

International Conference on Computer Science and Computational Intelligence (ICCSCI 2015)

Degree Centrality for Social Network with Opsahl Method

Yoga Yustiawan^{a,*}, Warih Maharani^b, Alfian Akbar Gozali^c

^{a,b,c}*School of Computing, Telkom University, Bandung Technoplex, Bandung, 40257, Indonesia*

Abstract

In this paper we study how to determine the nodes that most influential to a node in the network. Social Network Analysis (SNA) can measure the centrality of a node in order to obtain an influential nodes in the dissemination of information. One of the centrality measurement that can be applied is degree centrality. In this research, the method used is Opsahl method, combines two indicators, the number of neighborhood (degree) and the amount of weight relations (strength) of a node and uses tuning parameters. The weight relations are obtained from the number of relations as following/follower, mention and reply. Tuning parameters are parameters which are used to set the influence of both degree and strength to the degree centrality measurement results. Based on test results, the node who has a high strength value is derived from weight relations which are obtained from mentions and replies.

© 2015 Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Peer-review under responsibility of organizing committee of the International Conference on Computer Science and Computational Intelligence (ICCSCI 2015)

Keywords: social network; degree centrality; opsahl method; tuning parameter

1. Introduction

Social Network Analysis (SNA) is a wide strategy to investigate the social structure³. SNA is used to obtain patterns of relationships between nodes to ascertain underlying social structure^{4,10}. Node is represented as people and group (users), meanwhile relation shows the relationship between nodes^{7,9}. The problem is how to find the most influential node in the network. The focus is how to measure centrality in the network. This research will focus on

* Corresponding author. Tel.: +62-8222-000-6617; fax: +62-285-447-3248.

E-mail address: yogagm@students.telkomuniversity.ac.id (Y. Yustiawan), wmaharani@telkomuniversity.ac.id (W. Maharani), alfian@tass.telkomuniversity.ac.id (A.A. Gozali)

degree centrality measurement. Degree centrality measurement is implemented by Opsahl Method. Opsahl Method is chosen because this method uses 2 indicator to measure degree centrality, those are degree and strength. The parameter is used for combining those two indicators, it is called tuning parameter. Tuning parameter defines the influence between two indicators to a node in the network⁸.

Degree is the the sum of nodes which are connected to a node and strength is the sum of weight relations of a node. If we just consider one of those two indicators, we cannot make sure the centrality value of node is right or not to determine influential node in a network especially for social network. It is because there are communications between one node to another, we consider the communications as a strength of node. The communications also determine either the node is influential or not. The communications consist of follow, mention and reply. Opsahl Method can analyze the influence of two indicators which are used.

2. Research Method

Opsahl Method is implemented by a system. The system design consist of several steps such as data crawling, preprocessing, degree centrality measurement and rank process.

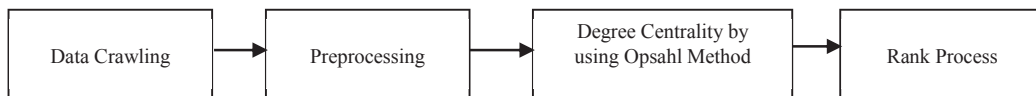


Fig. 1. System Design

2.1. Data Crawling

Dataset is taken from social network Twitter. We choose social network Twitter because our goal is to explore the relationship between user interests and their interactions in the social network, we need to collect information about users and link structure⁵. The crawling process is handled by NodeXL application². Dataset includes Twitter users and their relations.

1					
2	Vertex 1	Vertex 2	Relationship	Relationship Date (UTC)	Edge Weight
4	User 3	User 4	Followed	41947,26302	2 OK
5	User 3	User 5	Followed	41947,26302	2 OK
6	User 3	User 6	Followed	41947,26302	1 OK
7	User 7	User 3	Followed	41947,26302	1 OK
8	User 3	User 2	Follower	41947,26302	1 OK
9	User 9	User 8	Followed	41947,26302	1 OK
10	User 9	User 2	Followed	41947,26302	1 OK
11	User 9	User 10	Followed	41947,26302	1 OK
12	User 11	User 12	Followed	41947,26302	1 OK

Fig. 2. Twitter Dataset

2.2. Preprocessing

Preprocessing is a process to manage data before system manages the data. The flow of data preprocessing is shown as bellows²:

