

CS498HS4: Computational Advertising

Homework 1

Zhenye Na (zna2)
September 26, 2018

Question 1 (5 points)

(a)

Show the values that you get if you run two rounds of computing hub and authority values on the network of Web pages in Figure 14.16. (That is, the values computed by the k -step hub-authority computation when we choose the number of steps k to be 2.)

Show the values both before and after the final *normalization* step, in which we divide each authority score by the sum of all authority scores, and divide each hub score by the sum of all hub scores. (We will call the scores obtained after this dividing-down step the *normalized scores*. It's fine to write the normalized scores as fractions rather than decimals.)

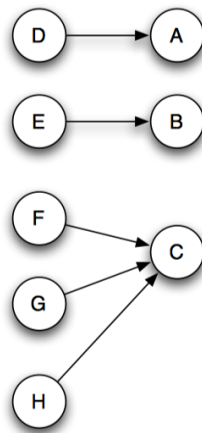


Figure 14.16: A network of Web pages.

Solution:

Number of steps k is 2 in this question, so we need update all hub scores and all authority scores twice.

auth scores

	A	B	C
initial state	1	1	1
$k=1$	1	1	3
$k=2$	1	1	9

hub scores

	D	E	F	G	H
initial state	1	1	1	1	1
$k=1$	1	1	3	3	3
$k=2$	1	1	9	9	9

Before Normalization:

The steps to reach this result is that, I first initialize hub scores and authority scores to be 1. A single application of the *Authority Update Rule* followed by a single application the Hub Update Rule produces the results of the original list-finding technique. So in 2 iterations, update $auth(p)$ to be the sum of the hub scores of all pages that point to it and update $hub(p)$ to be the sum of the authority scores of all pages that it points to.

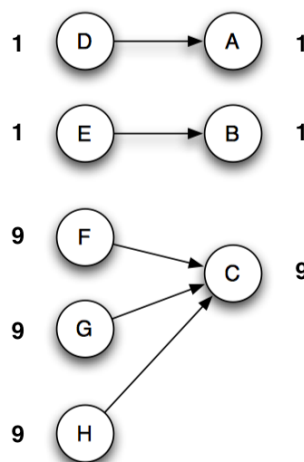


Figure 14.16: A network of Web pages.

After Normalization:

For *normalization*, I just divide each *auth*/*hub* score by the sum of *auth*/*hub* scores

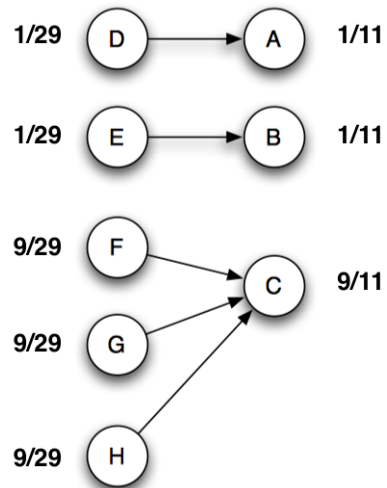


Figure 14.16: A network of Web pages.

(b)

After normalization:

In this case, because node E points to Node C, when we calculate Hub score for Node E, the “input” authority scores come from Node B and Node C. So this breaks the symmetry.

auth scores

	A	B	C
initial state	1	1	1
$k=1$	1	1	4
$k=2$	1	5	17

hub scores

	D	E	F	G	H
initial state	1	1	1	1	1
$k=1$	1	5	4	4	4
$k=2$	1	22	17	17	17

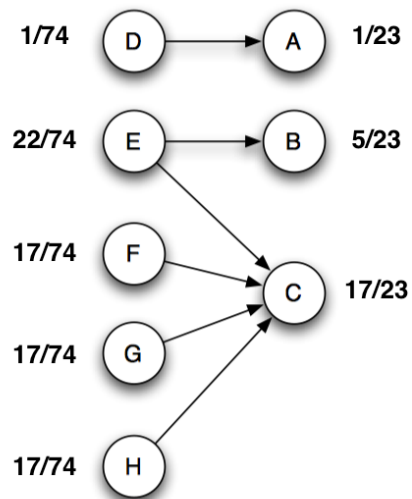


Figure 14.17: A network of Web pages.

(c)

Solution:

The input hub score for Node A is from Node D and input *hub* score for Node B is from Node E. However, Node E also points to Node C, which has lots of input Nodes. After Authority Update Rule and Hub Update Rule, Node B will have a higher authority score.

