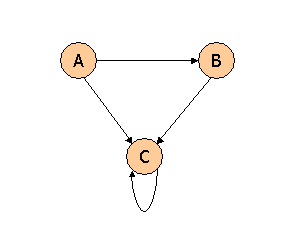
CS4225/CS5425 BIG DATA SYSTEMS FOR DATA SCIENCE

Tutorial 5: Large Graph Processing

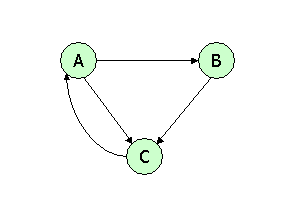
**===Part 1===**

**1.** Consider three Web pages with the following links:



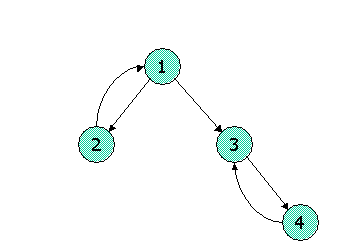
Suppose we compute PageRank with a β of 0.7, and we introduce the additional constraint that the sum of the PageRanks of the three pages must be 3, to handle the problem that otherwise any multiple of a solution will also be a solution. Compute the PageRanks a, b, and c of the three pages A, B, and C, respectively.

**2.** Consider three Web pages with the following links:



Suppose we compute PageRank with β=0.85. Write the equations for the PageRanks a, b, and c of the three pages A, B, and C, respectively.

3.  Consider the following link topology.



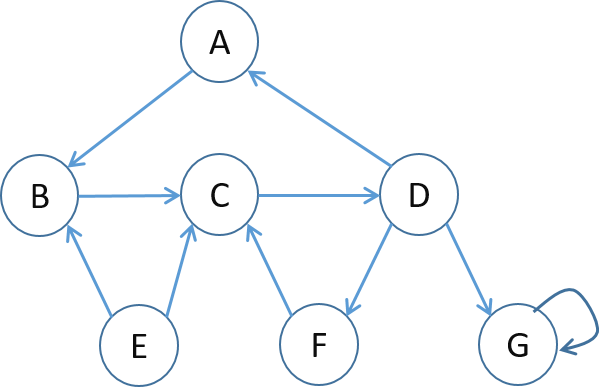
Compute the Topic-Specific PageRank for the following link topology. Assume that pages selected for the teleport set are nodes 1 and 2 and that in the teleport set, the weight assigned for node 1 is twice that of node 2. Assume further that the teleport probability, (1 - beta), is 0.3.

**===Part 2===**

1. Given the following graph,

1) how many dead ends are there in the graph? For each dead end (if any), please indicate the set of vertices forming the dead end.

2) how many spider traps are there in the graph? For each spider trap (if any), please indicate the set of vertices forming the spider trap.



2. Set up the PageRank equations for the below graph, assuming β = 0.8 (jump probability = 1−β). Denote the PageRank of node *x* by *r(x)*.

