Summer 6-1-2014

CentralityDifferent Influence Models of Node Centrality in Transactional Community

Minxue Huang
Economics and Management School, Wuhan University, China

Peixiang Sun
Economics and Management School, Wuhan University, China

Bangming Xiao
Economics and Management School, Wuhan University, China

Different Influence Models of Node Centrality in Transactional Community

Minxue Huang, Peixiang Sun, Bangming Xiao
Economics and Management School, Wuhan University, China

Abstract: This study investigates the various influence models of nodes' network centrality in the context of transactional community. Combining the Social Network Analysis (SNA) with Tobit regression, the research indicates that: i) a node's degree centrality (its followers) and betweenness centrality (the number of the shortest paths in which the node is included) have a positive impact on its network influence, ii) the closeness centrality (physical closeness in network) shows no significant impact on its influence. Theoretically, the results provide insight into the sources of influence and various influence models of different node centralities in transactional community network. In practice, influentials can be better identified according to different centralities, so as to distinguish the opinion leaders in a more accurate way.

Key words: transactional community, node centrality, network influence, opinion leaders

1. INTRODUCTION

With the development of science technology and social economy, the Internet has stepped into the Web2.0 era [1]. Internet has become an important part of people's life, with more and more people involved with activities such as online shopping, social networking and so on. Many companies conduct their social marketing via network and the key is to identify the influentials in the network. Many scholars have tried to resolve this issue by studying the influence of different network centrality. However, most of them studied it in a setting of static social network, which is formed through social ties among members. But for transactional community, it is a sparse and dynamic network, which is formed through members' transactional activity with each other. In other words, the members in the sparse transactional community are connected via similar products purchasing or the same products participating experience. Here our research questions are how the influence of central nodes is built and what is it's difference from the typical social network? As far as we know, very few literatures focused on this setting of transactional network. Given the importance of network centrality in network, we distinguished the centrality into three types: degree centrality, closeness centrality, and betweenness centrality [2]. This study tries to explore the different influence of three types of centrality in the setting of transactional network.

We organize the rest parts of paper as the follow: We review the related literature and propose the hypotheses. We collect the data of members' connections (following and follower) and activities (posting, replies, etc.). Practically, this study relies on China's largest online trading platform – Taobao.com, and selects one of the most established community modules - Taobao Gang. We analyzed the relationship networks by doing Social Network Analysis and investigated the difference of centralities with the Tobit regression. The results shows that different types of centralities exist between social network and transactional network.

2. RESEARCH FRAMEWORK AND HYPOTHESES

2.1 Transactional Community

Social network analysis in management has flourished over the past few decades, especially in group network and group-oriented analysis [3]. These researches focused on the relationships between members rather than their attributes, and concerned interactive structural patterns but not isolated individuals [3]. Network is defined as the set of nodes (they can be individuals, organizations and institutions), the connections between the

nodes, and the connections between nodes standing for relationships in networks.

Transactional community is a common form of virtual community. Based on the existing typologies of online communities and consideration of participants' needs, the following types of communities can be identified: discussion or conversation communities, task- and goal-oriented communities, virtual community and hybrid communities [5]. Other scholars divide the communities into interest communities, interpersonal communities, imaginary communities and transactional communities based on user demands [6].

Transactional community is different from general social communities, where virtual relationship comes from members' embeddedness in social relations and it is easily affected by emotional factors and normative social influence [7, 8]. While in transactional community, members are more likely to be affected by informational social influence because they are strangers to each other before joining the community [8].

2.2 Network influence

2.2.1 Definition of influence

Essentially, influence in virtual network is a kind of social influence. Deutsch et al. [8] divided social influence into normative social influence and informative social influence. We shall define a normative social influence as an influence to conform to people's positive expectation. An informational social influence may be defined as an influence to accept information obtained from another member as evidence about reality. Studies on network influence in virtual community are time-honored, mainly focusing on opinion leaders and new product diffusion, especially opinion leaders and hubs in network. Kumar et al. [9] raised customer influence effect and customer influence value to illustrate individual influence in social media and social network. Similarly, this paper classified influence into two aspects: "influence scope (i.e. page views)" and "influence depth (i.e. number of replies)". Page views reflect the coverage and range of members' influence in network. Number of replies embodies the interaction between the members and their audiences, reflecting information spreading depth.

