Report of Network Science Assignment

MADE BY

Ziyuan Ye

11610203@mail.sustc.edu.cn

UNDER THE GUIDANCE OF

Shan He

AND

Chengbin Hou

Department of Computer Science and Engineering

Southern University of Science and Technology

Shenzhen, China, July 2018

Abstract

This assignment focus on analsying the data set with the technique such as **Page Rank**, **Find Eigenvector Centrality** and **Find In and Out degree**. I try to combine the above methods to work out the top 10 rank of a data set which is called **Top Four Cities**.)) This data set seems very confusing initially. However, by using the above techniques, we gradually can find the top 10 nodes which has the highest rank in all the data set.

Data Set Source:

`/Dataset/topfourcites.txt`

Chapter 1

Introduce the PageRank Algorithm

1.1 Flowchart

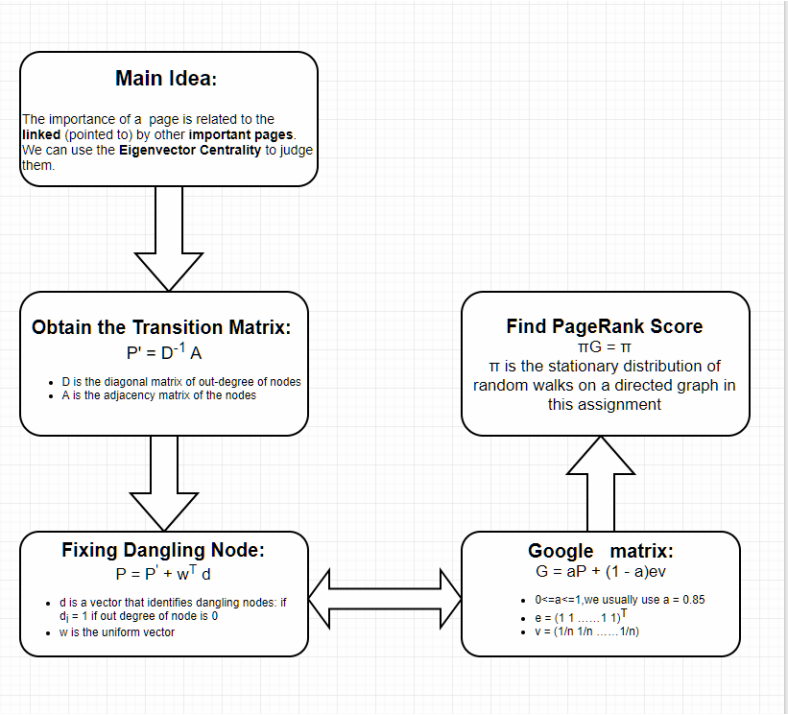

Figure: Flowchart

The flowchart cover each page rank step. It also cover the parameters and how I use select them.

1.2 Pseudo-code

Algorithm 1 Calculate *pagerank*

Require: $v = (1/n, 1/n....1/n), e = (1, 1, 1, 1.....11), W = (1/n, 1/n....1/n)$

```
d[i] is the out degree of node i
y ⇐ 1
if out degree of node = 0 then
    d[i] ⇐ 1
else
    d[i] ⇐ 0
end if
D' is the diagonal matrix of out − degree of nodes
P' ⇐ D' × A
P ⇐ P' + W^T × d
G ⇐ α × P + 1 − α × ev
π(0) = (1/n, 1/n....1/n)
while True do
    p(i + 1) ⇐ p(i) × G
    if abs(π(i + 1) − π(i)) < error then
        break
    end if
end while
return π
```

Chapter 2

Visualization

In this chapter, I visualize a tables combine with a graphs of the top ten node in each algorithm

Top ten in degree of node	
Node	value
Meyer DS 2004	8.03124821796318
Earl J 2004	6.382854503328885
Kalev A 2006	5.146559217353166
Cress DM 2000	3.498165502718872
Sampson RJ 2005	3.498165502718872
Charles CZ 2003	3.0860670740602987
King BG 2007	3.0860670740602987
Mcadam D 2001	2.880017859731012
Benford RD 2000	2.6739686454017257
Logan JR 2004	2.467919431072439

Table: Top ten in degree of node

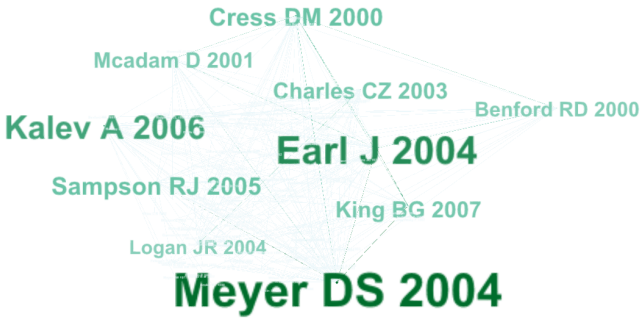

Figure In degree of top ten node

Top ten out degree of node	
Node	value
Mccarthy JD 1977	8.012879874411901
Mcadam Douglas 1982	6.008786197075345
Tilly C 1978	5.700464092869721
Wilson W J 1987	5.083819884458474
Denton Nancy A 1993	4.7754977802528495
Dimaggio PJ 1983	4.158853571841602
Meyer JW 1977	4.00469251973879
Granovetms 1973	3.850531467635978
Mccarthy JD 1996	3.2338872592247303
Gamson W 1990	3.0797262071219182

