Game of Thrones Most important character Rank

Introduction

Google developed page rank to rank websites. PageRank is a Ranking algorithm that works by assigning a score to each of the websites based on the number of links that are connected to that website, the more links the higher the rank the page is going to have, and the higher the rank it means that page is more valuable. These links do not have equal weights, the more important the link the more weight it has. When we search in Google about a topic, the pages with higher ranks come to the top to facilitate the search and web surfing. The positive aspect of PageRank is that it improves the quality of search engine results, helps to deliver more relevant information, and helps with a good user experience.

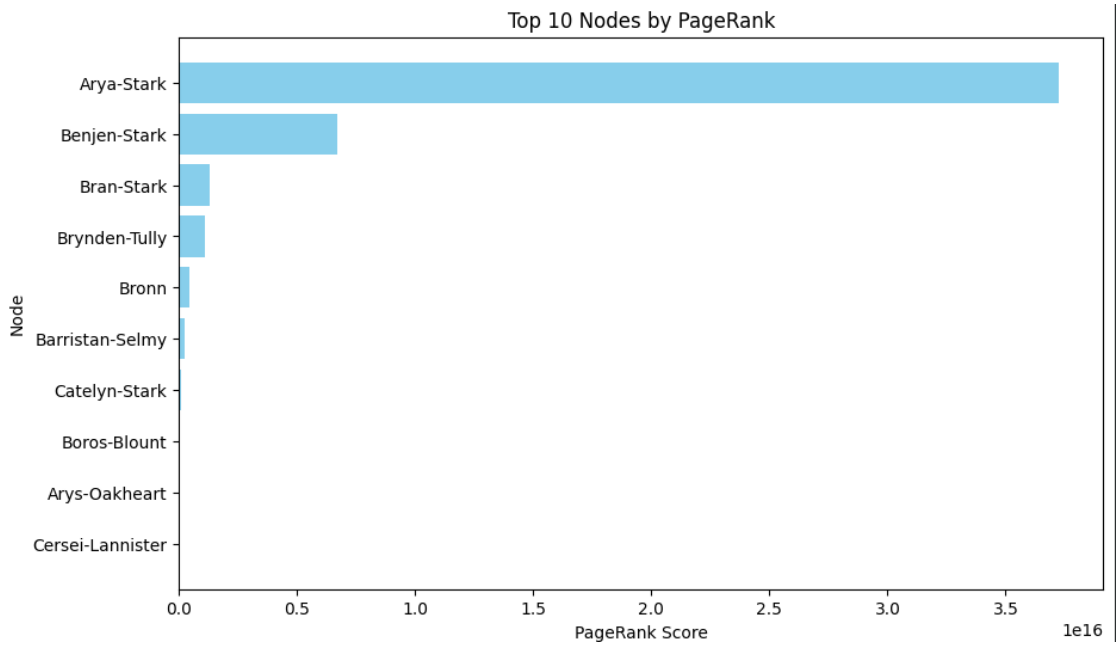
Python code implementation.

```python
import pandas as pd
import networkx as nx
import matplotlib.pyplot as plt
import numpy as np

url = "https://raw.githubusercontent.com/drolsonmi/math3280/master/Projects/GameOfThrones1.csv"
df = pd.read_csv(url)

# Create a list of unique nodes (sources and targets)
unique_sources = df['Source'].unique()
unique_targets = df['Target'].unique()
unique_nodes = np.union1d(unique_sources, unique_targets)

# Create a square transition matrix
transition_matrix = df.pivot_table(index='Source', columns='Target', values='weight', aggfunc='sum', fill_value=0)
transition_matrix = transition_matrix.reindex(index=unique_nodes, columns=unique_nodes, fill_value=0)
N = len(unique_nodes)

# Initialize vectors v and e
v = np.ones((N, 1)) / N  # Reshape to (N, 1)
e = np.ones((N, 1))

beta = 0.8

# Web surfer steps
for i in range(20):
    v = beta * np.dot(transition_matrix, v) + (1 - beta) * (e / N)
#print(v)
```

In the above code after getting the data from the link, I have evaluated the transition matrix, which represents the weighted connections between the characters of game of thrones TV show.

```python
# Sorting the Rank nodes
sorted_indices = np.argsort(v[:, 0])[::-1]  # Sort in descending order
sorted_nodes = unique_nodes[sorted_indices]
sorted_pageranks = v[sorted_indices]
sorted_pageranks
```

In the code sections PageRank algorithm is applied for calculating the importance of the characters which is done by iteratively updating the probability vector v. The beta parameter controls the probability of teleporting to any other character randomly.



