Degree Centrality, Betweenness Centrality, and Closeness Centrality in Social Network

Junlong Zhang[1,*] and Yu Luo[2]
[1]School of Business, Macau University of Science and Technology, Macau, China
[2]School of Public Administration, South China Normal University, Guangzhou, China
*Corresponding author

Abstract—**Social network theory is becoming more and more significant in social science, and the centrality measure is underlying this burgeoning theory. In perspective of social network, individuals, organizations, companies etc. are like nodes in the network, and centrality is used to measure these nodes' power, activity, communication convenience and so on. Meanwhile, degree centrality, betweenness centrality and closeness centrality are the popular detailed measurements. This paper presents these 3 centrality in-depth, from principle to algorithm, and prospect good in the future use.**

Keywords- social network; centrality; degree centrality; betweenness centrality; closeness centrality

I. INTRODUCTION

Among the individuals and individuals, organizations and organizations, individuals organizations there are a common set of connections, and the social scientists call these connections 'network' [1]. And one social network is a structure constituted by a set of 'nodes' connected with different 'ties' [2]. In 70's of 20th century, *The Strength of Weak Ties* from Granovetter was published on *American Journal of Sociology* [3], henceforth the theory of social network has been growing more and more mature. On the perspective of social network, 'nodes' can be many kinds of social actors, such as individuals, groups, companies, organizations. Meanwhile 'ties' can also represent different connections and relations. Under this paradigm, more and more social actors and relations are analyzed [4]. And according to a hundred years' big data analytics, social network gets fast-growing attentions in recent 20 years in social science [5]. The research contents of social network include "strong/weak ties", "social capital", "egocentric/whole network" and "centrality" etc. We use "social network" as searching keywords on Web of Science, to analyze the numbers of related publications in last 30 years from 5 top journals in management, including *Academy of Management Journal*(AMJ), *Academy of Management Review*(AMR), *Administrative Science Quarterly*(ASQ), *Journal of Management*(JM), *Strategic Management Journal*(SMJ). From this statistical result, we can see that the publications are increasing obviously (FIGURE I.). Social network is becoming increasingly popular in analysis [6].

Data in social science basically has three types - ideational data, attribute data and relational data [7]. Ideational Data usually makes motivations, meanings and definitions into conceptions, and typological analysis from Max Weber is a

good example [8]. Attribute data is about the index like attitude, characteristics, choices and preferences from the units of social actors (individuals or groups or organizations) [9]. The well-known correlation analysis and regression analysis use attribute data. And unlike attribute data, relational data doesn't describe single unit, furthermore it describe relations-among several or numerous social actors, from static to dynamic condition[10], also what kind of relational data depend on the what researchers focus on [11]. After getting social network data, we could use it to calculation, explanation and simulation. Centrality, the data on nodes' structural dimension, is one of most used and important [12].
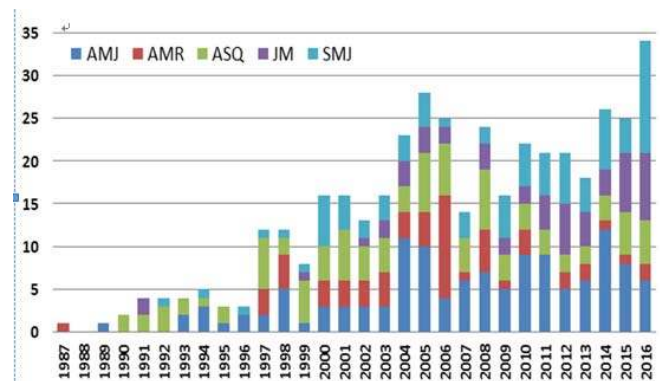

