

Experiment – 10

Title :- Implementation of Page rank/ HITS algorithm

Aim :- To Implement Page rank/ HITS algorithm

Theory :-

Hyperlink Induced Topic Search (HITS) Algorithm is a Link Analysis Algorithm that rates webpages, developed by Jon Kleinberg. This algorithm is used to the web link-structures to discover and rank the webpages relevant for a particular search.

HITS uses hubs and authorities to define a recursive relationship between webpages. Before understanding the HITS Algorithm, we first need to know about Hubs and Authorities.

- Given a query to a Search Engine, the set of highly relevant web pages are called **Roots**. They are potential **Authorities**.
- Pages that are not very relevant but point to pages in the Root are called **Hubs**. Thus, an Authority is a page that many hubs link to whereas a Hub is a page that links to many authorities.

```
# importing modules
import networkx as nx
import matplotlib.pyplot as plt

G = nx.DiGraph()

G.add_edges_from([('A', 'D'), ('B', 'C'), ('B', 'E'), ('C', 'A'),
                  ('D', 'C'), ('E', 'D'), ('E', 'B'), ('E', 'F'),
                  ('E', 'C'), ('F', 'C'), ('F', 'H'), ('G', 'A'),
                  ('G', 'C'), ('H', 'A')])

plt.figure(figsize =(10, 10))
nx.draw_networkx(G, with_labels = True)

hubs, authorities = nx.hits(G, max_iter = 50, normalized = True)
# The in-built hits function returns two dictionaries keyed by nodes
# containing hub scores and authority scores respectively.

print("Hub Scores: ", hubs)
print("Authority Scores: ", authorities)
```

Spyder (Python 3.9)

File Edit Search Source Run Debug Consoles Projects Tools View Help

/home/swpc-03/Documents/HITS(Exp10).py

temp.py X HITS(Exp10).py X

```

1 #!/usr/bin/env python3
2 # -*- coding: utf-8 -*-
3 """
4 Created on Tue Oct 11 10:17:29 2022
5
6 @author: swpc-03
7 """
8
9 # importing modules
10 import networkx as nx
11 import matplotlib.pyplot as plt
12
13 G = nx.DiGraph()
14
15 G.add_edges_from([('A', 'D'), ('B', 'C'), ('B', 'E'), ('C', 'A'),
16 ('D', 'C'), ('E', 'D'), ('E', 'B'), ('E', 'F'),
17 ('E', 'C'), ('F', 'C'), ('F', 'H'), ('G', 'A'),
18 ('G', 'C'), ('H', 'A')])
19
20 plt.figure(figsize=(10, 10))
21 nx.draw_networkx(G, with_labels=True)
22
23 hubs, authorities = nx.hits(G, max_iter=50, normalized=True)
# The in-built hits function returns two dictionaries keyed by nodes
# containing hub scores and authority scores respectively.
24
25 print("Hub Scores: ", hubs)
26 print("Authority Scores: ", authorities)
27
28
29

```

Help Variable Explorer Plots Files

Console 1/A X

```

hubs, authorities = nx.hits(G, max_iter=50, normalized=True)

```

File /home/swpc-03/Documents/HITS(Exp10).py in hits

LSP Python: ready conda: base (Python 3.9.12) Line 29, Col 1 UTF-8 LF RW Mem 83%

Conclusion :-

Thus we have successfully performed the HITS algorithm.