The graph displays the top N Game of Thrones characters ranked by PageRank scores. Each horizontal bar represents a character, and its length indicates their centrality in the network. Characters with higher PageRank scores are positioned at the top, signifying their significance in terms of interactions with other characters.

```
# character with the highest PageRank score
max_pagerank_character = max(G.nodes, key=lambda node: G.nodes[node]['pagerank'])
max_pagerank_score = G.nodes[max_pagerank_character]['pagerank']
print(f"The most important character is '{max_pagerank_character}' with a PageRank score of {max_pagerank_score:.4f}")
```

The most important character is 'Arya-Stark' with a PageRank score of 37266328399037448.0000.

Dead Ends

A dead end is a character (node) that, after many iterations, consistently has a PageRank score of 0. This would imply that the web surfer never visits or transitions to that character, suggesting that character is not connected to the rest of the network.

There are DeadEnds present in the data. The characters with PageRank scores of 0 in the output shows that there are dead ends in our data. That shows that they do not have connection to other characters.

```
                                    0
Source
Addam-Marbrand                      7.107358e+09
Aegon-I-Targaryen                   3.600062e+12
Aemon-Targaryen-(Maester-Aemon)     3.035836e+11
Aerys-II-Targaryen                  6.751887e+12
Aggo                                6.608127e+12
...                                          ...
Will-(prologue)                     0.000000e+00
Willis-Wode                         0.000000e+00
Wyl-(guard)                         0.000000e+00
Wylla                               0.000000e+00
Yoren                               0.000000e+00

[187 rows x 1 columns]
                                    0
Source
Addam-Marbrand                      8.884198e+09
Aegon-I-Targaryen                   4.500078e+12
Aemon-Targaryen-(Maester-Aemon)     3.794795e+11
Aerys-II-Targaryen                  8.439859e+12
Aggo                                8.260159e+12
...                                          ...
Will-(prologue)                     0.000000e+00
Willis-Wode                         0.000000e+00
Wyl-(guard)                         0.000000e+00
Wylla                               0.000000e+00
Yoren                               0.000000e+00

[187 rows x 1 columns]
                                    0
```

Spider Traps

In the realm of PageRank analysis, the term "spider traps" alludes to scenarios in which the PageRank scores of nodes, or characters in our context, exhibit tendencies toward infinity (divergence) or manifest oscillatory behaviour. This phenomenon arises primarily when cycles exist within the network structure without the appropriate application of damping factors.

The fundamental objective of the PageRank algorithm is to equitably distribute importance scores across nodes. However, in the absence of damping, these scores may escalate uncontrollably in certain instances, thereby engendering instability.

The code at the right side shows just one column of data that shows that there is Spider Traps in our data.

```
Source
Addam-Marbrand                  0.0
Aegon-I-Targaryen               0.0
Aemon-Targaryen-(Maester-Aemon) 0.0
Aerys-II-Targaryen              0.0
Aggo                            0.0
...                             ...
Will-(prologue)                 0.0
Willis-Wode                     0.0
Wyl-(guard)                     0.0
Wylla                           0.0
Yoren                           0.0

[187 rows x 1 columns]
```

Conclusion

The initial purpose of Google's PageRank algorithm was to rank websites. It now improves user experiences by delivering more relevant search results and assigning scores based on link amount and quality. Through recurrent computations, PageRank evaluates character relevance within the context of the Game of Thrones character network, with 'Arya Stark' emerging as the most significant character. Dead ends and characters with PageRank scores of 0 draw attention to isolated Characters in the Game of Thrones. Additionally, the concept of spider traps warns against unstable PageRank scores in cyclic network structures. The adaptability of PageRank goes beyond web page ranking; it also offers useful insights on connectivity and significance within characters in the Game of Thrones Tv show, making it an essential tool for analysing complex data sets.

Source:

Stox, P. (2023). Google Still Uses PageRank. Here's What It Is & How They Use It. *SEO Blog by Ahrefs*. https://ahrefs.com/blog/google-pagerank/