2018 ASEE Mid-Atlantic Section Spring Conference: Washington, District of Columbia Apr 6

Node Centrality and Ranking Tool

Dr. Paul Cotae, University of the District of Columbia

Dr. Paul Cotae, Professor of Electrical and Computer Engineering has more than 25 years of experience in the communication field (research and education). He received a Dipl. Ing. and a M.S. degrees in communication and electronic engineering in 1980 from the Technical University of Iassy and a Ph.D. degree in telecommunications from "Politechnica" University of Bucharest, Romania in 1991, and a Master in Applied Mathematics in 1998 from the University of Colorado at Boulder. From 1994 to 1998 he spent four years at the University of Colorado at Colorado Springs and the University of Colorado at Boulder as a Fulbright Scholar and Visiting Associate Professor doing research and teaching in ECE department and APPM department. He served also as a consultant to Navsys Corp., Colorado Springs, in 1997. From 2002 to 2008 he was with the Department of Electrical and Computer engineering at the University of Texas at San Antonio (UTSA). From 1984 to 2001, he was with the Department of Electrical Engineering, Technical University of Iassy, where he conducted research and teaching in the area of digital communications as a Full Professor at the same department. Since 2008, he has been with the University of the District of Columbia as a Professor. His current research interests include multiple access, modulation and coding, mobile communications, and digital communication systems. He has authored or coauthored more than 150 papers in these areas and four books. Dr. Cotae serves as an Associate Editor for IEEE Communication Letters, EURASIP Journal on Signal Processing, Elsevier International Journal of Computers and Electrical Engineering (JCEE), and he has been on the Technical Program Committee and Session chair of the IEEE Conferences GLOBECOM (2003-2011), VTC Spring 2005, 2006 and ICC 2005-2011. He is a Senior Member of IEEE, member of ASEE, member of HKN (Eta Kappa Nu) and SIAM. He is cited in Who's Who in American Education, Who's Who in America, and in Who's Who in the World. He is Faculty Fellow for the ONR-ASEE Summer Faculty Research Program 2009-2018.

Mr. Luis P. Aguinaga, University of the Distric of Columbia

Node Centrality and Ranking Tool

Luis Aguinaga luis.aguinaga@udc.edu **Paul Cotae** pcotae@udc.edu
Department of Electrical and Computer Engineering
School of Engineering and Applied Sciences
University of the District of Columbia
4200 Connecticut Avenue NW
Washington, DC 20008,USA

Abstract –The paper is about the topology of a network. Find a center is always important in geometry. There are different kinds of center but in general centers minimize or maximize a property of the geometrical figure. The ranking tool is just a way to figure out how far or close to a center is a particular node in the network. Node centrality is the crucial importance in different areas like transportation, wireless communication, Internet, virology and more. Each time when we have a network, knowing where are the centers locations in a network is the crucial importance to improve, optimize, protect or attack a network.

Keywords
Network, Centrality, Node, Link, Topology, Degree, Closeness, Betweenness, and Eigenvector

Introduction

Any network can be characterized by a vector of nodes $N = \{v_1, v_2, \cdots v_n\}$ and an adjacency matrix A which elements are the links l_{ij} between the nodes and, l_{ij} is the link between the nodes is v_i and v_j. In this paper we assume that all of the links do not have a preferred direction, and also all of them have the same value. In addition all of the nodes have the same importance. This means that all of the nodes will be treated as if they have the same characteristics. Under these conditions the matrix A is symmetric.

In any network a critical factor is which node is the more important among all nodes. Just thinking about if we have a disease it will spread jumping from one node to other. In this case cut the links or nodes that are more important will be delay or stop the disease. Or in the case of a computer network figure out which node is the most important one will be use to improve that node to increase our security, or if that is not our network that will be a good node to start and attack. In a traffic network this information will be use to decide which roads or intersections need to be enlarged in order to improve the network in a more efficient way. In a communication network this information can help us to reduce cost and make the less critical nodes more accessible in price.

For all of these reasons and, many more, it is determined which is the importance of the nodes will be useful. Here is when a node centrality and ranking tools make and apparition. We use the centralities to characterize each node in a particular network. Then we assign each node a rank base of how central is this node. In this paper we will talk about 4 node centralities measures :Degree, Closeness, Betweenness, and Eigenvector centrality. Then we will use these criteria independently and jointly to evaluate for different sample networks.

