Shortest Paths

Assumption: Unweighted graph, arrivals only

Algorithm:

$H \leftarrow \phi$

For (u,v) in stream do

If $d_H(u,v) > \alpha$ then

$H \leftarrow H \cup \{u,v\}$

return H

① Shortest path distances in H vs. true distances

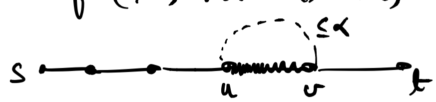
② Space ?

Lemma: $d_G(s,t) \leq d_H(s,t) \leq \alpha\, d_G(s,t)$

G: graph of all edges

Proof: 1st: $H \subseteq G$

2nd: If (u,v) not included, $\exists$ path of length α between u & v



H that satisfies this is called α-spanner

girth: length of shortest cycle

Lemma: H has girth at least $\alpha+2$

Proof (by contradiction) Suppose H had a cycle of length $\leq \alpha+1$

Consider last edge (u,v) added to cycle

$d_H(u,v) \leq \alpha \Rightarrow (u,v)$ would not be added

Thm: A graph with girth $2t+1$ has $O(n^{1+\frac{1}{t}})$ edges

Pf: m: # edges $\quad d = \frac{2m}{n}$ average degree

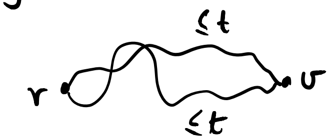
$\frac{d}{2}$ core of graph: iteratively remove nodes of degree $< \frac{d}{2}$

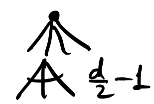
<u>Claim:</u> Resulting graph is non-empty

If we removed all vertices, we remove $< \frac{m}{n} \cdot n = m$ edges

Consider any of remaining nodes
Construct depth t tree by performing t steps of BFS

① Every node within depth t has unique path of length $\leq t$ to root





② Every node has degree $\geq \frac{d}{2}$

Tree has at least $\left(\frac{d}{2}-1\right)^t$ nodes $\leq n$

$$\left(\frac{m}{n}-1\right)^t \leq n$$

$$m \leq n^{1+\frac{1}{t}} + n = O\left(n^{1+\frac{1}{t}}\right)$$

$$\alpha + 2 = 2t+1$$

$$t = \frac{\alpha+1}{2} \Rightarrow \#\text{edges } O\left(n^{1+\frac{2}{\alpha+1}}\right)$$

<u>Note:</u> We ignored update time
Can achieve $O(1)$ update time by more sophisticated methods

Lower bound for shortest paths between all pairs of vertices

Exact distances need $\Omega(n^2)$ bits

$2^{\binom{n}{2}}$ labeled graphs on n vertices
must have distinct representation for each!

Why? For any pair $\exists$ (u,v) in one graph & not in other
$d(u,v) = 1$ for one
$d(u,v) > 1$ for other

So we need $\log\left(2^{\binom{n}{2}}\right) = \binom{n}{2}$ bits.

Lower bound for approximate distances

<u>Fact</u>: $\exists$ graph with $n^{1+\frac{1}{k}}$ edges and no cycles of length $\leq ck+1$ for some constant c

Erdős girth conjecture: $\exists$ graphs with $\Omega(n^{1+\frac{1}{k}})$ edges and girth $2k+1$

$2^{n^{1+\frac{1}{k}}}$ subgraphs

For every pair of subgraphs $\exists\ (u,v)$ s.t. $d(u,v) = 1$ in one
$d(u,v) > ck$ in other

In order to get ck approximation
we need $\log(2^{n^{1+\frac{1}{k}}})$ bits $= n^{1+\frac{1}{k}}$ bits.

Lower Bounds via Communication Complexity

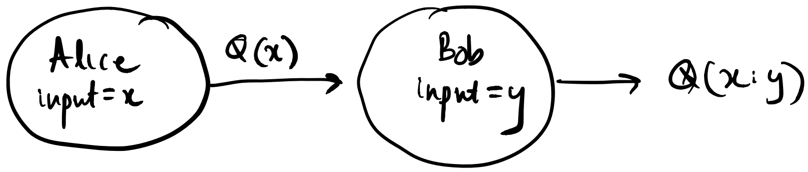
<u>Setup</u>: Alice has string $x \in \{0,1\}^a$
Bob has string $y \in \{0,1\}^b$
They need to compute $f(x,y)$

1. Communication over many rounds
   For streaming algorithms, we consider one-way communication
2. Deterministic, Randomized (public randomness)
3. Error probability $\varepsilon < \frac{1}{2}$ (say $\frac{1}{3}$)

Connection to Streaming

1. Small-space streaming algo imply low-communication 1-way protocol
2. Such a protocol does not exist (from Communication Complexity)

Stream: x, y



Disjointness Problem

Given 2 binary strings $x, y \in \{0,1\}^n$

$$\mathrm{DISJ}(x, y) = \begin{cases} 1 & \text{if } \exists i \;\; x_i = y_i = 1 \\ 0 & \text{otherwise} \end{cases}$$

<u>Lemma</u>: Every deterministic protocol for $\mathrm{DISJ}(x,y)$, $x,y \in \{0,1\}^n$ needs n bits of communication.

<u>Proof</u>: 2^n strings $\in \{0,1\}^n$

If we use less than n bits of communication

$\exists$ 2 strings that are mapped to same message

$x^{(1)}, x^{(2)}$ differ in some bit say i

$x_i^{(1)} = 1, \quad x_i^{(2)} = 0$

$y_i = 1, \quad y_j = 0 \quad j \neq i$

Thm Randomized protocol for $\mathrm{DISJ}(x,y)$ with success prob. $\geq \frac{2}{3}$, must use $\Omega(n)$ bits of communication

<u>Corollary</u>: Any randomized algorithm to estimate F_∞ within $(1 \pm \cdot 2)$ must use $\Omega(n)$ bits

<u>Pf</u>: If we can approximate F_∞, we would be able to compute disjointness

Given $x \in \{0,1\}^n \quad y \in \{0,1\}^n$

$S_x = \{i : x_i = 1\}$ $S_y = \{i : y_i = 1\}$

① $S_x \cap S_y = \phi$ $F_\infty \in \{0, 1\}$

② $S_x \cap S_y \neq \phi$ $F_\infty = 2$

<u>Lemma</u> (Yao) If there exists a distribution D over all possible input strings $(x, y) \in \{0,1\}^a \times \{0,1\}^b$ s.t. for any deterministic one-way protocol P with

$$\Pr_{(x,y)\in D}[P \text{ returns wrong answer on } (x,y)] \leq \varepsilon$$

the communication cost is $\geq K$ (public random bits) then any randomized one-way protocol with error $\leq \varepsilon$ on every input has communication cost $\geq K$

Any randomized protocol w. public randomness is distribⁿ over deterministic protocols!