22C:253 Lecture 18

Scribe: Zhihong Wang

November 14, 2002

Analysis of the primal-dual Steiner forest algorithm (cont.).

Claim 2 A cut S is active iff S is a connected component w.r.t the current set of chosen edges and $f(S) = 1$.

Proof:

($\Leftarrow$) If S is a connected component and $f(S) = 1$, then S is unsatisfied. Furthermore, S is minimal, because any proper subset $S' \subset S$ has an edge going out of S'.

($\Rightarrow$) Suppose S is active but S is not a connected componet, clearly, no currently chosen edge crosses $(S, \bar{S})$. Hence S is the union of two or more connected componets. Since S is active,$f(S) =$ 1. Hence for some $u \in S$ and $v \in \bar{S}$, $r(u, v) = 1$. Suppose $u \in C$ for some connected component C in S, then C is unsatisfied, implying S is not a minimal unsatisfied cut, a controdiction. □

Claim 3 $\sum_{e \in F'} C_e \leq 2 \cdot \sum_{S \in V} Y_S \cdot f(S)$

Proof:

$$\sum_{e \in F'} C_e = \sum_{e \in F'} \Big(\sum_{S: e \in \delta(S)} Y_S \Big) = \sum_{S \subseteq V} \Big(\sum_{e: e \in \delta(s) \cap F'} Y_S \Big) = \sum_{S \subseteq V} Y_S \cdot |\delta(S) \cap F'| = \sum_{S \subseteq V} Y_S \cdot deg_{F'}(S)$$

where $deg_{F'}(S) = |\delta(S) \cap F'|$, denoting the number of edges in F' that cross $(S, \bar{S})$, which has no relation to Y_S.

We need to show that

$$\sum_{S \subseteq V} Y_S \cdot deg_{F'}(S) \leq 2 \cdot \sum_{S \in V} Y_S \cdot f(S)$$

We will show something stronger, that is,

$$Changes\ in\ L.H.S \leq Changes\ in\ R.H.S.$$

Initially, L.$H.S = R.H.S. = 0$. Consider any arbitrary iteration and let Δ be the increate in Y_S, during that iteration,

$$Changes\ in\ L.H.S = \sum_{active\ S} deg_{F'}(S) = \Delta \cdot \sum_{active\ S} (S)$$

$$Changes\ in\ R.H.S. = 2 \cdot \sum_{active\ S} \Delta \cdot f(S) = 2 \cdot \Delta \cdot (\ number\ of\ active\ cuts\ S)$$

We want to show that

$$\sum_{activeS} deg_{F'}(S) \leq 2 \cdot \Delta \cdot (number\ of\ active\ cuts\ S)$$

that is, the average degree of active cuts w.r.t. F':

$$\frac{\sum deg_{F'}(S)}{number of active S} \leq 2$$

To finish the proof, we need one additional claim.

Claim 4 Let C be a component w.r.t. the currently chosen set of edges such that $f(C) = 0$, then $deg_{F'}(C) \neq 1$.

Proof: Suppose the claim is false, that is $f(C) = 0$ but $deg_{F'}(C) = 1$. So there exists a unique $e \in F'$ that crosses $(C, \bar{C})$.

Since $e \in F' \Rightarrow e$ is not redundant w.r.t. F'.

$\Rightarrow e$ is an edge on a unique $u - v$ path for some u,v, and $r(u, v) = 1$

$\Rightarrow$ W.l.o.g. $u \in C and v \in \bar{C}$.

$\Rightarrow f(C) = 1$ a controdiction. □

Claim 4 tells us that any inactive component C has $deg_{F'}(C) = 0$(i.e. it is isolated) or $deg_{F'}(C) \geq 2$. From this observation, the result follows. □

To show that the analysis is tight for this algorithm. Consider the following example:

$V = \{1, 2, 3, ..., n, (n + 1)\}(1, 2, ..., (n + 1)$ are labels on vertices), where $1, 2, 3, ..., n \in K_n$ and edges in K_n cost 2 each, edges from $(n+1)$ to each vertex in K_n have unit cost. And $S_1 = \{1, 2, ...n\}$.

The OPT = n. Cost of solution is $2 \cdot (n - 1)$.

Upper Bound on Integrality Gap Let OPT_f denote the optimal solution for primal problem.

$$OPT \leq \sum_{e \in F'} C_e \leq 2 \cdot \sum y_S \cdot f(S),\ and$$

$$\sum y_S \cdot f(S) \leq OPT_{dual} = OPT_f$$

$$\Rightarrow \frac{OPT}{OPT_f} \leq 2$$

thus giving a upper bound on integrity gap.

Lower Bound on Integrality Gap Consider a cycle on n vertices, with all edges of cost 1. The cost of dual solution found by algorithm is $\frac{n}{2}$, which is the optimal for the dual because there is a primal feasible solution with cost $\frac{n}{2}$. Therefore,

$$OPT_f = \frac{n}{2}, OPT = (n - 1)$$

$$\Rightarrow \frac{OPT}{OPT_f}\ (is\ essentially\ ) \geq 2$$

We will discuss "Facility Location Problem" nextly.

Facility Location Problem

Input:A set C (of cities), a set F (of facilities). The cost of opening facility $i \in F$ is f_i. The cost of servicing a city $j \in C$ using a facility $i \in F$ is C_{ij}.

Output: A set $I \subseteq F$ of open facilities and a function $\Phi : C \to I$ such that total cost

$$(\sum_{i \in I} f_i + \sum_{j \in C} C_{\Phi j})$$

is minimized.

We will discuss a factor-3 approximation algorithm using the primal-dual schema.