Approximation Algorithms - Tools and Techniques **Spring 2002**

Lecture 13 : Feb 22, 2002

Lecturer: Kamesh Munagala [1] *Scribe: Siddartha Y R*[2]

Cut Problems have many applications in areas such as VLSI design and clustering. They are $\mathcal{NP}-hard$. MIN-CUT is poly-time, as shown by the Max-Flow/Min-Cut Theorem.

The Minimum Multicut Problem

Given graph $G(V,E)$ with weight c_e on edge e and a set of node pairs $\{(s_1,t_1),(s_2,t_2),\ldots,(s_k,t_k)\}$; To find the minimum cost set of edges to be deleted from the graph so that each of the node pairs are disconnected - s_i is disconnected from t_i, for each i.

The problem is known to be $\mathcal{NP}-hard$. In this lecture we develop an approximation algorithm using LP rounding. To start with, we have an IP formulation. The IP variables are

$$x_e = \begin{cases} 1; & \text{if } e \in \text{CUT} \\ 0; & \text{otherwise} \end{cases}$$

The objective is to

$$\min \sum_{e\in E} x_e c_e$$

The constraints ensure that there is no path between s_i and t_i. Let P_i be the set of paths between s_i and t_i. Then

$$\sum_{\substack{e\in p\\ p\in P_i}} x_e \geq 1;\ \forall p \in P_i,\ \forall i \in \{1,\ldots,k\}.$$

For the LP relaxation, we allow $x_e \in [0,1]$. However, the number of constraints is the number of paths between s_i and t_i, for each i; and is not polynomial, and therefore, neither is the time complexity of a standard LP algorithm solving it. We use the separation theorem of Lecture. 12 to work around this problem.

From the Separation Theorem, we know that to solve such an LP problem in poly-time; it is required to determine the following in polytime:

[1]Dept of Comp. Sci., Stanford Univ.
[2]Dept of Comp. Sci. and Automation, Indian Institute of Science, Bangalore.
`siddu@csa.iisc.ernet.in`

1. If a given setting of x_es satisfies all the constraints, and
2. If it does not, to find the violated constraint.

Consider the graph G with the x_es *as edge lengths*. Determine the shortest path from each s_i to t_i with these edge lengths. If any of the shortest paths has length ≤ 1, then that path violates the constraint. It is clear that this can be done in poly-time. Then, from the Separation Theorem, this LP can be solved in poly-time.

An LP Formulation for Shortest Path

Imagine the graph edges to be formed from strings of length given by the edge length, connecting "rocks" representing the nodes. Let $\rho(t_i)$ represent the "height" to which the rock corresponding to t_i can be lifted, with s_i lying on the ground (and of course, without breaking any of the strings). Then, the shortest path between s_i and t_i can be viewed as the maximum height to which t_i can be lifted with s_i lying on the ground, i.e

$$\max \rho(t_i)$$

subject to

1. s_i is on the ground
$$\rho(s_i) = 0.$$
2. No string breaks. i.e, for any edge (u, v), whose length is $l(u, v)$,
$$\rho(v) \leq \rho(u) + l(u, v)$$
3. $$\rho(v) \geq 0;\ \forall v$$

The number of constraints is polynomial. To check that this encodes the shortest path, we show the following:

Claim 1 *If $\rho^*(t)$ is the maxima returned by the* LP *solution, then*

1. $\rho^*(t) \leq$ *Shortest path length form s to t.*
2. $\rho^*(t) \geq$ *Shortest path length form s to t.*

Proof:

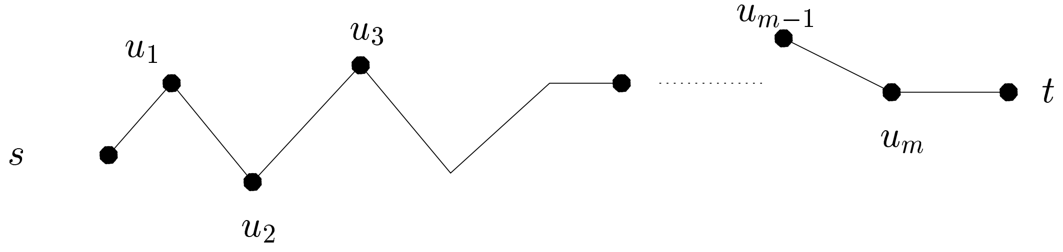

Figure 1: The shortest path between s and t.

