

MAC0323 Algoritmos e Estruturas de Dados II

Eduardo Pinheiro(8936798)

05-03-2016

1.6.15 Hubs and Authorities:

Hubs and authorities. Add to your generator from the previous exercise a fixed number of hubs, which have links pointing to them from 10% of the pages, chosen at random, and authorities, which have links pointing from them to 10% of the pages. Compute page ranks. Which rank higher, hubs or authorities?

Como o enunciado do problema não está bem definido, como o próprio professor comentou, abaixo listarei o que levei em consideração para resolver o problema, seguindo instruções passadas no Paca.

- Hubs e Authorities são adicionados ao número total, portanto se passado em linha de comando dez páginas dois hubs e três authorities o número total de páginas será quinze.
- Cada Hub e cada Authority terá exatamente 10% das páginas seja mandando link ou recebendo link.
- O algoritmo recebe 4 argumentos da linha de comando, sendo eles: número de páginas, número de links, número de hubs e número de authorities. tendo a forma:

```
$ java -algs4 Generator N M H A > file.txt
```

Para testar os links gerados:

```
$ java -algs4 Transition < file.txt | java -algs4 Markov 40
```

Ou ainda, gerando um histograma:

```
$ java -algs4 Transition < file.txt | java -algs4 RandomSurferHistogram 1000000
```

Which rank higher, hubs or authorities?

Em uma análise puramente intuitiva, chegamos a conclusão que um hub, ou seja, uma página que é apontada por muitas outras terá um page rank maior pois no modelo aleatório utilizado ela terá mais chances de ser sorteada. Quando passamos para uma abordagem empirica do problema obtemos o mesmo resultado, como esperado a página que recebe mais links tem um posicionamento melhor, como podemos ver nas imagens a seguir:

- utilizando: `java -algs4 Generator 100 60 4 2`



- utilizando: `java -algs4 Generator 50 40 4 2`



*Cada barra representa uma página e as barras vermelhas representam hubs.

**Histograma gerado por `RandomSurferHistogram.java`

Verificamos então que hubs tem uma posição maior no algoritmo de page rank.