CS 683: Advanced Algorithms	**Multicommodity Flow: Graph Partioning**
February 6, 2001	**Scribes: Catherine Grasso, Sharad Goel**

1 Introduction

Recall that in the multicommodity flow problem (MCF), we are given a graph G(V,E) and source-sink pairs (s_i, t_i), with the goal of maximizing the amount of flow that can be simultaneusly sent between each $s_i - t_i$ pair. We denote this maximum flow value by λ^*. Previously, we saw that by duality this maximum flow corresponds to a minimum length assignment on the edges. That is,

$$\lambda^* = y^* \stackrel{def}{=} \text{minimum length assignment} = \min \sum_e y_e \text{ subject to } \sum_{i=1}^{k} d(s_i, t_i) \geq 1.$$

Also recall that the "narrowest cut" $\eta(G, \{s, t\})$ of a graph G with a set {s,t} of source-sink pairs is,

$$\eta(G, \{s, t\}) = \min_{U \subset V} \frac{\text{e(U, V-U)}}{\text{number of pairs split by (U, V-U)}}.$$

Clearly, $\lambda^* \leq \eta(G, \{s, t\})$, however as we saw before it is possible to have strict inequality.

Our goal for this lecture is to show that in the all pairs MCF problem, $\eta(G, \{s, t\})$ does not get "too" big. In particular, we will show that $\eta(G, \{s, t\}) = O(y^* \log(n))$.

2 Leighton-Rao Algorithm

2.1 The Basic Idea

Let G(V,E) be a graph. In order to show that $\eta(G, \{s, t\}) = O(y^* \log(n))$, we will construct a cut (U, V-U). Our cut algorithm builds on two heuristics:

Heuristic One:
1) Randomly select a point in the graph.
2) Do BFS as long as the ratio between the edges added to nodes already visited is exponential.
3) Halt when you find a "channel", i.e. a region where the number of edges out of the set of previously visited nodes into the set of unvisited nodes drops off.

Termination of Heuristic One should result in a cut of ratio $O(y^* \log(n))$. If it does not, then some region S is a giant component. Then we must consider the long "tendrils", regions of nodes hanging off of S, in order to find the appropriate cut:

Heuristic Two:
1) Do BFS out from nodes in S to find "tendrils."
2) Cut one of these "tendrils."

2.2 The Algorithm

Theorem 2.1 *(Region Growing): Let G be a graph with edge lengths y_e, and total length y^*. Let $\Delta > 0$ be a parameter. We can partition the nodes of G into sets that induce subgraphs of radius $\leq \Delta$ (with respect to edge lengths) so that the total number of edges joining different sets is $\leq 4\log_2(n)\frac{y*}{\Delta}$.*

Proof. First, we descretize the edge lengths so we can think in terms of breadth first search (BFS). I.e, if we have an edge of length y_e, we replace that edge with $\lfloor \frac{y_e}{y*}m + 1 \rfloor$ edges connected in a line (where m is the number of edges of G). Note that each edge in our discretized graph G* accounts for $\leq$ y*/m units of length in our original graph G. Furthermore, if m* is the number of edges in G*, then

$$m^* \leq \sum_e \frac{y_e}{y^*}m + 1 = m\sum_e \frac{y_e}{y^*} + \sum_e 1 = 2m.$$

Now note that if $\Delta \leq \frac{4y^*}{m}\log_2(n)$, then we can make each node of G its own set. So, we may assume that $\Delta \geq \frac{4y*}{m}\log_2(n)$. Define $\epsilon = \frac{y*}{\Delta m}\log_2(n) \leq \frac{1}{4}$.

Choose some v $\in$ G* and start a BFS at v. Let $m_0 = \frac{m^*}{n}$ and m_i = number of edges within i BFS steps from v. Stop the BFS at the first i such that $m_{i+1} < (1+\epsilon)m_i$. This defines a region in G*. Delete this region from G*, choose one of the remaining nodes, and repeat the above procedure until we have partioned G*. Now, if we charge each crossing edge to the boundary of the region whose growth accounted for it, we have

$$\text{number of edges crossing regions} \leq \epsilon(m^* + m_o n) = 2\epsilon m^* \leq 4\epsilon m = 4\log_2(n)\frac{y*}{\Delta}$$

Consider a region R of G*. If i is the radius of R, then,

$$m^* \geq \text{number of edges in R} \geq (1+\epsilon)^i m_0 = (1+\epsilon)^i \frac{m^*}{n}.$$