1. Let the shortest path between s and t be given by $su_1u_2\cdots u_mt$ as shown in Figure. 1. Consider the $\rho(\cdot)$ returned by the LP for each node in this path. Then we have, from the LP constraints:

$$\begin{aligned} \rho(u_1) \leq & \quad \rho(s) + l(s, u_1) \\ \rho(u_2) \leq & \quad \rho(u_1) + l(u_1, u_2) \\ & \vdots \\ \rho(u_m) \leq & \quad \rho(u_{m-1}) + l(u_{m-1}, u_m) \\ \rho^*(t) \leq & \quad \rho(u_m) + l(u_m, t) \end{aligned}$$

Adding the above equations and canceling terms common to both sides, we have the required result.

$$\rho^*(t) \leq \sum_{(u_i, u_{i+1}) \in \text{ shortest path}} l(u_i, u_{i+1})$$

2. Let $\bar{\rho}(u)$ be the distance from s to u by the shortest path from s to any u, including t. Then, verify that this $\bar{\rho}(u)$ is a feasible solution for the above LP for u. This implies:

$$\rho^*(u) \geq \bar{\rho}(u)$$

which gives us the required result.

■

Now, we consider an alternate LP formulation for the MULTICUT problem, using the ideas developed for the shortest path problem above. We again have

$$x_e = \begin{cases} 1; & \text{if } e \in CUT \\ 0; & \text{otherwise} \end{cases}$$

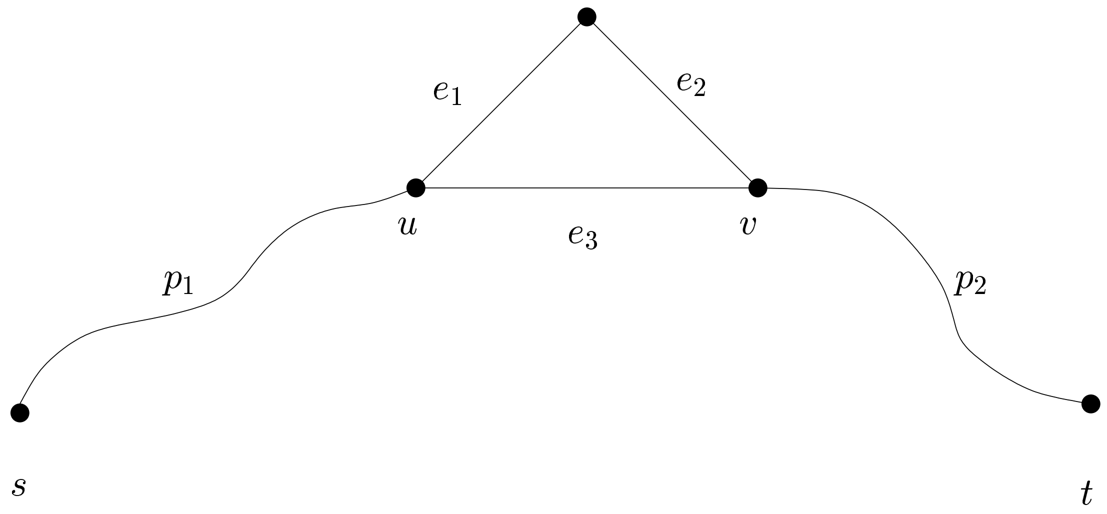

Figure 2: $x(e)$ forms a metric over V.

and $\rho_i(v)$ is the height of v with s_i on the ground. The LP seeks to minimise the 'cut' as defined by the x_es, while maintaining a 'distance' (with x_e as length again) more than 1 between each s_i and t_i.

