International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 3 Issue 4, April - 2014

Centrality Measures in Social Networking: Study and Analysis Using NetDraw 2.138 in UCINET6

Poonam Sharma
M.Tech.(IT), 4th semesester
University School Of Information and Communication Technology
Guru Gobind Singh Indraprasth University
Sector-16 C, Dwarka, New Delhi-110075

***Abstract*— Social network is a social structure of set of actors with ties between these actors. Study of social network of nodes and ties is Social Network Analysis (SNA) where nodes are representing as points and ties as lines. There are many tools available for SNA such as visone, gephi, cytoscape, UCINET6 etc. In this paper, we experiment with TARO dataset by Hage and Harary (1983) using social network analysis and visualization tool NetDraw2.138 in UCINET6 to find the most influential node in a network based on their centrality measures . Centrality measures play an important role in the field of SNA where centrality and pre centrality measures are used to rank the relative importance of nodes and edges of a graph. Centrality is a structural characteristic of an individual in the network where high centrality individuals can be termed as leaders and low centrality individuals as peripherals. Sometimes being peripheral can be advantageous as low centrality individuals are generally assigned with less workload and overhead in an organization [www.activatenetworks.net].**

Keywords—Degree, Betweenness, Closeness, Harmonic closeness, Eigen vector, 2 local Eigen vector,

I. INTRODUCTION

Now a day social network has been a significant topic for analysis and visualization in social science research. In recent years, there is a rapid growth in WWW and internet world due to which many social networking sites like Facebook, Twitter, ibibo, LinkedIn and many more came in picture. Social networking has become one of the most important parts of our daily life as it enables us to communicate with a lot of people. Social networking sites are created to assist in online networking. These sites are generally communities created to support a common theme [http://ciemcal.org/importance-of-social-networking-sites/].

Social networks are used to reflect real world relationship that allows users to share information and form connections between each other. Facebook is the best example of social networking sites. In January 2014, 1.23 billion users were active on the website every month, while on December 31, 2013, 945 million of this total was identified by the company as mobile users. The company celebrates its tenth anniversary in the week beginning February 3, 2014 [http://en.wikipedia.org/wiki/Facebook].

A social network is typically represented as a structural graph with individual persons is represented as vertices, relationship between pair of individuals is represented as edges and the strength between them is represented as weights on edges [4].

Centrality is an important concept in social network analysis (SNA) as it determines the position of an individual in social network i.e. an individual with high centrality measures is more important than the one with low centrality measures. An individual with higher centrality score may have more workload than the peripherals and most of the time peripherals are in advantage of this thing. All the centrality measures defined till now are shown in Fig.1.

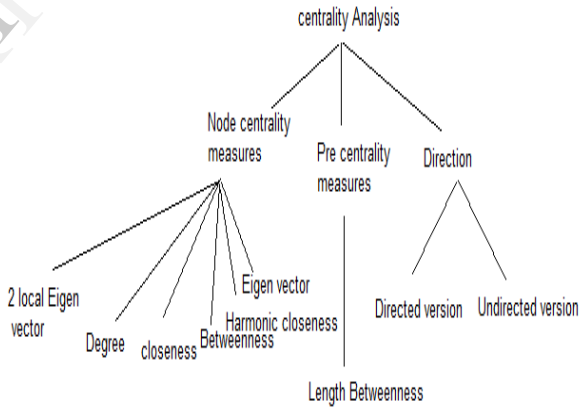

Fig.1: Classification tree of centrality analysis using NetDraw2.138 in UCINET6

II. TERMINOLOGIES AND REPRESENTATION

Degree centrality: Degree centrality is total number of ties to a node and the node with higher number of ties will have higher Degree centrality. Let us consider, the degree of a node Vi represents the number of words that co-occur with the word corresponding to Vi and N(Vi) be the set of nodes connected to Vi, the degree centrality of a node Vi is given by [3]:

$$C_D(V_i) = \frac{|\mathcal{N}(V_i)|}{|V| - 1}$$

Betweenness centrality: Betweenness centrality is a measure of importance of a node that assesses the shortest path between all node pairs that pass through the node of interest.

The betweenness centrality for a node of interest v is denoted by [5]:

$$C_B(v) = \sum_{\substack{s \neq v \neq t \\ s,v,t \in V}} \frac{\rho_{st}(v)}{\rho_{st}}$$

Where, ρ_{st} : number of shortest paths from s to t.
$\rho_{st}(v)$: number of shortest paths from s to t that pass through the node v.

