REPORT - Project 2.1

Finding the top leader

April 19, 2024

Name	Charvi Pahuja
Registration number	2023CSB1114

1 Dataset Generation

One day, as a class activity, students were instructed to engage in a question-and-answer session with their peers. Following the interaction, students were required to fill out a Google Form where they listed the students who impressed them during the activity. Each student could provide a maximum of 30 impressions. The responses from the Google Form were collected and organized into an Excel sheet. The first column represented the entry number of each student, while the corresponding rows contained the entry numbers of the students who impressed them.

2 Creation of network graph using data

A Python script was written to do this. Python dictionary pandas was used to read the Excel file and take input. Using the NetworkX dictionary, entry numbers from the first column were created as nodes. Taking this as i and j as the entry number in corresponding rows for all impressions, an edge (i,j) was added directed from i to j. This way, the network directed graph was created. The resulting network graph resembled a web of interconnected nodes, with each node representing a student and the edges representing the impressions shared between them. It looked like:

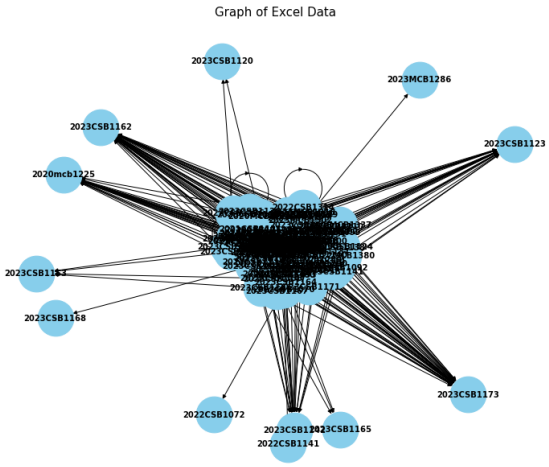

Figure 1: Network Graph

2.1 Random walk method to find PageRank

The random walk simulation begins by choosing a random node in the network and assigning 1 point to it. Randomly, another node is chosen from the nodes that our first node points to. At each step,

1 point is dropped to the node. If we arrive at a point that has no out-link, teleportation is done. It means out of all the nodes, another node is chosen at random. This is done for a lot of times. At the end, nodes are sorted in descending order based on the number of points that they have. This is the order of PageRank for the network. The first element of this list is the most famous one and is termed as the **Top Leader**. Our code prints this node.

2.2 Output of our code

Top leader of our network graph is found out to be **2023CSB1091** using the code.

```
In [1]: runfile('C:/Users/nares/OneDrive/Desktop/cs101/project2/2023CSB1114-Project 2.1.py', wdir='C:/Users/nares/
OneDrive/Desktop/cs101/project2')
TOP LEADER OF THE IMPRESSION NETWORK IS: 2023CSB1091
```

Figure 2: Output of Python script