

Exercise 1

$T_1 =$ "a beautiful toy"

$T_2 =$ "Toy one and toy two"

$T_3 =$ "a new toy"

$T_4 =$ "one two...toy"

We build the inverted list:

		$\log_2 \frac{4}{2} = 1$
a	→ 1, 3	$\log_2 \frac{4}{1} = 2$
and	→ 2	$\log_2 \frac{4}{1} = 2$
beautiful	→ 1	$\log_2 \frac{4}{1} = 2$
new	→ 3	$\log_2 \frac{4}{1} = 2$
one	→ 2, 4	$\log_2 \frac{4}{2} = 1$
Toy	→ 1, 2, 3, 4	$\log_2 \frac{4}{4} = 0$
two	→ 2, 4	$\log_2 \frac{4}{2} = 1$

	T_1	T_2	T_3	T_4
a	1x1	0	1x1	0
and	0	1x2	0	0
beautiful	1x2	0	0	0
new	0	0	1x2	0
one	0	1x1	0	1x1
Toy	1x0	2x0	1x0	1x0
two	0	1x1	0	1x1

$$\text{sim}(T_1, T_3) = 1 \quad (\text{without normalization})$$

$$\text{sim}(T_2, T_3) = 0$$

$$\text{sim}(T_4, T_3) = 0$$

So T_1 is the most similar text to T_3 .

As far as the query = [one one toy] is concerned, we interpret it as a document and thus compute its TF-IDF vector, where the IDF is the one of the previous page.

	a	and	beautiful	new	one	toy	two
q	0	0	0	0	2x1	1x0	0

$$\text{sim}(q, T_1) = 0$$

$$\text{sim}(q, T_2) = 2$$

$$\text{sim}(q, T_3) = 0$$

$$\text{sim}(q, T_4) = 2$$

The query is more similar to T_2 and T_4 .

Exercise 2

Consider the WAND-algorithm over the following four lists

$$t_1 \rightarrow \dots, 5, 6, 7, 8, 15$$

$$t_2 \rightarrow \dots, 1, 3, 4, 5, 7, 8, 14$$

$$t_3 \rightarrow \dots, 8, 13, 15$$

$$t_4 \rightarrow \dots, 3, 4, 7, 8, 9$$

$$t_2 \rightarrow \dots, 1, 3, 4, 5, 7, \textcircled{8}, 14$$

$$t_4 \rightarrow \dots, 3, 4, 7, \textcircled{8}, 9$$

$$t_1 \rightarrow \dots, 5, 6, 7, \textcircled{8}, 15$$

$$t_3 \rightarrow \dots, \textcircled{8}, 13, 15$$

not \Rightarrow

UB

2

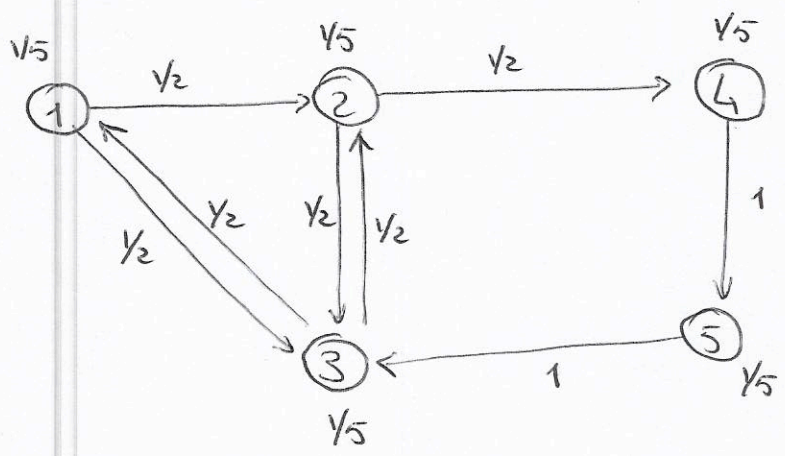
0.3

0.2

4

The threshold is set to 3.1, so the docID we choose as a pivot is the last one, hence 8, which actually occurs in all lists and thus it induces the full calculation of its score.

Exercise 3



uniform teleport
 $\alpha = 1/2$
 uniform start

node	PR
1	$\frac{1}{2} \left(\frac{1}{5} \cdot \frac{1}{2} \right) + \frac{1}{2} \cdot \frac{1}{5} = \frac{1}{20} + \frac{1}{10} = \frac{3}{20}$
2	$\frac{1}{2} \left(\frac{1}{5} \cdot \frac{1}{2} + \frac{1}{5} \cdot \frac{1}{2} \right) + \frac{1}{2} \cdot \frac{1}{5} = \frac{2}{10}$
3	$\frac{1}{2}$ $\frac{1}{2} \left(\frac{1}{5} \cdot \frac{1}{2} + \frac{1}{5} \cdot \frac{1}{2} + \frac{1}{5} \cdot 1 \right) + \frac{1}{2} \cdot \frac{1}{5} = \frac{4}{20} + \frac{1}{10} = \frac{6}{20}$
4	$\frac{1}{2} \left(\frac{1}{2} \cdot \frac{1}{5} \right) + \frac{1}{2} \cdot \frac{1}{5} = \frac{3}{20}$
5	$\frac{1}{2} \left(1 \cdot \frac{1}{5} \right) + \frac{1}{2} \cdot \frac{1}{5} = \frac{2}{10}$

The personalized PR is applied with respect to node 2, the α and the standing probability are the same as above, so I just change the additional term referring to the teleport step by putting $\frac{1}{2} \cdot 1$ to ② and \emptyset to all the others.

node	PPR
1	$\frac{1}{20} + 0 = \frac{1}{20}$
2	$\frac{1}{10} + \frac{1}{2} = \frac{6}{10}$
3	$\frac{4}{20} + 0 = \frac{4}{20}$
4	$\frac{1}{20} + 0 = \frac{1}{20}$
5	$\frac{1}{10} + 0 = \frac{1}{10}$

← This is the node most similar to ②