a $\log k$ approximate solution. This also means that for general graphs the problem is $\Omega(\log^{2-\epsilon} n)$ -hard to approximate.*

3 Summary

In this lecture we covered the edge-weighted group steiner tree problem when the input graph is a tree, and we saw a new technique that uses dependent randomness to round our fractional solution.

References

- [DHK09] ErikD. Demaine, MohammadTaghi Hajiaghayi, and PhilipN. Klein. Node-weighted steiner tree and group steiner tree in planar graphs. In Susanne Albers, Alberto Marchetti-Spaccamela, Yossi Matias, Sotiris Nikolettseas, and Wolfgang Thomas, editors, *Automata, Languages and Programming*, volume 5555 of *Lecture Notes in Computer Science*, pages 328–340. Springer Berlin Heidelberg, 2009.

- [GKR00] Naveen Garg, Goran Konjevod, and R. Ravi. A polylogarithmic approximation algorithm for the group steiner tree problem. *Journal of Algorithms*, 37(1):66 – 84, 2000.
- [KR95] Philip N. Klein and R. Ravi. A nearly best-possible approximation algorithm for node-weighted steiner trees. *J. Algorithms*, 19(1):104–115, 1995.