18110063 HARSHIT KUMAR

CS328 : HOMEWORK 2

1. Let set of users be

$\{u_1, \ldots u_n\}$

& their personalized pagerank vectors be

$v_1, \ldots v_n$

respectively

$V = \{v_1, v_2 \ldots v_n\}$

Any personalized pagerank computed from V, without accessing the web graph will be of the form:

$v' = \alpha_1 v_1 + \alpha_2 v_2 + \ldots \alpha_n v_n$

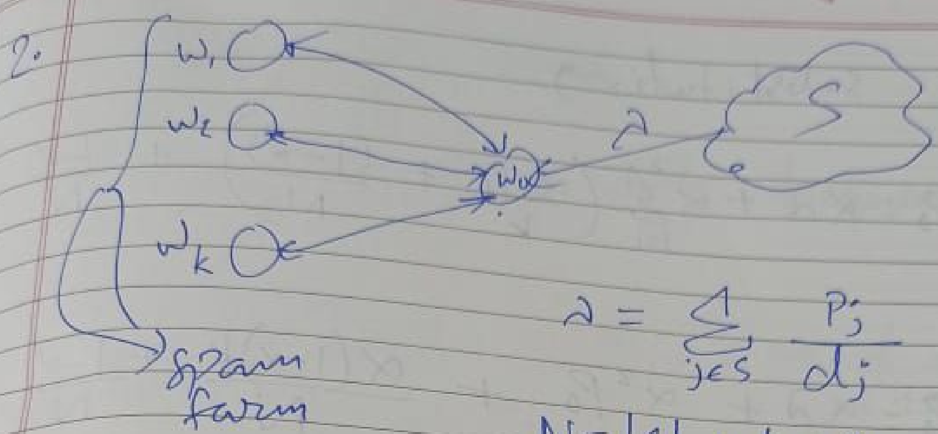
where $\alpha_1 \ldots \alpha_n \in \mathbb{R}$

$\therefore$ $v' \in \boxed{\text{span}(V)}$

~~all person~~ $\therefore$ span(V) is the set of all vectors computable from V

2.

2.



$$\lambda = \sum_{j \in S} \frac{P_j}{d_j}$$

$$N = |S| + k + 1$$

* P_0 depends on: pagerank from S
pagerank from spamfarm
pagerank from random jump

$$P_0 = \alpha \lambda + \alpha \sum_{i=1}^{k} P_i + \frac{(1-\alpha)}{N}$$

$$P_i = \frac{\alpha P_0}{k} + \frac{1-\alpha}{N}$$

(∵ ~~P_i received~~ $P_i \cdots P_k$ receive pagerank only from P_0 & random jump)

Substituting.

$$P_0 = \alpha\lambda + \alpha \sum_{i=1}^{k} \left(\frac{\alpha P_0}{k} + \frac{(1-\alpha)}{N} \right) + \frac{1-\alpha}{N}$$

$$P_0 = \alpha\lambda + \alpha^2 P_0 + \frac{\alpha(1-\alpha)k}{N} + \frac{1-\alpha}{N}$$

$$P_0(1-\alpha^2) = \alpha\lambda + \frac{\alpha k - \alpha^2 k + 1 - \alpha}{N}$$

$$P_0 = \frac{\alpha\lambda}{1-\alpha^2} + \frac{(1-\alpha)(\alpha k+1)}{N(1-\alpha^2)}$$

$$\boxed{P_0 = \frac{\alpha\lambda}{1-\alpha^2} + \frac{\alpha k+1}{N(1+\alpha)}}$$

3. n distinct items

Also, $\frac{C}{k^3}$ items with frequency k

$$\therefore \sum_{k=1}^{\infty} \frac{C}{k^3} = n \quad \left(\text{no. of distinct items}\right)$$

$$C \cdot \sum_{k=1}^{\infty} \frac{1}{k^3} = n$$

$$C \cdot a = n$$

, where a is a ~~constant~~ ~ 1.20205 (constant)

(Apéry's constant)

$$\therefore C \approx \frac{n}{1.202} \qquad \therefore \boxed{C = O(n)}$$

On fixing w & d, CS gives a better guarantee for this distribution.

$\therefore$ CS is better than CM always.