2.2.2 Source of influence

Participants in competition network are composed of self and contacts which give them competitive edges, resulting into a higher ROI (return on investment) [10]. Network brings information benefits in the form of access, timing and referral. Other things equal, individuals with gigantic and diversified network can guarantee useful information to flow more smoothly between their contacts and themselves. Heterogeneous networks can bring about more benefits than homogeneous networks, and heterogeneous social locations can provide opportunities via information and limitations through control. People connected to various groups and factions have advantage of getting a wealth of information because their vision is broader than those with limited networks.

To sum up, network structure, network relationships and location of members offer different information, control and power of members, and based on these structures, relationships and locations the members can exert their influence to other members.

2.3 The different influence patterns of members' centralities

Sociologists have extensively studied how the network structure affects social influence, albeit in a small social network. Krackhardt [11] confirmed that not only the number of connected nodes, but also the way they are embedded should be considered when evaluating the influence power of individuals. Coleman [12] pointed out that when two connected individuals are connected to the third party, the information dissemination is more effective and the relationship becomes more solid. Burt [13] defined this as "network closure", namely the third party creating a redundant path for information flow, thus strengthening mutual trust between the two relevant actors. In addition, as a "broker", individuals interconnecting separate clusters are more influential, because they get a greater control over information from other clusters. Members' network location, such as the number of neighbors, neighbors interconnected level, and brokerage, is also a good predictor of influence and adoption

tendency [14].

This study investigates the different impact of centralities on members' influence, which has never been fully studied before, particularly in the transactional community. Specifically we use degree centrality, closeness centrality, and betweenness centrality to measure and reflect members' centrality. Thus we propose the theoretical framework in this paper here, as below:

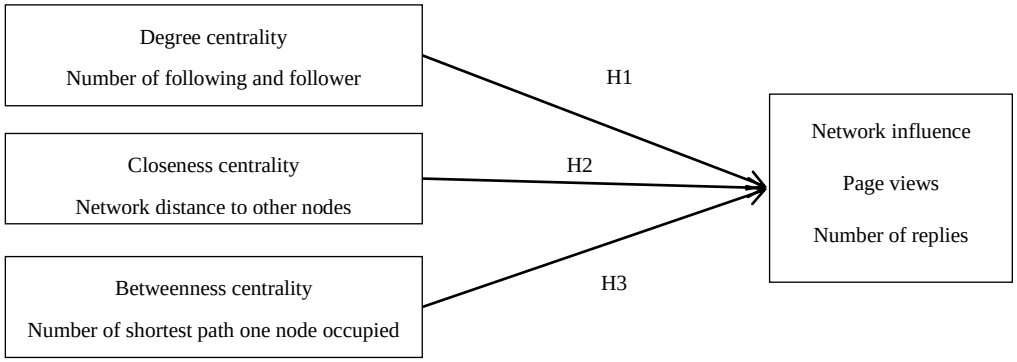

Figure 1. The research framework

The framework includes three different impact of centralities with three corresponding hypothesis. We will discuss them in the next section.

2.3.1 Members' degree centrality - following and follower

Degree centrality is the first-proposed and also the easiest centrality concept, which is defined as the number of connected nodes. Nodes with a high degree centrality are defined as network Hubs. In the virtual social network, degree centrality reflects the connectedness strength in local network, and the possibility of information flow is positively related to tie strength. In addition, information influence in the network is a two-way transmission, that is, information from other sides will generate an information flow to you, conversely, it also guide, aggregate, legalize information from you, and pass it on to others [10].

Therefore, nodes with a high degree centrality can have more information paths to exert influence on their connected neighbors. Based on this, we propose the first hypothesis:

H1: In transactional community, the higher the degree centrality is, the greater the network influence power will be.

2.3.2 Members 'closeness centrality-network distance

Node's farness is defined as its average distance to all other nodes, and closeness is defined as the inverse of farness [15], which measures the distance from one node to another. It reflects the reachability between nodes, and is inversely proportional to the average length of the path.