Therefore,

$$(1+\epsilon)^i \frac{m^*}{n} \leq m^*$$

and,

$$i * \log_2(1+\epsilon) \leq \log_2(n).$$

Since $\epsilon < 1/4, \log_2(1+\epsilon) \geq \epsilon$. So,

$$i \leq \frac{\log_2(n)}{\log_2(1+\epsilon)} \leq \frac{\log_2(n)}{\epsilon} = \Delta * \frac{m}{y^*}.$$

Now recall that each edge of G* accounts for $\leq$ y*/m units of length in G. So, if r is the radius of R in G, then,

$$r \leq \Delta * \frac{m}{y^*} * \frac{y^*}{m} = \Delta.$$

■

Now that we have proven the Region-Growing Lemma, we use it to find a "narrow cut."

Theorem 2.2 *Let y^* be as above. Then $\eta(G, \{s,t\}) \leq O(y^* \log(n))$.*

Proof. Set $\Delta = \frac{1}{2n^2}$ and run the region-growing algorithm on G. This results in one of two outcomes:
(1) Every region has $\leq \frac{2}{3}n$ nodes.
(2) Some region, S, is a giant component with $\geq \frac{2}{3}n$ nodes.
If (1), then we can find a low-ratio cut as follows:

Combinatorial Fact: Given $x_1, x_2, ..., x_k$ where $\sum_{i=1}^{k} x_i \leq n$ and each $x_i \leq \frac{2}{3}n$, it is possible to partition the x_i into two sets each of total sum $\leq \frac{2}{3}n$.

Using this fact we can repartition the nodes in the k components found by the region-growing algorithm to construct two sets U and V-U each of size $\leq \frac{2}{3}n$ with $\leq \frac{4y^* \log_2(n)}{\Delta}$ edges between them. Then the ratio of this cut $\eta_{(U,V-U)}(G, \{s,t\})$ is,

$$\eta_{(U,V-U)}(G, \{s,t\}) \leq \frac{4y * \log_2(n)}{\Delta} \frac{9n^2}{2} \leq \frac{18y^* \log_2 n}{\Delta n^2} = 36y^* \log_2 n \text{ since } \Delta = \frac{1}{2n^2}.$$

So, $\eta(G, \{s,t\}) \leq \eta_{(U,V-U)}(G, \{s,t\}) = O(y^* \log(n))$.

Otherwise we are in situation (2) and we claim that some cut between the layers of the BFS beyond S has the desired ratio of O($y^* \log n$). We find this cut by extending the BFS tree of S out to all of G*.

Let S be the giant region we found from our region growing algorithm (i.e. $|S| > \frac{2}{3}n$). Let m_i be the number of edges in between the ith layer of nodes found by doing BFS out from S and the (i-1)th layer of nodes found by doing BFS out from S. And let n_i be the number of nodes at $\geq i$ steps from S. If σ_i is the ratio of the "ith layer cut," then $\sigma_i \leq \frac{m_i}{n_i} \frac{3}{2n}$. Let $\tilde{\sigma} = \min_i \sigma_i$. Note that it is enough to show that $\tilde{\sigma} \leq O(y^* \log(n))$. Now $m_i \geq \tilde{\sigma} n_i \frac{2}{3}n$. Summing over all i, we have

$$2m \geq m^* \geq \sum_i m_i \geq \frac{2}{3} n \tilde{\sigma} \sum_i n_i = \frac{2}{3} n \tilde{\sigma} \sum_{u \in V(G)-S} dist_{G^*}(u, s). \tag{1}$$

Furthermore, we have that

$$1 \leq \sum_{\{u,v\}} d_G(u,v) \leq \sum_{\{u,v\}} (d_G(u,S) + d_G(v,S) + 2\Delta) < (n \sum_u d_G(u,S)) + \underbrace{\binom{n}{2} \frac{1}{n^2}}_{<\frac{1}{2}}.$$

So,

$$n \sum_u d_G(u,S) \geq \frac{1}{2} \text{ and } \sum_u d_G(u,S) \geq \frac{1}{2n}.$$

In G*, all distances are expanded by $\geq \frac{m}{y^*}$. So,

$$\sum_i n_i = \sum_u d_{G^*}(u,S) \geq \frac{1}{2n}\frac{m}{y^*}.$$

From equation (1) above,

$$2m \geq \frac{2}{3} n\tilde{\sigma} \sum_{u \in V(G)-S} dist_{G^*}(u,s) \geq \frac{1}{3}\tilde{\sigma}\frac{m}{y^*}$$

So,

$$\tilde{\sigma} \leq 6y^* \ll O(y^* \log n)).$$

■