Today: Computation issues.

PageRank: Power iteration vs. Monte Carlo.

**Power iteration**

\[
\pi_0(u) = \frac{1}{N}
\]

\[
\pi_k(u) = \frac{\kappa}{N} + \frac{1}{N} \sum_{(w, v) \in E} \frac{\pi_{k-1}(w)}{\text{outdeg}(w)} (1 - \alpha)
\]

\[
|\pi(u) - \pi_k(u)| \leq (1 - \alpha) \sum_{w: (w, v) \in E} \frac{|\pi_{k-1}(w) - \pi(w)|}{\text{outdeg}(w)}
\]

\[
\|\pi - \pi_k\| \leq (1 - \alpha) \|\pi - \pi_{k-1}\|
\]

\[
\Rightarrow \|\pi - \pi_k\| \leq \|\pi - \pi_0\| (1 - \alpha)^k \leq 2(1 - \alpha)^k
\]

\[
\leq \frac{\varepsilon}{N} \quad \text{(target)}
\]

\[
\Rightarrow \# \text{of iterations} = \mathcal{O}\left(\frac{\log \frac{1}{\varepsilon}}{\log 1 - \alpha}\right) + (\log N) = \mathcal{O}\left(\frac{\log N}{\log 1 - \alpha}\right)
\]

\[
\# \text{of computations} = \mathcal{O}(M)
\]
Monte Carlo: via random walk.

Uniformly choose a node, \( v \).
At each stage stop w.p. \( \alpha \).
walk w.p. \( 1 - \alpha \).

On the average, takes \( \frac{1}{\alpha} \) steps, to generate one sample.

Want a weight for each of the \( N \) nodes. Take \( \frac{\log N}{\alpha \varepsilon^2} \) samples.

Here computation is \( \frac{N \log N}{\alpha \varepsilon^2} \), compared to \( \log N \cdot M \) for \( \frac{1}{\alpha \varepsilon^2} \) power method.

Incremental Pagerank. Graph changes slowly—Edges arrive, say.

\((v_t, v_t) = \text{arriving edge at time } t.\)
\(d_t(v) = \text{outdegree of } v.\)
\(\Pi_t(v) = \text{Pagerank at time } t.\)
Start $R = O(\log N)$ random walks from every node.
At time $t$, for every r.w. through $u_t$, reroute it through new edge $(u, v_t)$ w.p. $1/d_t(u_t)$.

# network calls for each rerouting $O(1/\varepsilon)$

We then maintain $R$ random walks.

Assume edges are chosen by an adversary to maximise computation time.

For the $t^{th}$ re-routings $= O(NR/\varepsilon t)$ (Claim).

Total time over many r.w. = $O(NR \log N)$

Proof of claim:

Expected number of rerouting at time $t = \mathbb{E}\left[\frac{NR \Pi_t(u_t)}{\alpha d_t(u_t)}\right]$

$\Pi_t(u_t) = \frac{d_t(u)}{t}$.

Then $\sum_u \mathbb{E}\left[\frac{NR \Pi_t(u)}{\alpha d_t(u)}\right]. \text{Prob. that one of them goes via } u_t$.

$\frac{1}{d_t(u_t)}$, prob. that walk passes through $(u, v_t)$. 

$O(NR \log N)$
Personalized PageRank (PPR)

Given s (source) and t (target), estimate the personalized PageRank of t w.r.t. s.

Monte Carlo uses time \( \geq 1/s \).
Local update uses time \( d/s \), \( d = M/N = \text{average degree} \).

\[ \delta = \frac{c}{n} \]

Claim: We can answer PPR queries for an average per-query running time to be \( O\left(\frac{d + \log n}{\sqrt{s}}\right) \).

Proof idea: Small number of forward walks from s.
Small number of backward walks from t.
Where the walks intersect.