Closeness centrality: Closeness centrality is inverse of fairness that means it is some of shortest distances between a node and all other nodes. Let $\mathrm{distance}(V_i, V_j)$ be the shortest distance between Vi and Vj [3].

$$C_C(V_i) = \frac{|V| - 1}{\sum_{V_j \in V} \mathrm{distance}(V_i, V_j)}$$

Eigen vector: A high Eigen vector centrality means a node is most visited while traversing and is well connected [Giovanni Scardoni et. al]. Eigen vector centrality of a node 'u' can be defined as linear combination of Eigen vector centralities of its neighbours [6].

$$C_E(u) = \frac{1}{\lambda} \sum_{v=1}^{|V|} w_{u,v} C_E(U)$$

Where, λ : a constant
C_E : Eigen vector
W : network matrix

Harmonic Measure: It is an alternate method to measure the closeness that takes in account all the pathways which somehow connects a node to all other nodes rather than just geodesic[7].

Centrality measures	Complexity
Betweenness	O(VE)
Closeness	O(V(log V)E)
Degree	$O(V^2)$
Eigen vector	$O(V^3)$
Harmonic	O(V(log V)E)

Table 1: complexities associated with centrality measures[1]

NETWORK ILLUSTRATION:
We have used one dataset in this paper named as TARO which was first used by Hage and Harary (1983). It represents the relation of giving gifts among 22 households in Papuan village and it contains one square matric of size 22, symmetric and binary.

The network diagram for TARO dataset in NetDraw is shown in Figure 2 below and we select centrality measures for making analysis:

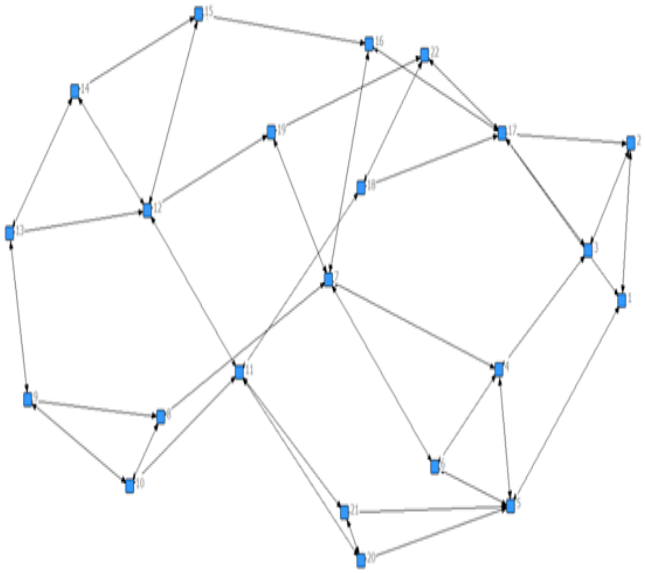

Fig. 2: Network digram for TARO dataset using NetDraw

After selecting centrality measures we set node shapes of network diagram by degree and then click on ok tab to select it.

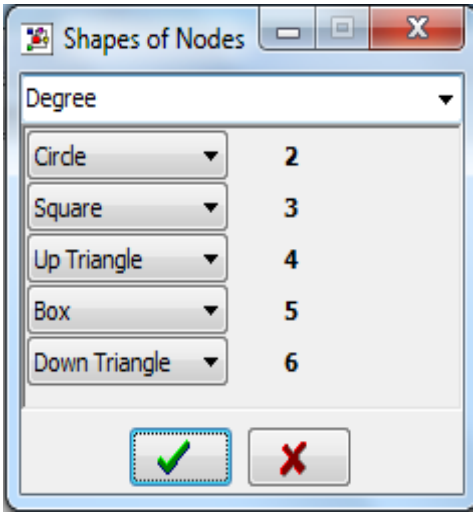

Fig. 3 Shapes of node by their degree value

In figure 4 we set symbol color by selecting the attribute degree and highest degree node is shown by green color.

