

Random Walk with Teleportation

Hetvi Bagdai

April 23, 2024

1 Introduction

Random walk with teleportation is a simplified version of the Pagerank algorithm. It simulates a random surfer navigating through the network.

2 Problem Statement

Provided with a dataset named, which represents an impression network where each row corresponds to a node and its connections to other nodes. The task is to implement random walk algorithm to identify leader within this impression network.

3 Algorithm Explanation

3.1 Random Walk with Teleportation

Random walk with teleportation involves the following steps:

1. Initialization: Start at a random node in the graph.
2. Iteration:
 - With a certain probability (teleportation probability), jump to a random node in the graph.
 - Otherwise, move to a random neighbor of the current node.
3. Node Visits: Keep track of the number of times each node is visited during the random walk.

4 Pagerank

Pagerank is an algorithm used to measure the importance of nodes in a network, particularly in the context of web pages on the internet. It was developed by Larry Page and Sergey Brin, the founders of Google, as a way to rank web pages based on their relevance and importance.

5 Explanation of Teleportation

Teleportation, in the context of algorithms like Pagerank and random walk with teleportation, refers to a mechanism that allows the algorithm to jump to a random node in the network with a certain probability, rather than following a link to a neighboring node as dictated by the network structure.

In the context of Pagerank and random walk algorithms, teleportation serves two primary purposes:

1. **Addressing Dead Ends:** In a directed graph, dead ends are nodes with no outgoing edges. Without teleportation, a random surfer (in Pagerank) or a random walker (in random walk) could get stuck at a dead end, as there are no links to follow. Teleportation ensures that the surfer or walker can always jump to any node in the graph, including those with no outgoing edges, thus preventing the algorithm from getting stuck.
2. **Injecting Randomness:** Teleportation introduces an element of randomness into the algorithm, simulating the behavior of a random user navigating through a network. By allowing the surfer or walker to jump to a random node with a certain probability, teleportation ensures that the algorithm does not become biased towards highly connected nodes or specific paths in the graph. This randomness helps to ensure the algorithm's convergence to the true importance of nodes in the network.

In practice, teleportation is implemented by introducing a probability parameter, typically denoted as α (alpha), which represents the probability of teleporting to a random node at each step of the algorithm. The value of α is often set between 0 and 1, with common choices being around 0.15 (as used in the original Pagerank algorithm) or 0.1. When α is 0, there is no teleportation, and the algorithm follows the network structure strictly. When α is close to 1, teleportation becomes more frequent, introducing more randomness into the algorithm.

Overall, teleportation is a crucial component of algorithms like Pagerank and random walk with teleportation, enabling them to handle dead ends, inject randomness, and produce more accurate estimates of node importance in large-scale networks.

5.1 Explanation of Code

The implementation of random walk with teleportation involves the following steps:

1. **Read the CSV file:** The code reads the data from a CSV file containing the structure of the impression network.
2. **Create a directed graph:** Using the NetworkX library, the code constructs a directed graph based on the data read from the CSV file.

3. **Implement random walk with teleportation:** A function is defined to simulate a random walk with teleportation on the graph. During each step, the algorithm randomly moves to a neighboring node or teleports to a random node based on the specified probability. It keeps track of the number of visits to each node.
4. **Output the top leaders:** The code identifies and prints the top leaders using random walk with teleportation and Pagerank algorithms.

6 Conclusion

This report provides an overview of random walk with teleportation and its implementation in Python. It demonstrates how this algorithm can be used to identify important nodes in a network. The provided code offers a practical example of applying random walk with teleportation in real-world scenarios.

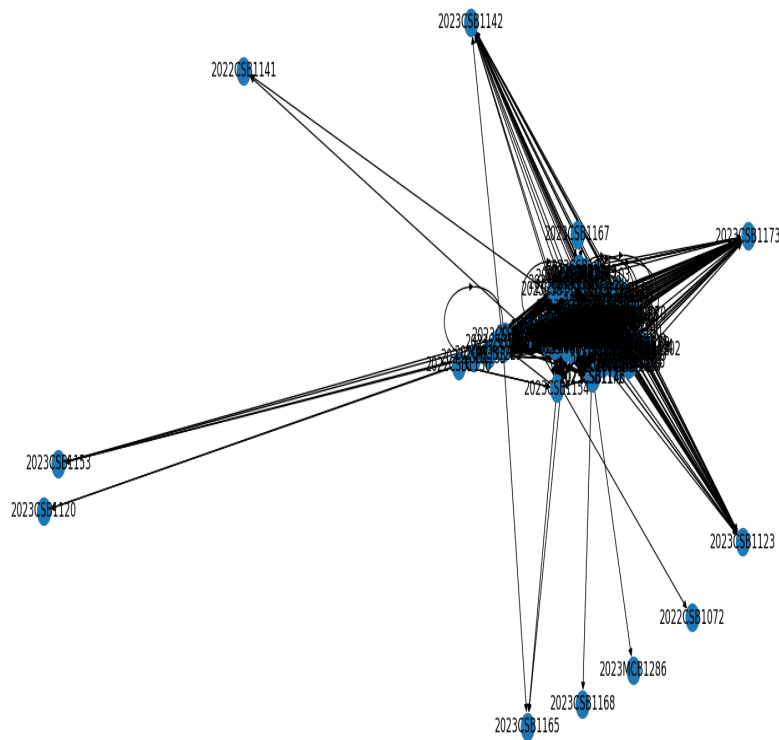


Figure 1: Network Graph