MATH 8500 Algorithmic Graph Theory, Spring 2017, OSU
Lecture 2: Multiway Cut
Instructor: Anastasios Sidiropoulos
Scribe: Austin Antoniou

1 An Approximation Algorithm for Multiway Cut

The Multiway Cut problem is similar in spirit to the Min Cut and $s-t$ Min Cut problems, but turns out to be significantly more complex.

Input: Suppose $G = (V, E)$ is a simple, undirected graph with edge weight $w : E \to \mathbb{R}^+$ and let $S = \{s_1, \ldots, s_k\} \subseteq V$ be a subset of the vertices (we will refer to the elements of S as "terminals").

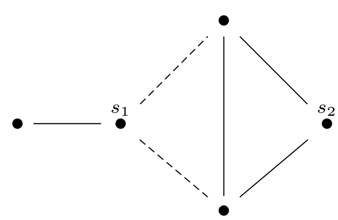
Goal: Find a subset $E' \subseteq E$ such that

(i) Each connected component of $G \setminus E'$ contains at most one terminal

(ii) The total weight $w(E') = \sum_{e \in E'} w(e)$ is minimized

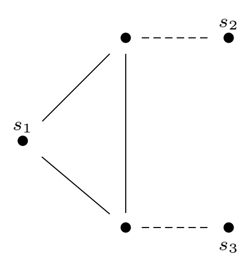
Note that, for $k = 2$, this is simply the $s-t$ Min Cut Problem.

Example 1. Consider the following graph with two terminals in which all edge weights are all 1:



The dashed edges above give an optimal multiway cut (notice, however, that this is not a minimum cut in the sense of the Min Cut Problem).

Now we consider another example with uniform edge weights and three terminals:



Again, the dashed edges give an optimal multiway cut in this example.

Theorem 1. *If $k = 3$, the multiway cut problem is NP-hard.*

We see now that, despite the apparent similarity of this problem to the s-t min cut problem, we should not expect to find an algorithm which yields a solution in polynomial time. We will instead seek a fast algorithm which yields an "approximate" solution: a cut which in not necessarily minimal, but is "not too far" from minimal. Our procedure for accomplishing this is the following.

Isolating Cut Heuristic:

For $i = 1, \dots, k$:

Construct the graph G_i as follows:

G_i has vertex set $V(G_i) = V \cup \{t_i\}$ and edge set $E(G_i) = E \cup \{\{s_j, t_i\} : j \neq i\}$.

For each $j \neq i$, extend the weight w by setting $w(\{s_j, t_i\}) = \infty$.

Compute the s_i-t_i min cut E_i in G_i.

End for.

By renumbering, assume that $w(E_1) \leq w(E_2) \leq \cdots \leq w(E_k)$.

Set $A = E_1 \cup \cdots \cup E_{k-1}$.

Return A.

Claim 1. *The set A as above is a valid (not necessarily minimal) multiway cut in G.*

Proof. For $i < k$ and $j \neq i$, s_i and s_j are in separate components of $G \setminus E_i$, which contains as a subgraph $G \setminus A$.

For $i = k$ and $j \neq i$, we have $j < k$, so s_k and s_j are in separate components of $G \setminus E_j$, which contains $G \setminus A$.

Thus the s_i all lie in distinct connected components of $G \setminus A$. □

Now we must formalize and prove our claim that A is an "approximately minimal" multiway cut. Suppose E^* is *some* optimal solution and let $V_1, \dots, V_k$ be the connected components of $G \setminus E^*$, where $s_i \in V_i$.

For any $U \subseteq V$, let the *boundary* of U be given by $\partial(U) := \{\{u, v\} : u \in U, v \notin U\}$.

Claim 2. *For all $i = 1, \dots, k$, $w(E_i) \leq w(\partial(V_i))$.*

Proof. This is true since $\partial(V_i)$ is an s_i-t_i cut but E_i is an s_i-t_i *min* cut, so E_i has smaller weight. □

Claim 3. $2w(E^*) = w(\partial(V_1)) + \cdots + w(\partial(V_k))$

Proof. We can prove this by "double counting" the edges in E^* (with multiplicity given by the weight of each edge). Each edge of E^* appears in the boundary of exactly two connected components, so each edge is associated to two terms on the right hand side of the equation in the claim. □

Theorem 2. *The Isolating Cut Heuristic has approximation ratio $2 - \frac{2}{k}$; in other words,*

$$\frac{w(A)}{w(E^*)} \leq 2 - \frac{2}{k}$$

Proof.

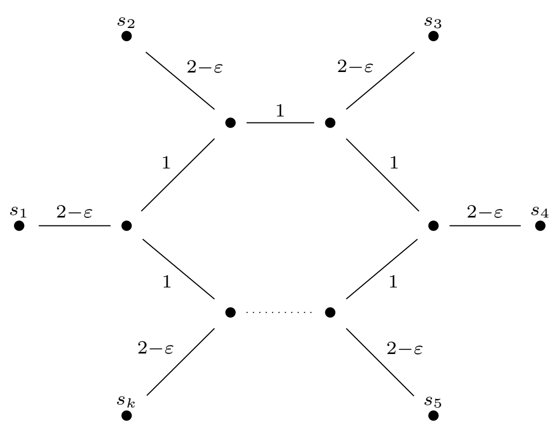
$$
\begin{aligned}
w(A) \le \sum_{i=1}^{k-1} w(E_i) &\le (1 - \frac{1}{k}) \sum_{i=1}^{k} w(E_i) && \text{since } w(E_k) \ge \tfrac{1}{k}\textstyle\sum_{i=1}^{k} w(E_i) \\
&\le (1 - \frac{1}{k}) \sum_{i=1}^{k} w(\partial(V_i)) && \text{by Claim 2} \\
&= (1 - \frac{1}{k})(2w(E^*)) && \text{by Claim 3}
\end{aligned}
$$

□

2 A tight example for the Isolating Cut Heuristic

We have now seen that the Isolating Cut Heuristic yields a solution which has no more than twice the optimal cost. Our proof of this was not terribly complex, which raises the question: is it possible that this procedure can, in general, yield a solution better than what we have just described? The next example shows us that the answer to this question is "no"; the bound $w(A) \le 2w(E^*)$ is tight.

Example 2. In the following graph



the optimal solution has total weight $w(E^*) = k$, given by choosing all of the edges with weight 1 (those that form the k-cycle).

However, the Isolating Cut Heuristic will cut each edge joining a terminal to the rest of the graph. At the end of the procedure, we will have removed $k-1$ edges of weight $2-\varepsilon$, so $w(A) = (k-1)(2-\varepsilon)$. Thus

$$
\frac{w(A)}{w(E^*)} = \left(\frac{k}{k-1}\right)(2-\varepsilon)
$$

which becomes arbitrarily close to 2 as k becomes large and as ε is chosen to be sufficiently small.