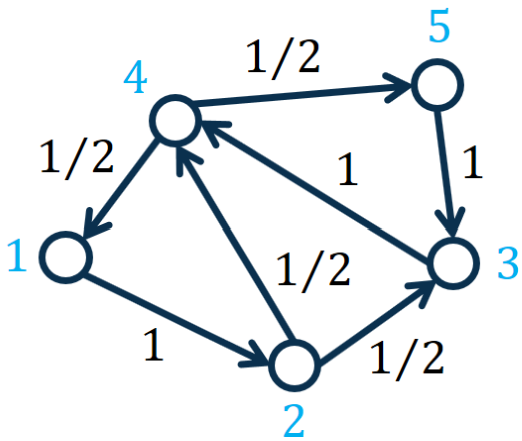
**CSC 4740/6740 Data Mining**

**Assignment 3**

**Due Date: 11:59 am, October 31, 2022**

1. (100 points) Implementing PageRank algorithm.



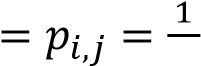
The PageRank values are given in the following table (given that the decay factor 𝑐𝑐 = 0.85):

|  |  |
| --- | --- |
| Nodes | PageRank Values |
| 1 | 0.1556 |
| 2 | 0.1622 |
| 3 | 0.2312 |
| 4 | 0.2955 |
| 5 | 0.1556 |

**PageRank:**

Compute the PageRank value of each node in the graph. Please refer to the slides for more details about the PageRank method. The key PageRank equation is as follows.

𝐫 = 𝑐𝐏⊤𝐫 + (1 −𝑐)𝟏/𝑛

where 𝐫𝐫 represents the 𝑛𝑛 × 1 PageRank vector with each element 𝐫𝑖representing the PageRank value of node 𝑖𝑖, 𝑛𝑛 represents the number of nodes in the graph, 𝐏represents the 𝑛 × 𝑛 transition probability matrix with each element 𝐏𝑖,𝑗 𝑑𝑖 representing the transition probability from node 𝑖 to node 𝑗, 𝑑𝑖 represents the degree of node 𝑖, 𝐏⊤represents the transpose of 𝐏**,** 𝑐 ∈ (0,1) represents a decay factor, 𝟏 represents a 𝑛 × 1 vector of all 1’s, and 𝑛 represents the number of nodes in the graph.

In this assignment, we set the decay factor 𝑐 = 0.85 and set the number of iterations to 30.

**Implementation**: Please implement the Power Iteration algorithm for computing PageRank in either Matlab or Python.

Or the output in your Python algorithm:

I chose Python for this assignment.

Screenshot:



Copied output:

Corresponding pagerank values for each node:

{1: 0.1556, 2: 0.1622, 3: 0.2312, 4: 0.2955, 5: 0.1556}