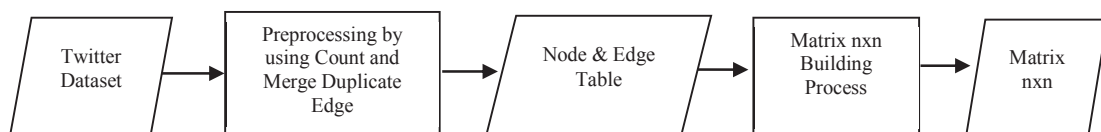


Fig. 3. Preprocessing

If preprocessing is done, then this will generate matrix nxn. This matrix will be used by a system. Matrix nxn is shown as bellows:

	User A	User B	User C	User D	User E	etc...
User A	0	1	0	0	0	
User B	1	0	1	2	2	
User C	0	1	0	2	2	
User D	0	2	2	0	2	
User E	0	2	2	2	0	
etc...						

Fig. 4. Matrix nxn

2.3. Degree Centrality Measurement by Using Opsahl Method

This process will do degree centrality measurement by using Opsahl Method for the preprocessing result data. There are several steps to do that. First, weighting process is a process to give a weight for each relation between nodes. Every relations have their own weights. Each of relations will add 1 to the weight when the relations occur. Second, this step for measuring the degree of the nodes. Based on matrix nxn, we find degree of each nodes which is incident with the central node. Third, after we found the degree, then we measure the strength of nodes. Last, if we have already have degree and strength, then we calculate degree centrality of each nodes.

2.4. Rank Process

After we have already calculated degree centrality of each nodes, then we sort out the degree centrality value from highest to lowest. We make rank of each nodes.

3. Results and Analysis

In this section, we will discuss about the result and analysis of implemented system. It consist of result and analysis for tuning parameter influence, result and analysis of central node influence and result and analysis of users with the same degree centrality values.

3.1. Result and Analysis of Tuning Parameter Influence

Table 1. Tuning Parameter (α) = 0 Graph

Node Rank	Node	C_D (Degree)	C_D^w (Strength)	$C_D^{w\alpha}$ with $\alpha=0$
1	User 2	397	802	397
2	User 19	197	377	197
3	User 12	182	284	182
4	User 21	163	271	163
5	User 26	112	143	112
6	User 16	106	143	106
7	User 87	106	278	106
8	User 86	97	367	97
9	User 107	95	464	95
10	User 151	88	117	88

Based on Table 1 above, we can see that the tuning parameter value = 0, the centrality of a node to be equal to the degree of a node. This happens because the Opsahl method is applying standard tuning parameter value = 0 for measuring centrality by only considering one indicator that ignores the degree and strength indicator which is owned by a node. It is identical to occur in degree centrality measurements performed by Freeman⁶. By considering only the value of any degree, then the node which has the greatest degree value will be the most popular node in the dataset community. The greater the degree of its value, the greater the value of its centrality. In other words, the value of degree of a node is proportional to the value of its centrality.

Strength value has no effect at all when using the tuning parameter 0 so this means that the weight of a relation is not giving effect to the calculation results centrality of a node. Nodes that have a strong relationship or high weight will not be a node influential if it has a low degree value.

Table 2. Tuning Parameter (α) = 0.5 Graph

Node Rank	Node	C_D (Degree)	C_D^w (Strength)	$C_D^{w^\alpha}$ with $\alpha=0.5$
1	User 2	397	802	564,26
2	User 34	85	1611	370,04
3	User 19	197	377	272,52
4	User 90	84	706	243,52
5	User 12	182	284	227,34
6	User 21	163	271	210,17
7	User 107	95	464	209,95
8	User 86	97	367	188,67
9	User 93	85	394	183,00
10	User 99	74	423	176,92

Based on Table 2 above, we can see that the value of the tuning parameter between 0 and 1, the value of centrality of a node can be affected by two indicators of degree and strength. The greater the degree and strength of a node, providing a positive impact that increases the value of the node centrality. For this dataset, it is seen that the User 34 has a very high strength, but User 34 remain in second place after the User 2 because User 34 has sufficient degree value much lower than User 2.