FIGURE I. PUBLICATIONS ABOUT SOCIAL NETWORK IN 30 YEARS

II. CENTRALITY

Centrality is such an important index because it indicates which node takes up critical position in one whole network. Central positions always get equated with remarkable leadership, good popularity or excellent reputation in the network [13,14]. As soon as the social actor gets a higher centrality, it mean he/she gets closer to the center of network, that higher power, influence, convenience from the network he/she may acquire [15,16]. Perhaps most typically choose degree centrality measures are degree centrality, betweenness centrality and closeness centrality. Freeman firstly probed into these 3 centrality in 1970's [17,18,19]. For a better understanding and visualization, we take an undirected single-line network graph as an example to interpret (FIGURE II.).
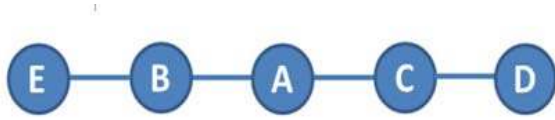
FIGURE II. EXAMPLE OF A NETWORK GRAPH

A. Degree Centrality

In a network graph, degree centrality is measured by the total amount of direct links with the other nodes, the fundamental formula Cd is equation (1) [20,21]. Since as time goes by, the size of the network may vary, to decrease this possible size effect to degree centrality measurement, Wasserman etc. make equation (1) standardized and put forward to equation (2) $C'd$ [22]. $\sum_{j=1}^{n} Xij$ means the number of links directly connectd with node N, and *n* means the total number of the nodes in focal network. We take a network graph like FIGURE II. as an example, the number of direct links connected with point A obviously is 2, in other words A point's Cd is 2, after standardization $C'd$ is approximately equal to 0.167. In the same way we can calculate Cd and $C'd$ of point B, point C, point D, point E. See the results on FIGURE III.

$$Cd(Ni) = \sum_{j=1}^{n} Xij \ (i \neq j) \qquad (1)$$

$$C'd(Ni) = \frac{\sum_{j=1}^{n} Xij}{(n-1)(n-2)} \ (i \neq j) \qquad (2)$$

Cd	1	2	2	2	1
$C'd$	0.083	0.167	0.167	0.167	0.083

FIGURE III. CALCULATIONS ON DEGREE CENTRALITY.

B. Betweenness Centrality

Betweenness centrality is to measure one node undertaking 'mediation' role in a network. If one node locates in the only way which others nodes have to go through, such as communication, connection, transportation or transaction, then this node should be important and very likely have a high betweenness centrality [17]. Fundamental formula Cb is equation (3), Wasserman etc. standardize Cb by dividing $\frac{(n-1)(n-2)}{2}$, and standardized formula $C'd$ see equation (4).

In these two equations, $\sum_{j<k} \frac{Gjk(Ni)}{Gjk}$ means the number of node N locate between any other two nodes in the network [23]. On shortest paths of all connections in a network, the higher the betweenness centrality the more frequent node N falls between any pairs of other nodes. Back to the network graph FIGURE II, A stands in the middle of E and C, E and D, B and D, B and C, so according to equation (3) for point A $Cb = 4$, $C'd = 0.53$. The whole results are on FIGURE IV.

$$Cb(Ni) = \sum_{j<k} \frac{Gjk(Ni)}{Gjk} \qquad (3)$$

$$C'b(Ni) = \frac{2\sum_{j<k} \frac{Gjk(Ni)}{Gjk}}{(n-1)(n-2)} \qquad (4)$$

Cb	0	3	4	3	0
$C'b$	0	0.4	0.53	0.4	0

FIGURE IV. CALCULATIONS ON BETWEENNESS CENTRALITY

C. Closeness Centrality

Closeness centrality, is meant to measure one node to the others nodes' sum distances, if the length of node N's shortest paths with other nodes in the network is small, then node N has a high closeness centrality [23,24]. It stands for the convenience and ease of connections between the focused node and the other nodes. The fundamental formula Cc is equation (5), $\sum_{j=a}^{n} d(Ni, Nj)$ means the total number of "steps" from node N to the other nodes in the network [17,18]. After putting closeness centrality on social actors in practice, Beauchamp suggests multiply Cc by *(n-1)*, so that we could get standardized equation $C'c$ [25]. Still in the sample graph FIGURE II, A stand both next to B and C, so the distance to B and C is the same- 1 "step". On the other hand, we can see that A stands 2 "step" away from E and D. So to node A, $Cc(Ni) = \frac{1}{1+1+2+2} = \frac{1}{6} \approx 0.167$, $C'c \approx 0.67$. Other nodes' results we can get on FIGURE V.