Table: Top ten out degree of node

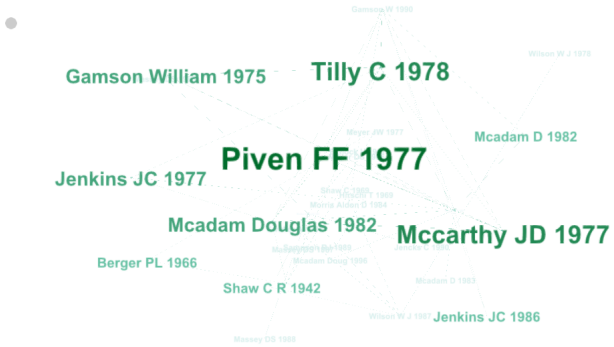

Figure Out degree of top ten node

Top ten Eigenvector centrality of node	
Node	value
Meyer DS 2004	0.2714494
Mccarthy JD 1977	0.27114936
Cress DM 2000	0.24436234
Earl J 2004	0.23883461
Tilly C 1978	0.23256826
Mccarthy JD 1996	0.222541
Mcadam Douglas 1982	0.2130869
Gamson W 1990	0.20264384
Tarrow S 1998	0.1793854
Mcadam D 2001	0.15747336

Table: Top ten Eigenvector centrality of node

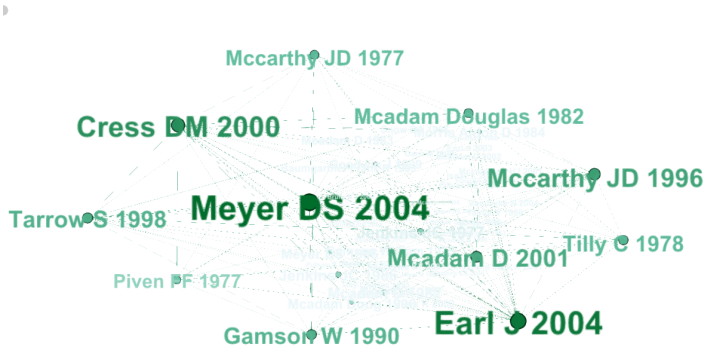

Figure Eigenvector centrality of top ten node

I try the iteration equals 100 and 200 and find the same result, which means that in at most 100 iterations the eigenvector centrality will come to converge.

Top ten Page Rank of node	
Node	value
Meyer DS 2004	0.78125451
Earl J 2004	0.31142786
Andrews KT 2004	0.17355846
King BG 2007	0.1697048
Mcadam D 2001	0.16243874
Cress DM 2000	0.15626345
Davis GF 2005	0.14311321
Mccammon HJ 2001	0.12880348
King BG 2005	0.11845306
Sampson RJ 2002	0.11700992

Table: Top ten Page Rank of node

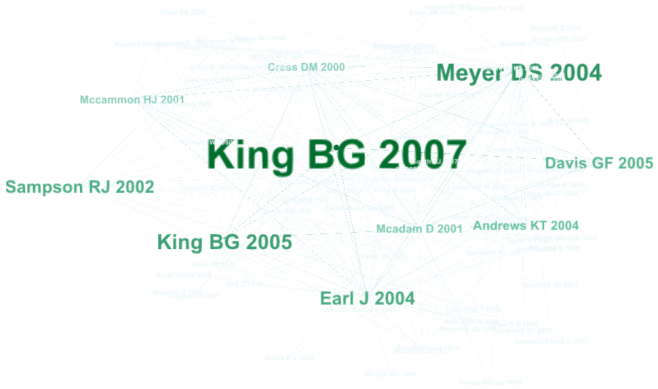

Figure Page Rank of top ten node

I try the iteration equals 100 and 200 and find the same result, which means that in at most 100 iterations the eigenvector centrality will come to converge.

Chapter 3

Conclusion

3.1 Abstract of In Degree, Out Degree and Eigenvector Centrality

As we learn, in degree means how many a node be pointed by other nodes. Out degree means how many a node pointing to other nodes. Eigenvector centrality means the importance of a node depends on the importance of its neighbors.

3.2 Comparison of In Degree, Out Degree and Eigenvector Centrality

As the table show above, we can find that top ten out degree node have numeral same nodes compare with eigenvector centrality. Thus, we can make an assumption that the higher out degree of a node is, the higher eigenvector centrality it might be. In the same way we can find that top ten out degree of nodes also similar with top ten eigenvector centrality. However the in degree and out degree rank result has few same nodes.

Finally,we can make a conclusion that the importance of a node is judge by both in degree and out degree rather than just one of them.

Chapter 4

Epilogue

Due to the limitation of the time, my result may not the same as the result called by library.

However, I cover all the requirements! Also, I try my best to approximate my result to the standard result. Fortunately, I found my result is similar to the standard one.

If it is possible, I am looking forward to be your students in the future.

Finally, I sincerely express my appreciation here to Chengbin and Shan He.

Code & Related Files:

`https://github.com/Voldet/NS_project`