$$\min \sum_e c_e x_e$$

$$\begin{array}{rcll} \text{s.t. } \rho_i(s_i) & = & 0 & \forall i \\ \rho_i(t_i) & \geq & 1 & \forall i \\ \rho_i(v) & \leq & \rho_i(u) + x_e & \forall e = (u, v) \\ \rho_i(v) & \geq & 0 & \forall v \end{array}$$

Claim 2 *x_e's returned by the* LP *form a metric over V.*

Proof: In Fig. 2, p_1 is a path from s_i to u and p_2 is a path from v to t_i. The LP solution satisfies[3]

$$len(p_1) + x(e_3) + len(p_2) \geq 1, \quad (1)$$

$$len(p_1) + x(e_1) + x(e_2) + len(p_2) \geq 1. \quad (2)$$

If $x(e_3) > x(e_1) + x(e_2)$, then, by reducing $x(e_3)$ to $x(e_1) + x(e_2)$, Eq. 1 is not violated, since Eq. 2. As the LP is minimising $x(e_3)$, it sets $x(e_3)$ such that it is just sufficient to avoid violating the (only) constraint it appears in; i.e. Eq. 1. Hence it sets $x(e_3) = x_(e_2) + x(e_1)$. The metric so obtained is called a "Spreading Metric". ∎

[3]In the following, x_e, etc will also be represented as $x(e)$ whenever it improves readability.

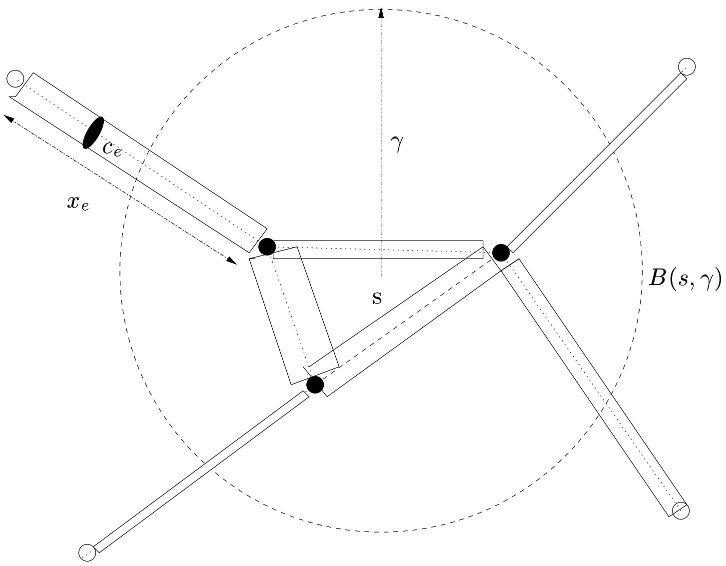

Figure 3: The graph as a system of pipes

View the edges of the graph as "pipes" with cross-sectional area c_e and length x_e. The objective function $\sum_e x_e c_e$ is just the total volume of the system of pipes.

Let $B(v,\gamma)$ denote a ball of radius γ centered at vertex v.

$$B(v,\gamma) = \{u \in V | dist(u,v) \le \gamma\}.$$

This ball may cut a number of edges and possibly some vertices, as shown in Fig. 3. Let $\delta_B(s_i,\gamma)$denote the edge-set cut by the ball. Consider the volume of the pipe system that lies inside the ball, given by $V(v,\gamma)$,

$$V(s_i,\gamma) = \sum_{u,v \in B(s_i,\gamma)} c(u,v)x(u,v) + \sum_{u,v \in \delta_B(s_i,\gamma)} c(u,v)(\gamma - dist(s_i,u)).$$

where, for an edge $(u,v) \in \delta_B(s_i,\gamma), u \in B(s_i,\gamma)$ and $v \notin B(s_i,\gamma)$. See that $V(s_i,\gamma)$ is a non-decreasing function of γ. It may be discontinuous, but the discontinuity is only caused when the ball crosses a vertex; i.e. if there exist no vertex v such that $dist(s_i,v) = r$, then $V(s_i,\gamma)$ is continuous at r (Thus, there are a finite number of discontinuities). Consider the rate of change of $V(s_i,\gamma)$ at such a region of continuity. The rate of change is just the net cross-sectional area of all the pipes that the ball is presently cutting. Call this rate of change $C(s_i,\gamma)$, and we have

$$C(s_i,\gamma) \triangleq \frac{dV(s_i,\gamma)}{d\gamma} = \sum_{e \in \delta_B(s_i,\gamma)} c_e.$$