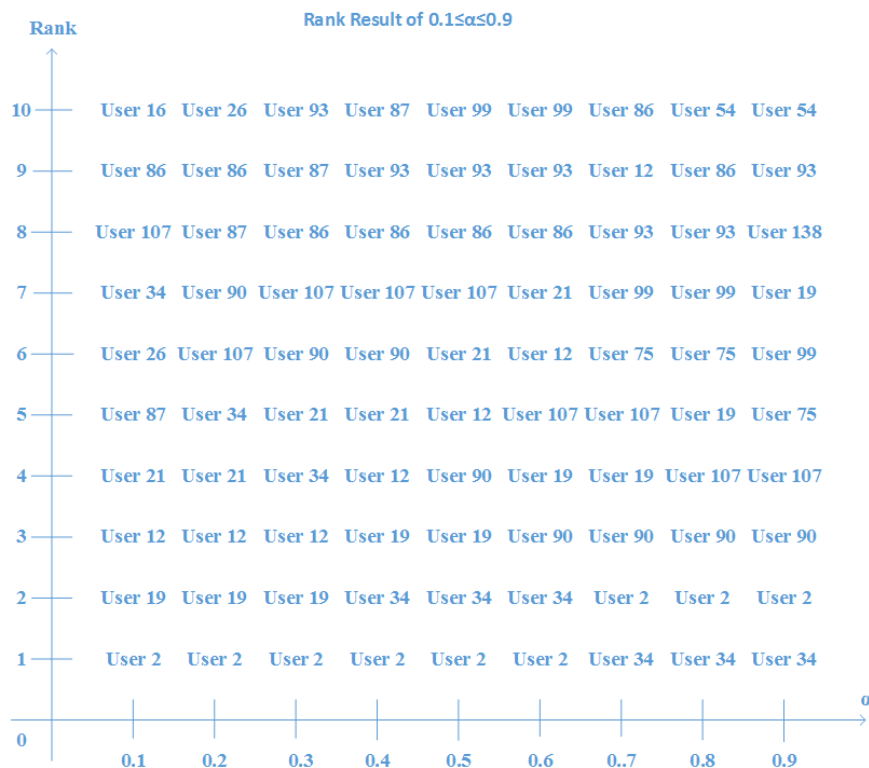


Fig. 5 Rank Result of $0.1 \leq \alpha \leq 0.9$

Based on Fig. 5 above, it appears that the position of the user rank User 2 decreased when the tuning parameter values above 0.6. This is due to the tuning parameter value above 0.7 means that the value of the tuning parameter start approaching 1 and prove that the closer to the tuning parameter value is equal to 1, the greater influence strength indicator than the degree of a node. User 2 decreased ranking due to its lack of relation weights compared with relation weights owned by the User 34. Likewise, it happens on User 90 who initially only be in rank 4, could eventually rise to the rank third on the tuning parameter values from 0.6 to 0.9. This proves that the User 90 have big strength or a strong relationship. The weight of the relation itself consists of weight relations follow, mention or reply. For relations such as follow can be divided into two namely those following and follower relation. From the fourth relation, a relation of the most influential on the relation weight is mention and reply. This is evident on the node that has a high relation weights which is ranked big three beat node that has a high degree. Here below are the details of the weight table relationships possessed by some important node is able to achieve three upper rankings at the time of the tuning parameter values between 0 and 1, among others:

Table 3. Relation Weights of 3 Upper Ranked Nodes ($0 < \alpha < 1$)

Node	Following	Follower	Mention	Reply
User 2	416	220	45	121
User 19	111	184	7	75
User 12	34	173	6	69
User 34	110	53	449	999
User 90	75	53	212	366

As shown in Table 3, User 34 and User 90 have high relation weight value of mention and reply so value positively strength of the both nodes. This shift affects the three upper rank of which are previously achieved by User 19 and User 12. Based on the results of the relation weights details on Table 3 strengthen the influence of these two types of relations which are mention and reply.

Table 4. Tuning Parameter (α) = 1 Graph

Node Rank	Node	C_D (Degree)	C_D^w (Strength)	$C_D^{w\alpha}$ with $\alpha=1$
1	User 34	85	1611	1611
2	User 2	397	802	802
3	User 90	84	706	706
4	User 107	95	464	464
5	User 75	65	452	452
6	User 138	37	440	440
7	User 99	74	423	423
8	User 54	62	400	400
9	User 93	85	394	394
10	User 19	197	377	377

Based on Table 4 above, it can be seen that the value of centrality of a node is equal to the value of the strength of a node. So that the value of the tuning parameter = 1, the indicators considered by Opsahl method for determining the value of centrality of a node is just the value of its strength alone. It is identical to occur in degree centrality measurements made by Barrat et.al.¹. By not considering the degree of a node in the calculation of the value of centrality, the greater the strength value owned by a node, the node will be increasingly popular in the dataset community. In other words, the value of centrality of a node is proportional to the strength value that is owned by the node. Because the value of this parameter tuning into account only the value of strength alone, then the impact on the importance of the effect of relation weight to the determination of the ranking results of the calculation of centrality of a node. Nodes that have higher relation weights have a great opportunity into a node with the top rank.