$$Cc(Ni) = \frac{1}{\left[\sum_{j=a}^{n} d(Ni, Nj)\right]} \ (i \neq j) \qquad (5)$$

$$C'c(Ni) = \frac{n-1}{\left[\sum_{j=a}^{n} d(Ni, Nj)\right]} \quad (i \neq j) \quad (6)$$

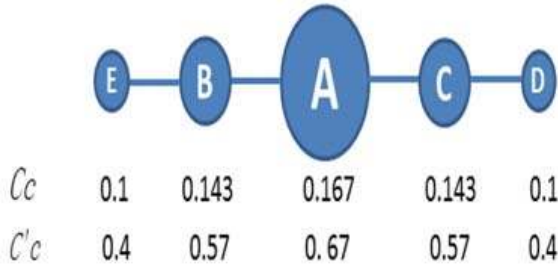|	E	B	A	C	D
Cc	0.1	0.143	0.167	0.143	0.1
$C'c$	0.4	0.57	0.67	0.57	0.4

FIGURE V. CALCULATIONS ON CLOSENESS CENTRALITY.

III. CROSS-OVER STUDY OF 3 CENTRALITY

Degree, betweenness and closeness these 3 centrality measures are derived from the adjacency matrix and so constitute some different algorithm on the same underlying data. Thus the questions like how to distinguish these 3 measures or how to use exactly are easily generated. Although the conceptual differences exist among these 3 main centrality measures, from calculations we can see these measures do not differ that dramatically. Some scholars have tried different centrality measures to seize network centrality, and found betweenness centrality and degree centrality have positive effect on investment [17]. Meanwhile, some studies examined correlations among degree, betweenness and closeness, also examined these relations in circumstances of systematic error, random error, and incomplete statistics [18]. Overall, Freeman recommend that we could use degree centrality to measure activity of transferring and communication, betweenness centrality to calculate mediation or control of interest, and closeness centrality to estimate level of efficiency and convenience [19,20]. Relatively, degree centrality and betweenness centrality have been used more frequently. Some cross-over studies suggest that in some condition, 3 centrality measures have an interesting associated correlation (TABLE I.).

TABLE I. DEGREE BETWEENNESS AND CLOSENESS

	High Degree	High Betweenness	High Closeness
Low Degree		Inndividual's few ties are crucial for network flow	Individual ties to some active/important actors
Low Betweenness	Individual's connections are redundat-communication bypassed him/her		Maybe network exist many paths , individual are near to many actors but so are many others
Low Closeness	Embedded in cluster that is away from the rest of network	Individual monopolizes the ties from a bit of people to many other. Relatively rare.	

IV. CONCLUSION

The analyses from attribute data for social actors (individuals, organizations etc.) have been well-known and widely used, such as regression analysis and correlation analysis. Social network, on the contrary, focus more on the social actors' relation data in the network. These relations construct various network structures, positions, clusters and transfer, from which we can better research on phenomena and circumstances, from a general election to transnational financial corporations. Originally, there are three things to form social network at once: phenomena, contexts for theory building and testing and analytical tools [12]. So it's not hard to speculate there is a great researching possibility in it. Top journals like Nature, Science reckon social network a promising future [26,27,28]. Indeed there are an optimistic growing amount of research productions in recent years, and we can see its research trends is more and more interdisciplinary.

REFERENCES

[1] Anthony Giddens, Sociology(6th)[M]. Polity Press, Cambridge, 2009.

[2] Granovetter, Mark S. "The Strength of Weak Ties." American Journal of Sociology 78.Volume 78, Number 6(1973):347-367.

