COL 751 : Assignment - 1

Total marks: 60

1 Reporting Distances in Undirected Graphs

Attempt either of the following questions.

1.1 Distance Oracle [30 marks]
Let $G = (V, E)$ be an n-vertex undirected unweighted graph, and $R \subseteq V$ be a random subset of size $O(n^{1-a} \log n)$, where $a \leqslant \frac{1}{2}$ is a fractional parameter. Let $H = (V, E_H)$ be an auxiliary directed weighted graph derived from G as follows: *For each $x \in V$, compute a set $B(x, n^a)$ of closest n^a vertices to x, and add to H an edge from x to each $y \in B(x, n^a)$ of weight $dist(x, y, G)$.*

For each $x \in V$, let $p(x)$ be the vertex in R closest to x, and for $L \geqslant 0$, let $N(x, L)$ denote the set of all those vertices y whose distance from x is at most L. Further, for any $L > 0$, let

$$V_L = \{x \in V \mid N(x, L) \subseteq B(x, n^a)\}.$$

Question (a) Prove that with probability $1 - \frac{1}{n^2}$, for each $x \in V \setminus V_L,\ dist(x, p(x), G) \leqslant L$.

Question (b) Consider a pair of vertices (x, y) in G satisfying $dist(x, y, G) = D$, and let P be any arbitrary x to y shortest path in G. Prove that the following holds:

1. If all the vertices in P lie in $V_{\epsilon \frac{D}{2}+1}$, then there exists a vertex in $V(P)$ whose hop-distance[1] from x as well as y in H is at most $\lceil 1/\epsilon \rceil$.
2. If P contains a vertex from $V \setminus V_{\epsilon \frac{D}{2}+1}$, then there exists a vertex in $V(P) \setminus V_{\epsilon \frac{D}{2}+1}$ whose hop distance from either x or y is at most $\lceil 1/\epsilon \rceil$.

Question (c) Explain how to use above two parts to design an oracle of $O(n^{2-a} \log n)$ size that for any given $x, y \in V$ reports a $(1 + \epsilon, 2)$-approximation to (x, y) distance in G in $O(n^{a\lceil 1/\epsilon \rceil})$ time. (Hint: Use truncated BFS traversal in graph H from x and y).

1.2 Distance Estimation [10 marks]
Design an $O(n^{2.5} \log n)$ time algorithm that, given an n-node graph G with non-negative edge weights, computes for all pairs of vertices u, v an estimate $D(u, v)$, such that with high probability, for all pairs of vertices u, v for which there is a shortest path that uses at least $\sqrt{n}$ nodes, $D(u, v)$ is the distance between u and v.

[1] For any x, y, the (x, y)-hop-distance in a graph H is defined as the minimum number of edges needed to traverse a shortest x to y path in H.

2 Graph Spanners [20 marks]

Given an unweighted n-vertex graph $G = (V, E)$, a 6-additive spanner $H \subseteq G$ is a subgraph satisfying that $dist(u, v, H) \leqslant dist(u, v, G) + 6, \forall u, v \in V$.
In this exercise, we will construct, in a step by step manner, a 6-additive spanner H with $\widetilde{O}(n^{4/3})$ edges.[2]

(a) The first part defines a degree threshold Δ_1, and adds all edges incident to vertices with degree at most Δ_1 to H. How large can Δ_1 be (i.e., so that the edge bound of $O(n^{4/3} \log n)$ is maintained)?

(b) Next, the algorithm takes care of all shortest paths $\pi(u, v)$ that have at least one high-degree vertex with degree at least Δ_2. To do that, it samples each a random set Q of size $O(n \log n/\Delta_2)$. A BFS tree rooted at each vertex $q \in Q$ is added to H. What will be value of Δ_2?

(c) Finally, it remains to take care of paths that have no vertex with degree at least Δ_2. On each path $\pi(u, v)$, observe that all edges incident to low-degree vertices (vertices with degree at most Δ_1) are in H. Hence, when adding a shortest-path, we only "pay" for the number of missing edges (i.e. those edges that are incident to vertices with degree in range $[\Delta_1, \Delta_2]$).

We will take care of these paths in $O(\log n)$ phases. For every $i \in \{0, 1, \ldots, 2 \log n\}$, let Q_i be a uniformly random set of size $O(n \log n/2^i \Delta_1)$. For every vertex u that has a neighbor, say w, in Q_0, add one edge between u and w to the spanner.

In each phase $i \geqslant 1$, we take care of all paths $\pi(u, v)$ that have $x \in [2^{i-1}, 2^i]$ edges that are missing in H. This is done as follows. For each $t_1 \in Q_0$ and each $t_2 \in Q_i$, add to H the shortest $t_1 - t_2$ path in G that has at most 2^i missing edges in H. That is, among all paths between t_1 and t_2 in G that have at most 2^i edges that are not in H, pick the shortest one and add its edges to H.

Provide the missing details in above construction, and prove that H is a +6-spanner with $O(n^{4/3})$ edges.

3 Survey and Open Questions [10 marks]

Provide a brief half page survey of some of the most foundational works in graph spanners, distance oracles, and related hardness results. Also state at least 3-5 open questions. You may refer to survey here.

[2] $\widetilde{O}(\cdot)$ hides polylogarithmic factors.