Node Rank	Node	C_D (Degree)	C_D^w (Strength)	$C_D^{w\alpha}$ with $\alpha=1.5$
1	User 34	85	1611	7013,48
2	User 90	84	706	2046,76
3	User 251	10	347	2044,06
4	User 278	11	302	1582,39
5	User 138	37	440	1517,32
6	User 287	14	313	1479,96
7	User 75	65	452	1191,93
8	User 2	397	802	1139,89
9	User 107	95	464	1025,45
10	User 54	62	400	1016,00

Here is a test result to obtain the rank results of the tuning parameter values range from 1.5 to 10, ie (written by node ID):

Rank	User Rank (written by User ID)								
	$\alpha=1.5$	$\alpha=2$	$\alpha=2.5$	$\alpha=3$	$\alpha=3.5$	$\alpha=4$	$\alpha=4.5$	$\alpha=5$	$\alpha=5.5$
1	34	34	34	34	34	251	251	251	251
2	90	251	251	251	251	34	34	34	34
3	251	278	278	278	278	278	278	278	278
4	278	287	287	287	287	287	287	287	287
5	138	90	138	138	138	138	138	281	281
6	287	138	90	90	281	281	281	138	138
7	75	75	297	281	294	294	294	294	294
8	2	294	284	294	75	286	286	286	286
9	107	281	286	286	286	75	75	75	75
10	54	104	75	75	291	291	291	291	291

[illegible]

According to Table 6 and 7, the rank result which is produced is quite stable, but still there is little change in the rank position. At the time of the tuning parameter value is 1.5, three upper rank occupied by User 34, User 90 and User 251. Furthermore, when the value of the tuning parameter increased, there is three stable upper rank occupied by User 34, User 251 and User 278 for the tuning parameter values from 2 to 3.5. There are changes in the rank position User 251 with User 34 when tuning parameter value reaches 4. The stability of the three upper rank last until the tuning parameter value reaches 10. Three uppermost nodes reach the three highest rank because it has a fairly high strength values and the degree values are quite low if we differentiate it with the value of its strength. These conditions prove that the tuning parameter values above 1, the value of high strength and low degree value of a node increase the value of its centrality to the node so make the node becomes the most influential and reach the top rank. Strength values of nodes which are the 3 upper rank can be seen in Table 5 below:

Table 8. The Relation Weights of 3 Upper Ranked Nodes ($\alpha > 1$)

Node	Following	Follower	Mention	Reply
User 34	110	53	449	999
User 251	15	4	22	306
User 278	11	7	12	272
User 90	75	53	212	366

In the Table 8, the nodes that have similar characteristics which are mention and reply relation weights greater than the weight of following and followers. It is proved that the mention and reply more influential for the centrality measurement with tuning parameter above 1. It is also proved that it is a high strength value positively the centrality value which is caused by the presence of particularly high relation weights such as mention and reply as well as the value of degree which is low strengthen the centrality value which is caused the degree value inversely proportional to the strength of a node value in the formula calculation of degree centrality in this Opsahl method.

3.2. Result and Analysis of Central User Influence

In this test result, central user in this dataset is User 2 is eliminated in the community, by cutting User 2 and other node relations with User 2. After testing, this produces rank comparison between graphs that include central user and does not include central user.

Based on those result shows that the overall value of the tuning parameter, there are insignificant changes in rank positions of the top 10. This is proven by looking at the rank position of the nodes under the User 2 rank contained in tables only go up 1 rank caused by the loss of User 2 node. This means that the relation between the nodes with User 2 are not too strong. Relation weights are owned by the nodes to the User 2 are low and do not cause significant changes. For example, the tuning parameter value of 0.5, rank 2 is occupied by User 34 and rank 3 occupied by User 19. Relation weight is owned by the User 34 to central user User 2 is 87 and the strength of the User 34 is 1611. Relation weights is owned by the User 19 to central user User 2 is 2 and the strength of the User 19 is 377. Thus, in case of cutting User 2 relation to the both nodes, then the value of strength User 34 becomes 1524 and the value of strength User 19 becomes 375. It is seen that does not occur in significant reductions in both the node. Both nodes are also supported by the remaining value of degree are still high because the degree value is only reduced one which is caused by cutting relation with User 2. In the tuning parameter value of 0.5, the value of degree and strength value of a node has a strong influence on the degree centrality of users. So User 34 can maintain its ranking above User 19. This is due to the influence of both the value of degree and strength of the User 34 greater than the value of the degree and strength of the User 19.