Newman[16]proposed that closeness can measure the time it takes for information to be transmitted sequentially from one node to all other nodes; Trier [17]pointed out that closeness reflects the extent to which members can access the network quickly and efficiently, and the possibility of obtaining information in the study of fast digital communication networks.

Thus, closeness reflects indirect relationship between the members. It is also faster and more likely for other nodes to access information with a higher closeness. Based on this, the second hypothesis is as follows:

H2: In transactional community, the higher the closeness centrality is, the greater the network influence power will be.

2.3.3 Members' betweenness centrality - the shortest path occupied

Betweenness centrality is defined as the number of the shortest paths between any two nodes in which the members is included. Freeman [18] introduced betweenness centrality to quantify nodes' control over the

communication between different nodes. That is, betweenness centrality is something about the control power over the information flow, which reflects how useful and necessary a node is in social networks.

Burt's structural hole theory proposes that individuals who connect different groups have greater influence because they have a stronger control over the information coming from different groups [13]. Nodes with high betweenness centrality work as "brokers", affecting the information flow and the communication between different individuals and groups [19]. Thus, members with high betweenness centrality are important for information flow and dissemination, which can increase its network influence. Based on this, we propose Hypothesis 3:

H3: In transactional community, the higher the betweenness centrality is, the greater the network influence power will be.

3. RESEARCH METHOD AND DATA

3.1 Data

We got data from a community called "Taobao food official gangs" through crawlers. The data was composed of the members' ID, time to join in, followers and relative information of posts (i.e. time, ID, members and contents). The data was collected from 20th July to 23th July in 2013, getting 51666 members. Then we ruled out isolated ones and got 12314 members forming 20621 links, 27433 posts and 91723 replies in a community network. We dealt with the data using software such as SQL Server and matlaB. Then we fitted number of members with posts, eliminated unqualified samples and finally got 2055 valid posts.

3.2 Model and Variables

According to research framework, we build two specific models——influence breadth model and influence depth model.

Model 1: influence scope model

$$readNO_{ij} = \alpha_1 ExistedDays_j + \alpha_2 InDays_i + \alpha_3 Gender_i + \alpha_4 jifen_i + \alpha_5 Similarity + \alpha_6 Degree_i + \alpha_7 Betweenness_i + \alpha_8 Closeness_{ii}$$

Model 2: influence depth model

$$replyNO_{ij} = \beta_1 ExistedDays_j + \beta_2 InDays_i + \beta_3 Gender_i + \beta_4 jifen_i + \beta_5 Similarity_i + \beta_6 Degree_i + \beta_7 Betweenness_i + \beta_8 Closeness_i$$

Among them, the dependent variable is the page views and numbers of replies, independent variables including intrinsic properties (gender, the number of days since member i joined the gangs, the days posts j already exists etc.), three kinds of centrality and similarity, as shown in Table 1:

Table 1. Model Variables and Descriptions

$readNo_{ij}$	Page views of member i's posts j; reflect the coverage and range of members' influence in network.
$replyNo_{ij}$	The frequency replied by others in the network, embodies the interaction of members and audience, reflecting information spread depth
$InDays_i$	The number of days since member i joined the gangs
$ExistedDays_j$	The number of days since posts j already exists
$Gender_i$	The gender of gangs member i
$jifen_i$	The credit of member i in the gangs
$Degree_i$	The number of following and follower of member i
$Betweenness_i$	The number of shortest path between two other members in the network that member i occupied
$Closeness_i$	The network distance of member ito other members in the network
$Similarity_i$	The similarity of member i with all other members in the gangs based on common concern with the same gangs member

3.3 Model estimation and hypothesis testing

Considering the feature of our data, we chose Tobit regression to explore the potential relations. In TOBIT model, the independent variables are observable and the dependent variable can only be observed in a limited way. To be specific, the dependent variable is constrained within certain limits, that is 0 or actual observed value, because the number of posts is not negative. Descriptive statistics in the model are demonstrated in Table 2:

Table 2. The statistics of variables

Variables	N	Mean	Standard deviation	Sum	Minimum	Maximum
readNo	2055	0.0003075	0.00127	0.63184	0	0.03123
replyNo	2055	0.0004854	0.00198	0.99751	0	0.04231
ExistedDays	2055	0.28842	0.22279	592.70647	0	1
InDays	2055	0.34104	0.25572	700.83945	0	1
Gender	2055	0.76399	0.42473	1570	0	1
jifen	2055	0.02959	0.0387	60.80341	0	0.19875
Similarity	2055	0.004	0.00568	8.22259	0	0.02979
Degree	2055	0.00153	0.00205	3.13405	0	0.01772
Betweenness	2055	0.0006433	0.00119	1.32191	0	0.00871
Closeness	2055	0.40168	0.24796	825.45862	0	0.79737

Note: all the data are normalization.

Furthermore, the correlation analysis results of variables are shown in Table 3:

Table 3. Correlation analysis results of variables

Pearson correlation coefficient, N = 2055 Prob> |r| under H0: Rho=0									
	readNo	replyNo	ExistedDays	InDays	Gender	jifen	Degree	Betweenness	Closeness
replyNo	0.3992 <.0001	1							
ExistedDays	0.3034 <.0001	0.1174 <.0001	1						
InDays	0.0257 0.2438	-0.0030 0.8936	0.1091 <.0001	1					
Gender	0.0011 0.9611	-0.0214 0.3314	-0.0701 0.0015	0.0139 0.5278	1				
jifen	-0.0007 0.9766	0.0483 0.0285	-0.0570 0.0097	-0.0201 0.3628	-0.2174 <.0001	1			
Degree	0.1441 <.0001	0.0845 0.0001	0.1252 <.0001	0.1577 <.0001	0.0785 0.0004	-0.1422 <.0001	1		
Betweenness	0.1093 <.0001	0.0642 0.0036	0.0125 0.5716	0.2297 <.0001	0.1727 <.0001	-0.1373 <.0001	0.7300 <.0001	1	
Closeness	-0.0002 0.9941	0.0010 0.9629	-0.0222 0.3148	0.0499 0.0237	0.0619 0.0050	-0.0972 <.0001	0.1394 <.0001	0.2068 <.0001	1
Similarity	0.0146 0.5092	-0.0308 0.1631	-0.0204 0.3547	0.1385 <.0001	0.1706 <.0001	-0.1428 <.0001	0.6150 <.0001	0.5029 <.0001	0.1997 <.0001

From the table, we exclude the endogenous interference because the correlation between dependent variable and member centrality is lacking or low. Furthermore, there exists tiny correlation between several independent variables; it cannot bring about the potential multicollinearity. Hence, we conduct a TOBIT regression using SAS based on the reliability of the regression model. Results are as follows in table 4:

Table 4. Estimation results

Variables	readNo		replyNo	
	Coefficient	Approximate Pr> |t|	Coefficient	Approximate Pr> |t|
Intercept	-0.00094	<.0001	-0.00142	<.0001
ExistedDays	0.00274	<.0001	0.00270	<.0001
InDays	-0.00011	0.40680	-0.00027	0.29170
jifen	0.00085	0.32690	0.00452	0.00700
Gender	0.00006	0.45010	0.00002	0.91390
Similarity	-0.02048	0.00700	-0.05714	0.00010
Degree	0.07523	0.00380	0.14901	0.00340
Betweenness	0.10650	0.00980	0.16008	0.04480
Closeness	-0.00001	0.95440	0.00011	0.69130
_Sigma	0.00138	<.0001	0.00260	<.0001

According to the regression results, we can find that both the degree centrality and betweenness centrality have a significant positive impact on page views and replies of the posts, but the closeness centrality have no effect. Further more, in model 1, betweenness centrality has a greater effect than degree centrality(0.10650>0.07523), which is consistent with our theory deduction that members with high betweenness centrality can facilitate dissemination of information between different network modules. While in model 2, the effect of degree centrality is more significant, which speaks volume for that members with high degree centrality has a strong influence on the members who are directly connected with them, also we can conclude that the stronger the relationship，the larger the flow of information in the network. In addition, closeness centrality is essentially about the physical location of someone in the network and it reflects the weak ties among members, so there is no significant effect on the dissemination of information. To sum up, hypothesis 1 and hypothesis 3 of our article are verified, while hypothesis 2 is not supported.