Instead of setting $V(s_i, 0)$ to zero, set it to $\frac{\phi}{k}$ where

$$\phi \triangleq \sum_e c_e x_e$$

and k is the number of source-destination pairs in the problems. Now, define a new "volume" measure as

$$\tilde{V}(s_i, \gamma) \triangleq \frac{\phi}{k} + V(s_i, \gamma).$$

The net volume is thus increased by $k\frac{\phi}{k} = \phi$ to 2ϕ (as remarked before, the volume corresponds to the cost of the cut in our original problem).

> **Mean Value Theorem**
>
> If $f(X)$ is continuous except for a finite number of points of discontinuity and differentiable in the interval $[a, b]$ (a weaker condition is sufficient, actually) then, there exists a point $c \in (a, b)$ such that
>
> $$f'(c) = \frac{f(b) - f(a)}{b - a}$$

Theorem 1 *There exists an* $r \in [0, \frac{1}{2}]$ *such that*

$$\frac{C(s_i, r)}{\tilde{V}(s_i, r)} \leq 2 \ln 2k.$$

Proof: Let $\tilde{V}(s_i, \gamma)$ correspond to f and $C(s_i, \gamma)$ corresponds to f'. We prove the theorem by contradiction. Assume

$$\begin{array}{rrcll}
 & \frac{f'(\gamma)}{f(\gamma)} & > & \ln 2k & \forall \gamma \in [0, \frac{1}{2}]. \\
\Rightarrow & \int_0^{\frac{1}{2}} \frac{f'(\gamma)}{f(\gamma)} d\gamma & > & 2\ln(2k) \cdot \frac{1}{2} & \\
\Rightarrow & \ln(f(\gamma))\Big|_0^{\frac{1}{2}} & > & \ln 2k & \\
\Rightarrow & \ln\left(\frac{f(\frac{1}{2})}{f(0)}\right) & > & \ln 2k & \\
\Rightarrow & f(\frac{1}{2}) & > & 2kf(0) = 2k\frac{\phi}{k} = 2\phi &
\end{array}$$

But we know that the total (modified) volume was previously obtained as 2ϕ and this volume includes edges incident on t_i which is distant at least 1 from s_i and hence, is not covered by the above range of $\gamma \in [0, \frac{1}{2}]$. We have a contradiction. ∎

We now develop an algorithm for MULTICUT using the above ideas and results.

We consider the s_i one at a time in arbitrary order. For each s_i grow a ball $B(s_i, \gamma)$ until the condition of Theorem 1 is satisfied. We know that such a γ occurs in $[0, \frac{1}{2}]$. We remove the ball $B(s_i, \gamma)$ from further consideration, pick another s_j which is still connected to t_j, and repeat. We finally output the union of the edges cut. The balls we obtain satisfy the following conditions

1. t_i lies outside the ball $B(s_i, \gamma)$.
2. $C(s_i, \gamma) < 2 \ln 2k \tilde{V}(s_i, \gamma)$.
3. $\nexists j$ such that s_j and t_j lie inside the ball $B(s_j, \gamma)$.

For each cut, it holds that ratio of C and $\tilde{V}$ is less than $2 \ln 2k$. Since the volumes removed are disjoint and each cut satisfies the above relation, the total cut is at most

$$\begin{aligned} &\leq \quad 2 \ln 2k \times \text{ Total Volume} \\ &\leq \quad 2 \ln 2k \times 2\phi \\ &\leq \quad 4 \ln 2k \times \phi \end{aligned}$$

This is the best known LP rounding for MULTICUT. LP rounding for MULTICUT has an integrality gap of $\ln k$.