4. Conclusion

In the tuning parameter values between 0 and 1 ($0 < \alpha < 1$), the value of degree and the value of strength of a node influences the rank of a node. In the tuning parameter values above 1 ($\alpha > 1$), the value of strength has bigger influence than the value of degree to the rank of a node. Stability of user ranks in the calculation of the degree centrality value with Opsahl method occurs when the value of the tuning parameter is above 1.

Relations that influence most of the relation weighting or strength value of a node are mention and reply. It has been shown in the test results, that the value of strength which increases the centrality value of node, most of it comes from the relations like mention and reply.

Appendix A. Rank Result of $0.1 \leq \alpha \leq 0.5$

Rank	$\alpha=0,1$	$\alpha=0,2$	$\alpha=0,3$	$\alpha=0,4$	$\alpha=0,5$
1	User 2	User 2	User 2	User 2	User 2
2	User 19	User 19	User 19	User 34	User 34
3	User 12	User 12	User 12	User 19	User 19
4	User 21	User 21	User 34	User 12	User 90
5	User 87	User 34	User 21	User 21	User 12
6	User 26	User 107	User 90	User 90	User 21
7	User 34	User 90	User 107	User 107	User 107
8	User 107	User 87	User 86	User 86	User 86
9	User 86	User 86	User 87	User 93	User 93
10	User 16	User 26	User 93	User 87	User 99

Appendix B. Rank Result of $0.6 \leq \alpha \leq 0.9$

Rank	$\alpha=0,6$	$\alpha=0,7$	$\alpha=0,8$	$\alpha=0,9$
1	User 2	User 34	User 34	User 34
2	User 34	User 2	User 2	User 2
3	User 90	User 90	User 90	User 90
4	User 19	User 19	User 107	User 107
5	User 107	User 107	User 19	User 75
6	User 12	User 75	User 75	User 99
7	User 21	User 99	User 99	User 19
8	User 86	User 93	User 93	User 138
9	User 93	User 12	User 86	User 93
10	User 99	User 86	User 54	User 54

References

1. Alain B, Marc B, Romualdo PS, Alessandro V. The architecture of complex weighted networks. *PNAS*. 2004; 101(11): 3747-3752.
2. Derek LH, Ben S, Marc AS. Analyzing Social Media Networks with NodeXL. First Edition. Burlington: Morgan Kaufmann. 2011: 53-67.
3. Evelien O, Ronald R. Social network analysis: a powerful strategy, also for the information sciences. *Journal of Information Science*. 2002; 28(6): 441-453.
4. Hildrun K, Theo K. A New Centrality Measure for Social Network Analysis Applicable to Bibliometric and Webometric Data. *COLLNET Journal of Scientometrics and Information Management*. 2007; 1(1): 1-7.
5. Juan B, Zhiguang Q, Jia H. Detecting Community and Topic Co-Evolution in Social Networks. *TELKOMNIKA Indonesian Journal of Electrical Engineering*. 2014; 12(5): 4063-4070.
6. Linton CF. Centrality in Social Networks Conceptual Clarification. *Social Networks*. 1978; 1: 215-239.
7. Stanley W, Katherine F. Social Network Analysis Method and Applications. First Edition. New York: Cambridge University Press. 1994: 17-21.
8. Tore O, Filip A, John S. Node centrality in weighted networks: Generalizing degree and shortest paths. *Social Networks*. 2010; 32: 245-251.
9. Xingqin Q, Eddie P, Qin W, Yezhou W, Cun-Quan Z. Laplacian centrality: A new centrality measure for weighted networks. *Information Sciences*. 2012; 194: 240-253.
10. Zudha AR, Warih M, Adiwijaya. The Analysis and Implementation of Degree Centrality in Weighted Graph in Social Network Analysis. Information and Communication Technology (ICoICT), 2013 International Conference of. Bandung: IEEE; 2013. p. 72-76.