Question 2 (10 points)

(a)

Before Normalization:

hub scores

	A	B
initial state	1	1
$k=1$	3	2
$k=2$	11	7

auth scores

	C	D	E	F
initial state	1	1	1	1
$k=1$	3	3	5	2
$k=2$	11	11	18	7

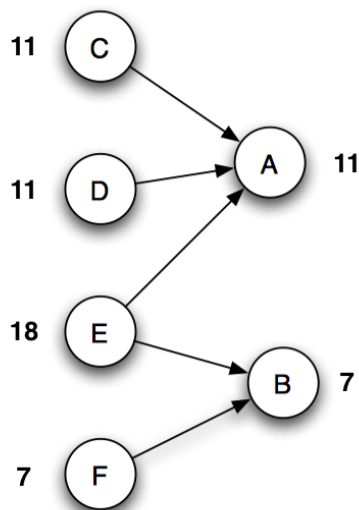


Figure 14.18:

After Normalization:

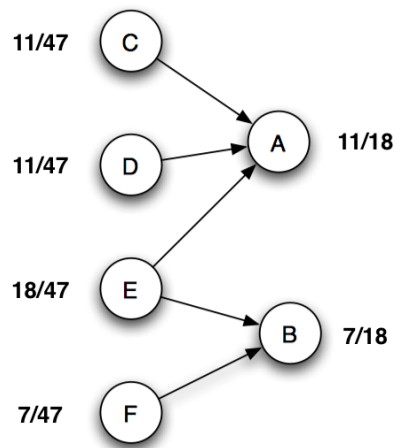


Figure 14.18:

(b)

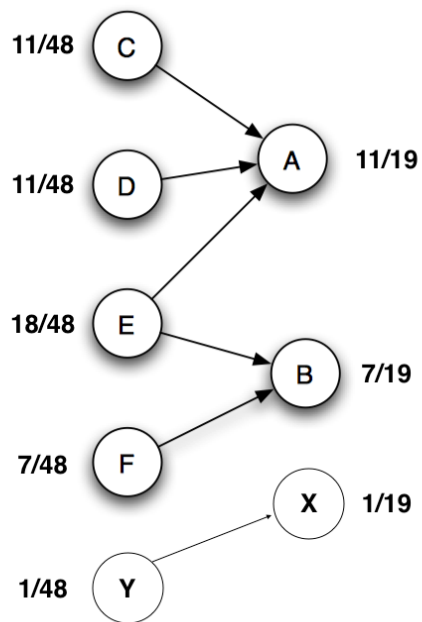
Option 1:

auth scores

	A	B	X
initial state	1	1	1
$k=1$	3	2	1
$k=2$	11	7	1

hub scores

	C	D	E	F	Y
initial state	1	1	1	1	1
$k=1$	3	3	5	2	1
$k=2$	11	11	18	7	1



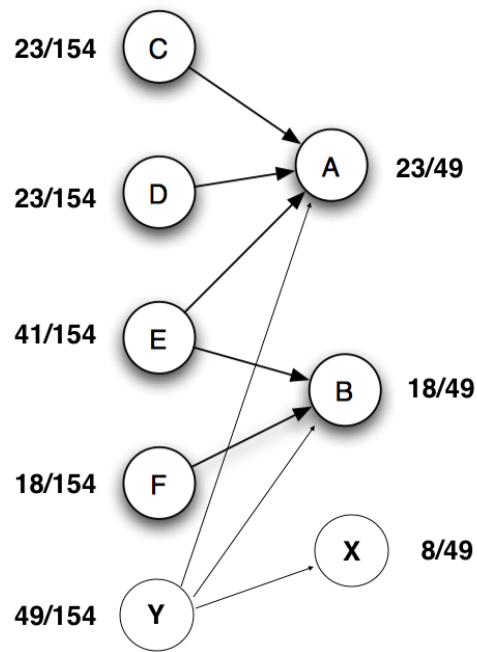
Option 2

auth scores

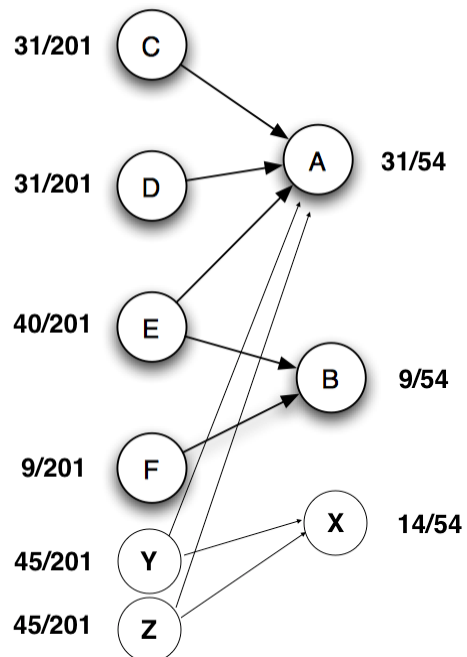
	A	B	X
initial state	1	1	1
$k=1$	4	3	1
$k=2$	23	18	8

hub scores

	C	D	E	F	Y
initial state	1	1	1	1	1
$k=1$	4	4	7	3	8
$k=2$	23	23	41	18	49



(c)



The strategy for placing Node X, Y and Z is that since we would like to place Node X at the second position when we rank the pages, then "incoming" hub scores to Node X is supposed to

as high as possible, it will not be higher than Node A, but can be higher than Node B. So Node Y and Z should not connect to Node B. This will keep the "incoming" hub scores of Node B the same but increase them of Node X.

Question 4 (2 points)

- Review movies with high authority scores.
- No, the effects in small graphs are more obvious than larger ones.