COL 758: Advanced Algorithms

Lecture 11

Distance Oracle Problem

Problem:

Given a graph $G = (V, E)$, build a compact data structure $\mathscr{D}$ such that

for any query pair $(u, v) \in V \times V$,

the distance $dist(u, v)$ can be reported in $O(1)$ time.

Trivial Solution: Dist matrix $O(n^2)$ space $O(1)$ query time $O(mn)$ processing time

What if G has 10^{10} vertices, but storage is 10^{15} bytes

Can't store Dist matrix of $O(n^2)$ size

Approximate Distance Oracle

Problem:

Given a graph $G = (V, E)$, build a compact data structure $\mathscr{D}$ such that

for any query pair $(u, v) \in V \times V$,
(t)-approximation of distance $dist(u, v)$ can be reported in $O(1)$ time.

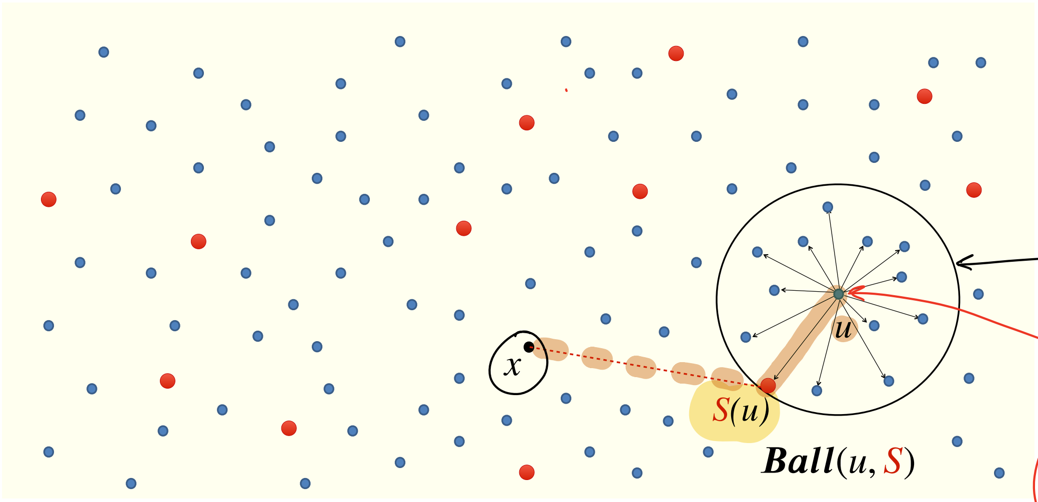
Goal: $dist(u,v) \leq \hat{d}(u,v) \leq (t) \cdot dist(u,v)$

Aim:

- Sub-quadratic space.
- S~~ub-cubic preprocessing tim~~e.
- $O(1)$ query time.

Main Idea

Define $S(u) :=$ The special vertex in "S" which is closest to u, for each vertex u



Locality of u

Normal vertex

(Has good connectivity in neighborhood)

For graph $G = (V, E)$

- Compute a *small* set S of ***special*** vertices.
- For each "special" vertex $s \in S$, store distance to ALL vertices in G
- For each "ordinary" vertex $u \in V \backslash S$, store distance to vertices present in its ***locality***.

dist from u upto first special vertex

Reporting Approximate Distance

$Ball(u, S) :=$ All vertices x with $dist(x, u) \cancel{\leq} dist(x, S(u))$

$S(u)$: Special vertex closest to u

To show:

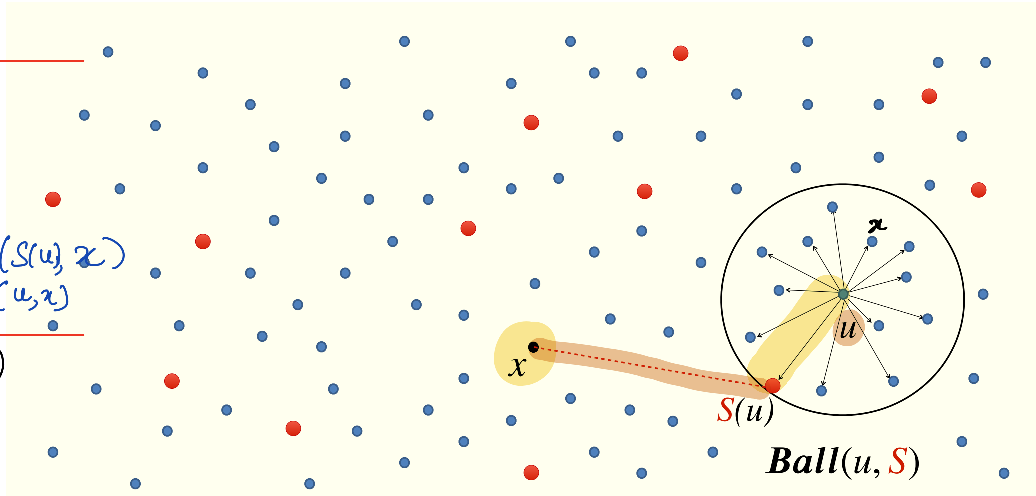
- $dist(x, u) \leq \hat{d}$

$x \notin Ball(u, S)$

$dist(u, S(u)) + dist(S(u), x) \geq dist(u, x)$

- $\hat{d} \leq 3\, dist(x, u)$

(?)



Query(x, u)**:**

If $x \in \boldsymbol{Ball}(u, S)$: **Return** exact distance

Else:

Return $\underbrace{dist(u, S(u)) + dist(S(u), x)}_{\hat{d}}$

↑ stretch $= 3 = O(1)$

(i) $dist(u, S(u)) \leq dist(u, x)$

(Bcoz $x \in Ball(u, S)$)

(ii) $dist(S(u), x) \leq dist(S(u), u) + dist(u, x)$

$\leq 2\, dist(u, x)$

(i) + (ii) $\Rightarrow \hat{d} \leq 3 \cdot dist(u, x)$

How to cleverly choose S to optimise space?

$$\text{Total Size: } |S|n + \sum_{u \in V} \textbf{Ball}(u, S)$$

Trick: Put each vertex in S with probability $p = \frac{1}{\sqrt{n}}$ (simplest way to choose "S")

$\text{Exp}(|S|) = pn^2$

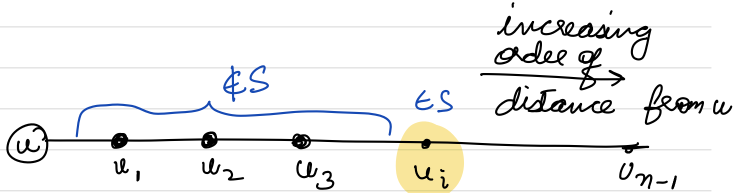
$\text{Exp}(\text{Ball}(u, S)) \leq \frac{1}{p}$ (why?)

(only non-trivial thing)

Expected total size $= pn^2 + \frac{n}{p}$

$= O(n^{1.5})$

for $p = 1/\sqrt{n}$



coin toss: p = prob of Heads

Ques: Expected no. of tosses to get first H?

$$\sum_{i=1}^{\infty} i\, p(1-p)^{i-1} = \frac{1}{p} \quad \text{(why?)}$$

To prove :

$$\sum_{i=1}^{\infty} i\, p(1-p)^{i-1} = \frac{1}{p}$$

Proof :

LHS :

$$p \underbrace{\left(\sum_{i=1}^{\infty} i\, x^{i-1} \right)}$$

$x = 1 - p$

$$= p \frac{d}{dx} \left(\underbrace{\sum_{i=1}^{\infty}} x^{i} \right)$$

$x + x^2 + x^3 + \ldots$

$$= p \frac{d}{dx} \left(\frac{1}{1-x} \right)$$

$$= -p \frac{d}{dp} \frac{1}{p}$$

$$= -p \left(-\frac{1}{p^2} \right) = \frac{1}{p} = \text{RHS}$$

Result for Distance Oracle

We can compute a distance oracle satisfying:

- **Expected Size:** $E\left(|S|n + \sum_{u \in V} \boldsymbol{Ball}(u, S) \right) = O(n^{3/2})$
- **Stretch:** 3
- **Query time** $= O(1)$

How to reduce

Size of oracle

(Homework, Not part of syllabus)