17 Nov 2023

Approx Max-Flow Min-Cut for the Sparsest Cut problem

Undirected graph G.

k pairs of terminals $\{(s_i, t_i)\}_{i=1}^{k}$

Edges have capacities $c(e)$.

If vertex set is partitioned into (A, B)

$\mathrm{cap}(A,B)$ = combined capacity of edges from A to B

$$\mathrm{sep}(A,B) = \#\{ i \mid \{s_i, t_i\} \cap A = 1 \}$$

$$\mathrm{sparsity}(A,B) = \frac{\mathrm{cap}(A,B)}{\mathrm{sep}(A,B)}$$ ← NP-hard to minimize!

If, in G, it is possible to route r units of flow simultaneously from s_i to t_i $\forall i$ then $\mathrm{sparsity}(A,B) \geq r$ $\forall$ cut A,B that separates at least one terminal pair.

$$\mathcal{Q} = \{ k\text{-tuples } (P_1, P_2, \ldots, P_k) \text{ of } s_i\text{-}t_i \text{ paths} \}$$

(PRIMAL)

$$\max \sum_{Q \in \mathcal{Q}} x_Q$$

$$\text{s.t.} \quad \sum_{e \in E} n_Q(e) x_Q \leq c(e) \quad \forall e \in E$$

$$x_Q \geq 0$$

(DUAL)

$$\min \sum c(e) y_e$$

$$\text{s.t.} \quad \sum_{e \in E} n_Q(e) y_e \geq 1 \quad \forall Q$$

$$y_Q \geq 0$$

2 interpretations of dual:

(1) y_e's are "edge lengths":

Dual constraint $\sum_{e \in Q} n_Q(e)\, y_e \geq 1 \quad \forall Q$

Summarized by saying

$$\sum_{i=1}^{k} \left(\text{shortest path length from } s_i \text{ to } t_i \text{ w.r.t. } y_e\right) \geq 1$$

(2) y_e's represent a "fractional cut"

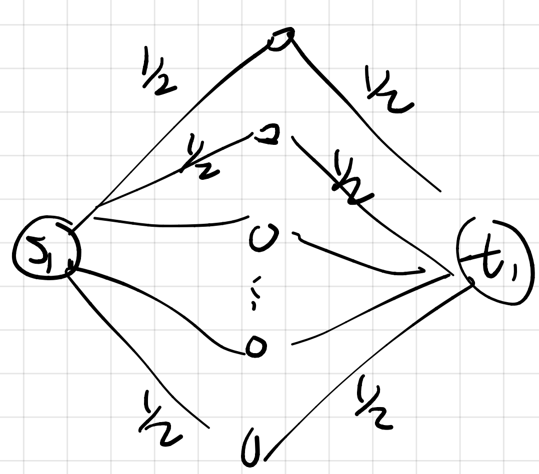
generalizing the case of an actual cut (A,B)

with $\text{sep}(A,B) = j$ corresponding to

$$y_e = \begin{cases} 1/j & \text{if } e \text{ has exactly one endpoint in } A \\ 0 & \text{otherwise.} \end{cases}$$

Tossing $\boxed{\text{independent}}$ coins to round y_e to a cut is bad!

E.g. $k=1$



This is one of the opt dual solutions. But indep coin tosses with these probabilities have $(3/4)^{n-2}$ probability of separating s_1–t_1.

Rounding:

1. Sample unif rand. $q \in \{1, 2, \ldots, \lfloor \log(2k) \rfloor\}$
2. Sample random subset $W \subseteq \{s_1, t_1, s_2, t_2, \ldots, s_k, t_k\}$ including each of the $2k$ vertices independently with probability 2^{-q}.
3. Compute shortest path distance $d(u,v)$ between all pairs of vertices using edge lengths y_e.

   Let $d(W, v) := \min\{d(u,v) \mid u \in W\}$.
4. Sample $r \in (0,1)$ unif random and let

$$A = \{v \mid d(W,v) \leq r\}$$

$$B = \{v \mid d(W,v) > r\}.$$

Game Plan:

1. Bound $\mathbb{E}[cap(A,B)]$ above.
2. Bound $\mathbb{E}[sep(A,B)]$ below.
3. Argue that ① & ② imply $\exists$ cut whose sparsity is $O(\log k)$ times $\sum_e c(e) y_e$.

① $\mathbb{E}[cap(A,B)] = \sum_{e=(u,v)} c(e) \Pr(e \text{ is cut by } A, B)$.

We cut $e=(u,v)$ precisely when $d(W,u) \leq r < d(W,v)$ or $d(W,v) \leq r < d(W,u)$.

Conditional on any W, this probability is

$$|d(W,u) - d(W,v)| \longleftarrow \text{interpreted as } 0 \text{ if } W = \emptyset.$$

Always,
$$d(W,u) - d(W,v) \le y_e$$
$$d(W,v) - d(W,u) \le y_e$$

$$\therefore \quad \mathbb{E}[\operatorname{cap}(A,B)] \le \sum c(e)\, y_e.$$

(2) $$\mathbb{E}[\operatorname{sep}(A,B)] = \sum_{i=1}^{k} \int_0^1 \Pr(A,B \text{ separates } s_i \text{ from } t_i \mid r = x)\, dx$$

Note A,B separates s_i from t_i iff

$$d(W,s_i) \le r < d(W,t_i)$$
or $$d(W,t_i) \le r < d(W,s_i)$$

Let $$T = \{s_1, t_1, \dots, s_k, t_k\}.$$
$$U_s = \{u \in T \mid d(s_i,u) \le r\}$$
$$U_t = \{u \in T \mid d(t_i,u) \le r\}$$

If $U_s \cap W \ne \emptyset$, $U_t \cap W = \emptyset$ then

$s_i \in A$, $t_i \in B$.

If $U_s \cap W = \emptyset$ $U_t \cap W \ne \emptyset$ then

$s_i \in B$, $t_i \in A$.

For $r < \frac{1}{2} d(s_i, t_i)$ $U_s \cap U_t = \emptyset$, both are non-empty.

Let $\ell = \#(U_s \cup U_t)$. If $2^q \le \ell < 2^{q+1}$

which happens w. prob. $\frac{1}{\log(2k)}$,

then $W \cap (U_s \cup U_t)$ is a sum of

ℓ indep random Bernoulli RV's

and each has expectation 2^{-8}.

$$\mathbb{E}\left[\# W \cap (U_s \cup U_t)\right] = \frac{\ell}{2^8} \in [1,2).$$

Lemma. A binomial distribution with exp val in $[1,2)$ has $\geq e^{-2}$ probability of sampling 1.

Conclude: $\forall\ 0 < r < \frac{1}{2} d(s_i, t_i)$,

$$\Pr(\text{separate } s_i, t_i \mid r) \geq e^{-2} \cdot \frac{1}{\log(2k)}.$$

$$\sum_{i=1}^{k} \int_0^1 \left(\text{this}\ \Uparrow\ \text{probability}\right)$$

$$\geq \sum_{i=1}^{k} \frac{1}{2} d(s_i, t_i) \cdot e^{-2} \cdot \frac{1}{\log(2k)}$$

$$\geq \frac{1}{2e^2 \log(2k)} \sum_{i=1}^{k} d(s_i, t_i) \geq 1$$