CMPSCI 711: More Advanced Algorithms

Graphs 4: Insert-Only Matchings

Andrew McGregor

Overview:

- A $2+\epsilon$ approx for the max weighted matching in $O(\epsilon^{-1} n \log n)$ space.

[Paz and Schwartman. SODA 17]

Last Compiled: February 15, 2018

Graph Matchings

Definition

A matching in graph $G = (V, E)$ is a subset of edges $M \subset E$ such that no two edges share an end point.

Problem

Find a matching M that maximizes $|M|$. If edges are weighted, we want to maximize $w(M) = \sum_{e \in M} w(e)$.

We'll assume all weights are $1, 2, \ldots, \text{poly}(n)$.

Unweighted Matching

- Let $M \leftarrow \emptyset$
- For each new edge e: add e to M if no edges in M share an endpoint with e

Theorem

Algorithm uses $O(n \log n)$ space and returns a 2 approximation to the maximum weighted matching.

Proof.

- Let $\textsc{opt} = \{o_1, o_2, \ldots\}$ be set of edges in the optimal solution.
- Let M be final set of selected edges and note M is maximal.
- For each $e \in \textsc{opt}$, charge \$1 to an edge $f \in M$ that shares an endpoint of e. There must exist such f because M is maximal. Every edge in M gets charged as most \$2. Hence,

$$2|M| \geq \text{ charges received } = \text{ charges made } = |\textsc{opt}|$$

□

Weighted Matching Algorithm

- $H \leftarrow \emptyset$ and $\phi(v) \leftarrow 0$ for all $v \in V$
- For each $e_i = \{u, v\} \in \{e_1, e_2, \ldots, e_m\}$:
  - If $w(e_i) > (1 + \epsilon)(\phi(u) + \phi(v))$:

$$H \leftarrow H \cup \{e_i\}$$

$$\phi(u) \leftarrow \phi(u) + (w(e_i) - \phi(u) - \phi(v))$$
$$\phi(v) \leftarrow \phi(v) + (w(e_i) - \phi(u) - \phi(v))$$

- Construct a greedy matching in H by considering edges in the reverse order they were added.

Lemma

$|H| = O(\epsilon^{-1} n \log n)$

Lemma

Algorithm is a 2 approximation when $\epsilon = 0$.

Corollary

Algorithm is a $2(1 + \epsilon)$ *approximation.*

Algorithm stores at most $O(\epsilon^{-1} n \log n)$ edges.

- Consider an arbitrary vertex v in the graph.
- The value of $\phi(v)$ is set to at least 1 when the first edge incident to v is added to H. Every time another edge that is incident to v is added to H, the value of $\phi(v)$ increases to at least

  $$\phi(v) + (w(e_i) - \phi(u) - \phi(v)) > \phi(v) + \epsilon(\phi(u) + \phi(v)) \geq (1+\epsilon)\phi(v) .$$

- If $\phi(v)$ becomes larger that the max edge weight, no more edges incident to v are added to H.
- At most $\log_{1+\epsilon} \mathrm{poly}(n)$ edges incident to v are added to H.

Algorithm returns a 2 approximation if $\epsilon = 0$: Part 1

- Let max weight matching have edges M^*. Let M be the matching returned and define $M_i = M \cap \{e_i, \ldots, e_m\}$. Define edge weights

$$w_i(e) = w(e) - \phi_i(u) - \phi_i(v)$$

where $\phi_i(\cdot)$ are the values just before ith edge in the stream. Note

$$w_{i+1}(e) = \begin{cases} w_i(e) & \text{if } e \text{ doesn't share endpoint with } e_i \\ w_i(e) - w_i(e_i) & \text{if } e \text{ shares one endpoint with } e_i \\ w_i(e) - 2w_i(e_i) & \text{if } e = e_i \end{cases}$$

if e_i was added to H and $w_{i+1} = w_i$ otherwise.

- Will show $w_i(M^*) \leq 2w_i(M_i)$ for all i by induction on decreasing i.
- Base case: $w_m(M^*) \leq 2w_m(M_m)$ because all $w_m(e) \leq 0$ for all edges except possibly e_m.
- Induction hypothesis: $w_{i+1}(M^*) \leq 2w_{i+1}(M_{i+1})$.

Algorithm returns a 2 approximation if $\epsilon = 0$: Part 2

- If $e_i \notin H$ then $w_i = w_{i+1}$ and so

$$w_i(M^*) = w_{i+1}(M^*) \leq 2w_{i+1}(M_{i+1}) = 2w_i(M_{i+1}) = 2w_i(M_i)$$

- Otherwise, assume $e_i \in H$ and let N be set of edges intersecting e_i in G. Then,

$$w_i(M^*) \leq w_{i+1}(M^*) + 2w_i(e_i) \leq 2w_{i+1}(M_{i+1}) + 2w_i(e_i)$$

  since at most two edges are in $N \cap M^*$ and other weights stay same.

- Since M_i has at least one edge in N and hence

$$w_i(M_i) \geq w_{i+1}(M_i) + w_i(e_i) \geq w_{i+1}(M_{i+1}) + w_i(e_i)$$

  and therefore $w_i(M^*) \leq 2w_i(M_i)$ as required.

Algorithm returns a $2(1+\epsilon)$ approximation

Define a new set of edge weights w' as follows: Run the algorithm with $\epsilon > 0$ and when we encounter e, define

$$w'(e) = \begin{cases} w(e)/(1+\epsilon) & \text{if } \phi(u)+\phi(v) < w(e) \leq (1+\epsilon)(\phi(u)+\phi(v)) \\ w(e) & \text{otherwise} \end{cases}$$

- Running algorithm with rule "add to H if $w'(e) > \phi(u)+\phi(v)$" is same as using rule "add to H if $w(e) > (1+\epsilon)(\phi(u)+\phi(v))$"
- We know using the first rule finds matching M with

$$w'(M) \geq w'(M^*_{w'})/2 \ .$$

  where $M^*_{w'}$ is the edges in the optimal matching with respect to w'.
- Since $w(\cdot)/(1+\epsilon) \leq w'(\cdot) \leq w(\cdot)$,

$$w(M) \geq w'(M) \geq w'(M^*_{w'})/2 \geq w'(M^*_w)/2 \geq \frac{w(M^*_w)}{2(1+\epsilon)}$$

  where M^*_w is the edges in the optimal matching with respect to w.