4. CONCLUSION AND DISCUSSION

Based on transactional community, this paper studies the impact of different centralities on members' network influence. It shows that transactional communities are different from social communities, such as its structure, involvement members and involvement needs. Besides, the driving force behind the structure evolution and development is also different, as social community is mainly influenced by social norms, while transactional community is based on information.

4.1 Theoretical contribution and implication

Theoretically, this study proves the different impact patterns of centralities in transactional communities: betweenness centrality has a strong influence on the impact breadth, degree centrality has a strong influence on the impact depth, while, closeness centrality has no significant influence. Meanwhile, the study also deepens the understanding of the sources of power together with Burt theory. In addition, previously, hubs are defined as members with a high degree centrality, while this study found that members with high betweenness centrality are also influential, and the influence mechanism is different. Nodes with high betweenness connect unrelated network modules, occupy a lot of structural holes and have strong control over information flow, thus contributing to build the entire network and information flows. Especially in information-driven transactional community, the members with high betweenness centrality are very important.

In practice, we can identify different hubs with different centralities and develop opinion leadership in the transactional community. Specifically, it can guide companies to choose the right hubs as its breaking point for brand promotion in the virtual network and accelerate new product adoption. In addition, companies can also strengthen its own network power so as to attract more fans, occupy more network "structural holes" and thus improve their ability to use network resources to influence customers.

4.2 Limitations

Networks are evolving, but the data we use is cross-sectional without consideration for dynamics; therefore, this research can be improved by analyzing the different influence of networks at different times.

ACKNOWLEDGEMENT

This research was supported by the National Nature Science Foundation of China under Grant 71372127.

REFERENCES

[1] Kaplan, A. M. and M. Haenlein (2010). "Users of the world, unite! The challenges and opportunities of Social Media." Business horizons53(1): 59-68.

[2] Newman, M. (2009). Networks: an introduction. Oxford University Press.

[3] Wasserman, S. (1994). Social network analysis: Methods and applications, Cambridge university press.

[4] Brass, D. J., et al. (2004). "Taking stock of networks and organizations: A multilevel perspective." Academy of management journal 47(6): 795-817.

[5] Stanoevska-slabeva, K. (2002). Toward a community-oriented design of Internet platforms. International Journal of Electronic Commerce, 6, 71-95.

[6] Armstrong, A., & Hagel, J. (2000). The real value of online communities. Knowledge and communities, 85-95.

[7] Granovetter, M. S. (1973). "The strength of weak ties." American journal of Sociology: 1360-1380.

[8] Deutsch, M., & GERARD, H. B. (1955). A study of normative and informational social influences upon individual judgment. The journal of abnormal and social psychology, 51(3), 629.

[9] Kumar, V., et al. (2013). "Creating a measurable social media marketing strategy for hokey pokey: Increasing the value and ROI of intangibles and tangibles." Marketing Sci.

[10] Burt, R. S. (2009). Structural holes: The social structure of competition. Harvard University Press.

[11] Krackhardt, D. (1998). Simmelian ties: Super strong and sticky. Power and influence in organizations, 21-38.

[12] Coleman, J. S. (1988). Social capital in the creation of human capital. American journal of sociology, S95-S120.

[13] Burt, R. S. (2005). Brokerage and Closure: An Introduction to Social Capital: An Introduction to Social Capital. Oxford University Press.

[14] Katona, Z., et al. (2011). "Network effects and personal influences: The diffusion of an online social network." Journal of Marketing Research48(3): 425-443.

[15] Sabidussi, G. (1966) The centrality index of a graph. Psychometrika 31, 581–603.

[16] M.E.J. Newman (2005), "A measure of betweenness centrality based on random walks", Social Networks 27: 39–54

[17] Trier, M. (2008). Towards Dynamic Visualization for Understanding Evolution of Digital Communication Networks. Information Systems Research, 19(3), 335-350.

[18] Freeman, Linton (1977). "A set of measures of centrality based upon betweenness". Sociometry 40: 35–41.

[19] Borgatti, S. P., et al. (2009). "Network analysis in the social sciences." Science323(5916): 892-895.