ACO Comprehensive Exam Fall 2024

Aug 16, 2024

1 Design and Analysis of Algorithms

Given an unweighted, undirected simple graph $G = (V, E)$ and two nodes $s \neq t \in V$, the min s-t cut problem is find a subset of edges $S \subseteq E$ of smallest size $|S|$ such that the graph $(V, E \setminus S)$ has no path from s to t. Consider the following LP to find the min s-t cut:

$$\min \quad \sum_{e \in E} x_e \quad \text{subject to}$$
$$\forall \text{paths } P \text{ from } s \text{ to } t \text{ in } G, \quad \sum_{e \in P} x_e \geqslant 1$$
$$\forall e \in E, \quad x_e \geqslant 0$$

1. (3 points) Prove that the optimal objective value of this LP is at most the min s-t cut value. Give a polynomial-time separation oracle for this LP.
2. (3 points) Given an optimum solution x_e^* for $e \in E$ to the above LP, we will now round it to find an s-t cut of size at most $\sum_{e \in E} x_e^*$.

   Consider the graph $G = (V, E)$ with length of each edge e being x_e^*. These edge weights induce a shortest path metric on V. For $r \geqslant 0$, let $B_r \subseteq V$ denote all the vertices at distance at most r from node s. Show that for a uniform random $r \in [0, 1]$, the expected number of edges in the cut $(B_r, V \setminus B_r)$ is at most $\sum_{e \in E} x_e^*$.
3. (2 points) Show how we can derandomize the above procedure to obtain an s-t cut of size at most $\sum_{e \in E} x_e^*$.
4. (2 points) Write the dual LP of the above LP. Name the algorithmic problem that this dual LP corresponds to and interpret the dual variables.

2 Combinatorial Optimization

Let K be a complete graph with vertex set V and $T \subseteq V$ be a subset of vertices with even cardinality and let k be a positive integer. Let H be an edge-disjoint union of any k T-joins of K. We would like to show the following statement:

(*) For any $s \in T$, there exists a vertex $v \in V$ such that there are at least k-edge disjoint paths from s to v in H.

Prove the following statements below to infer the statement above.

1. (1 point) Given $s \in T$, it is sufficient to find a vertex $v \in V$ such that the minimum s-v cut in H is at least k.
2. (2 points) Show that if $S \subseteq V$ is such that $|S \cap T|$ is odd, then $\delta_H(S) \geqslant k$.
3. (1 point) Consider the Gomory-Hu tree of H and show that it is sufficient to find an edge of the Gomory-Hu tree incident at s that has weight at least k.
4. (6 points) Using the Gomory-Hu tree or otherwise, show that there exists a vertex v such that minimum s-v cut in H is at least k.

3 Probabilistic Combinatorics

Problem. Let $p = p(n) \in [0, 1]$ and let $G \sim \mathbb{G}(n, p)$. Let D be the random variable equal to the largest size of a set of vertex-disjoint triangles in G. Prove the following statements:

(a)[4 points] If $p \ll n^{-2/3}$, then $D = o(n)$ with high probability.

(b)[6 points] If $p \gg n^{-2/3}$, then $D = (1 - o(1))(n/3)$ with high probability. (Hint: use Janson's inequality.)

4 Solutions

(Algorithms)

1. Consider an optimal s-t cut $S \subseteq E$ with $|S| = C^*$. Define a feasible solution to the LP by setting $x_e = 1$ for all $e \in S$ and $x_e = 0$ for all $e \notin S$. The reason this is feasible is because it is non-negative and for any s-t path P, at least one edge must belong to S (since this is a cut), ensuring $\sum_{e \in P} x_e \geqslant 1$. Thus, the LP's optimal objective value is at most C^*.

   Given a solution (x_e^*) for the LP, the separation problem asks us to check feasibility and if not then return a violated constraint. First, we can easily check non-negativity of all the edge variables x_e^* since there are only $|E|$ of them. Next, to check $\sum_{e \in P} x_e \geqslant 1$ for every path P, we find the shortest s-t path in the weighted graph where the length of each edge e is x_e^*. This can be done using Dijkstra's algorithm. If the shortest path length is strictly less than 1, then this path corresponds to a violated constraint; otherwise, no such path exists.

2. Consider any edge $e = (u, v)$. Let d_u and d_v denote the distances of these vertices from vertex s when the edge lengths are x_e^*. Observe that $|d_u - d_v| \leqslant x_e^*$ since we can reach from one vertex to another using edge e. Hence, the probability that edge e is in the random cut is at most x_e^*. By linearity of expectation, the expected number of edges in the cut is the sum of edge-cut probabilities, which is at most $\sum_{e \in E} x_e^*$.

3. Since the expected value of the cut size is at most $\sum_{e \in E} x_e^*$, there exists an r in $[0, 1]$ such that the cut size is at most $\sum_{e \in E} x_e^*$. Although there are an infinite number of potential r values, note that as we increase r from 0 the cut changes at most $|V|$ times. Hence, we can try all these $|V|$ potential r values and return the minimum of these cuts.

4. The dual LP can be written as:

$$
\begin{aligned}
\max \quad & \sum_{P \text{ is an } s\text{-}t \text{ path}} y_P && (1) \\
\text{subject to} \quad & \sum_{P \ni e} y_P \leqslant 1 \quad \forall e \in E && (2) \\
& y_P \geqslant 0 \quad \forall P && (3)
\end{aligned}
$$

   The dual LP corresponds to the maximum s-t flow problem, where dual variables y_P represent the flow along path P, and the constraints ensure that the total flow through each edge does not exceed edge capacity 1.