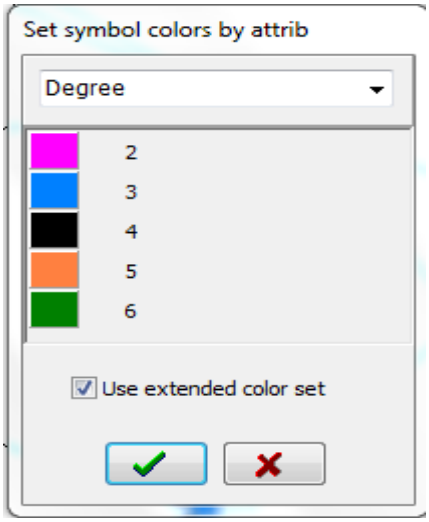

Fig.4: Symbol colors by degree attributes

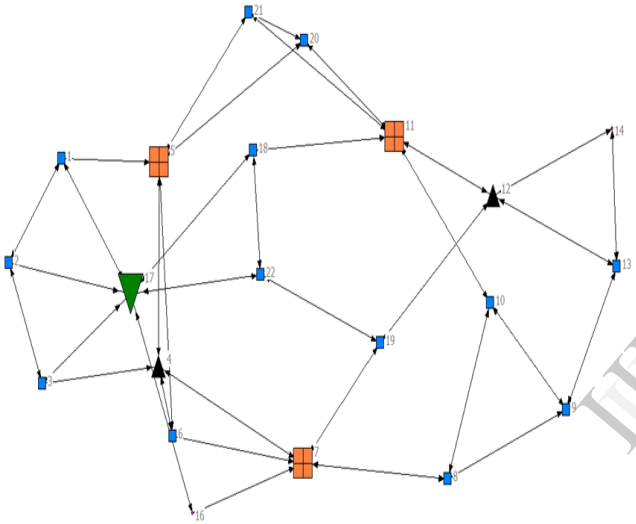

Fig.5: Network diagram after a shape and color is set

Here in figure 5 a network diagram is shown where nodes are sized and shaped by their degree and the size of node 17 is largest in comparison to all other nodes in network diagram as this node is having highest degree value 6 which is colored as green and shaped as down triangle by using NetDraw visualization tool.

OBSERVATION AND RESULT:

Table 2 shows the centrality measures values for all 22 nodes individually such as degree, betweenness, closeness, harmonic closeness, Eigen vector, 2 local Eigen vector.

	A	A	B	C	D	E
1	ID	Degree	Betweenne	Closeness	HarmonicCl	Eigenvector
2	1	3	8.267	57.000	9.733	0.224
3	2	3	0.500	60.000	9.233	0.206
4	3	3	5.833	56.000	9.700	0.216
5	4	4	15.433	49.000	10.833	0.269
6	5	5	23.600	52.000	10.833	0.300
7	6	3	4.517	53.000	9.917	0.224
8	7	5	41.450	45.000	11.667	0.271
9	8	3	15.167	53.000	9.833	0.136
10	9	3	5.500	61.000	9.067	0.105
11	10	3	11.000	54.000	9.750	0.134
12	11	5	46.383	44.000	11.833	0.261
13	12	5	34.133	48.000	11.417	0.228
14	13	3	5.667	62.000	9.100	0.124
15	14	3	2.333	59.000	9.167	0.134
16	15	3	12.183	53.000	9.750	0.150
17	16	3	23.550	48.000	10.667	0.201
18	17	6	38.183	47.000	11.833	0.334
19	18	3	15.350	49.000	10.417	0.210
20	19	3	17.533	47.000	10.667	0.185
21	20	3	5.650	52.000	9.917	0.204
22	21	3	5.650	52.000	9.917	0.204
23	22	3	7.117	51.000	10.083	0.194

Table 2: centrality measures for all the nodes in TARO network

After looking at all above figures and tables, we made an observation Table 3 for highest centrality measures as the node with highest centrality measures is always considered as an important node in social network diagram.

Measures	Highest Value	Node ID
Degree	6	17
Betweenness	46.383	11
Closeness	62	13
HarmonicCl	11.833	11,17
Eigen vector	0.334	17
2-local	18	17

Table 3: Observation Table

In this paper, we have analyzed the dataset by using weak ties between nodes in social network diagram rather than strong ties as strong ties are always bidirectional in nature. The most influential node in the social network can be the one that belongs to the largest number of communities which is known as “Community Bridge”. Here we see that node 17 is most repeated node and has highest centrality measures. Hence node 17 is most prominent or influential node which is acting as a bridge node and second important node is node 11.

REFERENCES

1. George Nomikos, Panagiotis Pantazopoulos, Merkourios Karaliopoulos, Ioannis Stavrakakis. "The Multiple Instances of Node Centrality and their Implications on the Vulnerability of ISP Networks". 2013.
2. Giovanni Scardoni, Carlo Laudanna, Gabriele Tosadori, Franco Fabbri, Mohammed Faizaan. "CentiScaPe: Network centralities for Cytoscape".
3. Florian Boudin. "A Comparison of Centrality Measures for Graph-Based Key phrase Extraction". International Joint Conference on Natural Language Processing, pages 834–838, Nagoya, Japan. 2013.
4. Kazuya Okamoto, Wei Chen, and Xiang-Yang Li. "Ranking of Closeness Centrality for Large-Scale Social Networks". Illinois Institute of Technology and Microsoft Research Asia.
5. Woochang Hwangy, Youngrae Choy, Aidong Zhangy and Murali Ramanathan. "Bridging Centrality: Identifying Bridging Nodes In Scalefree Networks". 2006.
6. Wei Wei, Jürgen Pfeffer, Jeffrey Reminga, and Kathleen M. Carley. "Handling Weighted, Asymmetric, Self-Looped, and Disconnected Networks in ORA". CMU-ISR-11-113. 2011.
7. Ekta Nankani, Simeon Simoff. "Predictive analytics that takes in account network relations: A case study of research data of a contemporary university". Proc. of the 8th Australasian Data Mining Conference (AusDM'09)