[3] Brass, Daniel J., et al. "TAKING STOCK OF NETWORKS AND ORGANIZATIONS: A MULTILEVEL PERSPECTIVE." Academy of Management Journal 47.6(2004):795-817.

[4] Contractor, Noshir S., and K. Faust. "Testing Multitheoretical, Multilevel Hypotheses about Organizational Networks: An Analytic Framework and Empirical Example." Academy of Management Review 31.3(2006):681-703.

[5] CHEN Y S. The trajectory of sociology over two centuries: a cultural study using millions of books[J]. Sociological Studies, 2015, 1: 23-48.

[6] Borgatti, Stephen P., and D. S. Halgin. "On Network Theory." Organization Science 22.5(2011):1168-1181.

[7] Sensitivity analysis[M]. New York: Wiley, 2000.

[8] Kalberg, Stephen. "Max Weber's Types of Rationality: Cornerstones for the Analysis of Rationalization Processes in History." American Journal of Sociology 85.Volume 85, Number 5(1980):1145-1179.

[9] Knoke David, H. James, Kuklinski, Network Analysis[M]. Beverly Hills: Sage Publication, 1982.

[10] Liu Jun, Lectures on Whole Network Approach-Apractical Guide to UCINET(2nd)[M]. Shanghai: Truth & Wisdom Press, 2014.

[11] Knoke D, Yang S. Social network analysis[M]. Sage, 2008.

[12] Li, Mingxiang. "Social network and social capital in leadership and management research: A review of causal methods." Leadership Quarterly 24.5(2013):638–665.

[13] Jar-Der Luo, Social Network Analysis(2nd)[M]. Beijing: Social Science Academic Press(China), 2010.

[14] Krackhardt D. Brass D. Interorganizational Networks: The Micro Side. Ins. Wasserman and Galaskiewicz, Advances in Social Network Analysis[M]. Thousand Oaks, California: Sage, 1994.

[15] Sparrowe, Raymond T., R. C. Liden, and S. J. Wayne. "Social Networks And The Performance Of Individuals And Groups." Academy of Management Journal 44.2(2001):316-325.

[16] Hochberg, Yael V, A. Ljungqvist, and Y. Lu. "Whom You Know Matters: Venture Capital Networks and Investment Performance." The Journal of Finance. 2007(52):251–301.

[17] Freeman, Linton C. "A set of measures of centrality based on betweenness." Sociometry 40.1(1977):35-41.

[18] Freeman, Linton C. "Centrality in social networks conceptual clarification." Social Networks 1.3(1978):215-239.

[19] Freeman, Linton C., D. Roeder, and R. R. Mulholland. "Centrality in Social Networks: ii. Experimental Results." Social Networks 2.2(1979):119-141.

[20] Bolland, John M. "Sorting out centrality: An analysis of the performance of four centrality models in real and simulated networks." Social Networks 10.3(1988):233-253.

[21] Marvin E. Shaw. "Group Structure and the Behavior of Individuals in Small Groups." The Journal of Psychology 38.1(1954):139-149.

[22] Nieminen, J. "On the centrality in a graph. " Scandinavian Journal of Psychology 15.1(1974):332-336.

[23] Wasserman S., Faust K., Social network analysis: Methods and applications[M]. New York: Cambridge Univesity Press, 1994.

[24] Beauchamp, Murray A. "An improved index of centrality." Behavioral Science 10.2(1965):161-3.

[25] Sabidussi, Gert. "The centrality index of a graph." Psychometrika 31.4(1966):581-603.

[26] Palla, Gergely, et al. "Uncovering the overlapping community structure of complex networks in nature and society. " Nature 435.7043(2005):814-8.

[27] Borgatti, S. P., et al. "Network analysis in the social sciences. " Science 323.5916(2009):892-5.

[28] Barabási, A. L. "Scale-free networks: a decade and beyond. " Science 325.5939(2009):412-3.