(Combinatorial Optimization)

(a) By Menger's theorem, we have that there are k edge disjoint path from s to v iff the minimum s-v cut is at least k.

(b) Observe that $\delta_H(S)$ is a T-cut and every T-join must intersect this T-cut. Since H contains k edge-disjoint T-joins, each of these T-joins contains at least one distinct edge in $\delta_H(S)$.

(c) By property of the Gomory-Hu tree, the weight of any edge $\{u, v\}$ is the value of the min-cut separating the two vertices. Thus if $\{s, v\}$-edge has weight at least k, the minimum s-v cut must have weight at least k.

(d) Consider the Gomory-Hu tree of H, say $\mathcal{T}$. Consider the components of $\mathcal{T} \setminus \{s\}$. Since $s \in T$ and $|T|$ is even, at least one such component with vertex set C must have $|C \cap T|$ be odd. Let v be the neighbor of s in this component defined by C. Then $\delta_H(C) \geqslant k$ by part (b) and moreover, $\delta_H(C)$ is the capacity of the edge $\{s, v\}$ by the property of Gomory-Hu tree. By part (c), the capacity of $\{s, v\}$ edge is exactly the value of the minimum s-v cut and thus it is at least k.

Probabilistic Combinatorics. *(a)* Let $\mathcal{T}$ be the set of all triangles in K_n. For each $T \in \mathcal{T}$, let A_T be the event that $T \subseteq G$ and let $\mathbf{1}_T$ be the corresponding indicator random variable. Define

$$X \;=\; \sum_{T \in \mathcal{T}} \mathbf{1}_T \quad \text{and} \quad \mu \;=\; \mathbb{E}[X] \;=\; \binom{n}{3} p^3.$$

Since $D \leqslant X$ by definition, we have $\mathbb{E}[D] \leqslant \mu = \Theta(n^3p^3)$. If $p \ll n^{-2/3}$, then $n^3p^3 = o(n)$, so we can fix a function $f = f(n)$ such that $n^3p^3 \ll f \ll n$. By Markov's inequality,

$$\mathbb{P}[D \geqslant f] \;\leqslant\; \frac{E[D]}{f} \;=\; O\left(\frac{n^3p^3}{f}\right) \;=\; o(1).$$

Hence, $D \leqslant f = o(n)$ with high probability, as desired.

(b) We shall apply Janson's inequality to prove the following claim:

Claim. *If $p \gg n^{-2/3}$, then the graph $G \sim \mathbb{G}(n,p)$ is triangle-free with probability at most $e^{-h(n)}$ for some function $h(n) \gg n$.*

Proof. To begin with, we replace p by $\min\{p, n^{-0.6}\}$ WLOG so that $n^{-2/3} \ll p \ll n^{-1/2}$. (This step is not necessary, but simplifies the argument.) For a pair of triangles T, $T' \in \mathcal{T}$, write $T \sim T'$ if $T \neq T'$ and the events A_T, $A_{T'}$ are not independent, i.e., $E(T) \cap E(T') \neq \varnothing$. In other words, $T \sim T'$ if and only if T and T' have exactly 2 common vertices. Now we compute

$$\Delta \;=\; \sum_{T \sim T'} \mathbb{P}[A_T \cap A_{T'}] \;=\; \underbrace{\binom{n}{3}}_{\text{choices for } T} \; \underbrace{3(n-3)}_{\text{choices for } T'} \; p^5 \;=\; \Theta(n^4p^5).$$

Since $p \ll n^{-1/2}$, it follows that $n^4p^5 = o(n^3p^3)$, so $\Delta = o(\mu)$. By Janson's inequality,

$$\mathbb{P}[X = 0] \;\leqslant\; \exp\left(-\mu + \frac{\Delta}{2}\right) \;=\; e^{-(1-o(1))\mu} \;=\; e^{-\Theta(n^3p^3)}.$$

As $p \gg n^{-2/3}$, we have $n^3p^3 \gg n$, as desired. ∎

Let $G \sim \mathbb{G}(n,p)$, where $p \gg n^{-2/3}$, and fix the function h given by the claim. Since h is super-linear, we may choose $m = m(n)$ so that $m \ll n$ and $h(m) \geqslant 100n$. We also choose $m(n)$ so that $p \gg n^{-2/3}$ implies that $p \gg m^{-2/3}$ as well (e.g., set $m(n) = n/g(n)$ and $p(n) = g(n)n^{-2/3}$ for any asymptotically growing function $g(.)$. Now for every fixed m-element subset $S \subseteq V(G)$, the induced subgraph $G[S]$ is distributed as $\mathbb{G}(m,p)$. Next we choose Since $p \gg m^{-2/3}$, we may apply the claim with $G[S]$ in place of G to conclude that $G[S]$ is triangle-free with probability at most $e^{-h(m)} \leqslant e^{-100n}$. By the union bound, it follows that G has a triangle-free induced subgraph on m vertices with probability at most

$$\binom{n}{m} e^{-100n} \;\leqslant\; 2^n e^{-100n} \;=\; o(1).$$

That is, with high probability every set of m vertices in G contains a triangle. Assuming this happens, we can form a family of at least $(n-m)/3$ vertex-disjoint triangles by iteratively picking a triangle and removing its vertices from the graph until fewer than m vertices remain. Therefore, $D \geqslant (n-m)/3 = (1-o(1))n/3$ with high probability, as desired.