Degree Centrality (DC)

In this first Centrality we will rank as most important node the one which has more connections. For example in the case of a disease the most important thing is limit the links that pathogen can be use to spread the infection. It is easy to image that in this case the number of connections that a node has is the most important thing.
We define dc_i as:

$$dc_i = \sum_{j=1}^{n} dc_{ij} \qquad (1)$$

$$where\ dc_{ij} = \begin{cases} 1\ if\ there\ is\ a\ link\ beteween\ i\ and\ j \\ 0\ if\ there\ is\ not\ a\ \ link\ beteween\ i\ and\ j \end{cases}$$

In this case the network can be characterize by the matrix *DC* or an equivalent vector *DC*

$$DC = \begin{bmatrix} 0 & dc_{12} & \cdots \\ dc_{21} & 0 & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \equiv DC = \begin{bmatrix} dc_1 \\ dc_2 \\ \vdots \end{bmatrix}$$

Closeness Centrality (CC)

Now let's say that the most important thing is that a node is close to another node. Or in other words how many channels are between these nodes using the closes path. This idea has relevance, for example, in a distribution network. The most important warehouse will be the one that is close the others warehouses. We can move material from one warehouse to another one and obviously the transport will be cheaper if the warehouses or nodes are close in distance. Of course in this particular example the links will not have the same value. But for now we will use the same values for any link. We define cc_i as

$$cc_i = \left(\sum_{j=1}^{n} cc_{ij} \right)^{-1} \qquad (2)$$

$$where\ cc_{ij} = d(v_i, v_j)$$

$d(v_i, v_j)$ is the distance between the nodes v_i and v_j; and it is the sum of the links that we need to cross for going from the node v_i to the node v_j
In this case the network can be characterize by the matrix *CC*

$$CC = \begin{bmatrix} 0 & cc_{12} & \cdots \\ cc_{21} & 0 & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}$$

Take in account that

$$cc_{ij} = 0 \Leftrightarrow i = j\ or\ cc_{ij} \geq 1 \Leftrightarrow i \neq j$$

Betweenness Centrality (BC)

In this case the importance of the node depends if the node is part of a path that is use for others nodes. First we define $\|G(v_j, v_k)\|$ as the cardinality of the set that include all of the geodesics between to nodes *j* and *k*. (a geodesic is the minimal distance between two nodes). Also we defined $\|G_{v_i}(v_j, v_k)\|$ which is the cardinality of the subset that include all of the geodesics between nodes *j* and *k* that include the node *i* in their path. This will be relevant in the case of a communication network. If we have a router or antenna that handle the information between many nodes this particular node will be highly important. In this case be close or have connections is not as relevant as be part of the path between nodes. We define bc_i as:

$$bc_i = \sum_{j \neq i \neq k} \frac{\left\| G_{v_i}(v_j, v_k) \right\|}{\left\| G(v_j, v_k) \right\|} \tag{3}$$

The centrality of the node *i* dependent of its relationship with the other nodes for *bc* we do not have a characteristic matrix we have a tensor of order 3 in which the elements of the tensor are:

$$bc_i^{jk} = \frac{\left\| G_{v_i}(v_j, v_k) \right\|}{\left\| G(v_j, v_k) \right\|} \Rightarrow BC = BC_i^{jk} \tag{4}$$

Take in account that

$$0 \leq bc_i^{jk} \leq 1$$

Eigenvector Centrality (EC)

If we use the characteristic matrix *A=DC* which is the matrix defines using equation (1) by the theorem:

> Theorem: (Perron-Frobenius) If a matrix A has a real eigenvalue λ_{max} greater than or equal to the magnitude of any of its other eigenvalues. There is a nonnegative eigenvector v_A associated with λ_{max}

The value ec_i is simple the value of the component *i* of the vector v_A. λ_{max} is also called the spectral radius of *A* and the vector v_A of length 1 is called the Gould vector. This model is particularly use in virology networks. If you put the disease in the most ranking node the disease will be spread faster around the network.

Combine Centrality

Sometimes is useful and more precise use more than one centrality value. In this analysis we use the linear combination of the four centralities describe above. And define.

$$cm_i = w_1 \cdot dc_i + w_2 \cdot cc_i + w_3 \cdot bc_i + w_4 \cdot ec_i \tag{5}$$

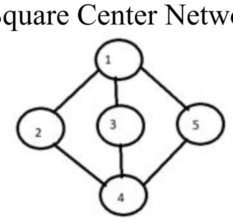
This combine centrality *cm* uses some weights w_n to characterize the degree of importance of each centrality value. Take in account that none of the centralities were normalizing and each value is limited by:

$$\begin{aligned} 1 &\leq dc_i \leq n-1 \\ \frac{1}{(n-1)^2} &\leq cc_i \leq \frac{1}{n-1} \\ 0 &\leq bc_i \leq \binom{n-1}{2} \\ 0 &< ec_i < 1 \end{aligned} \tag{6}$$

This normalization values will be useful if we want that all of the centralities values will be in the same scale.

Examples **Table 1 relative weighs use in all examples**

a) Square Center Network



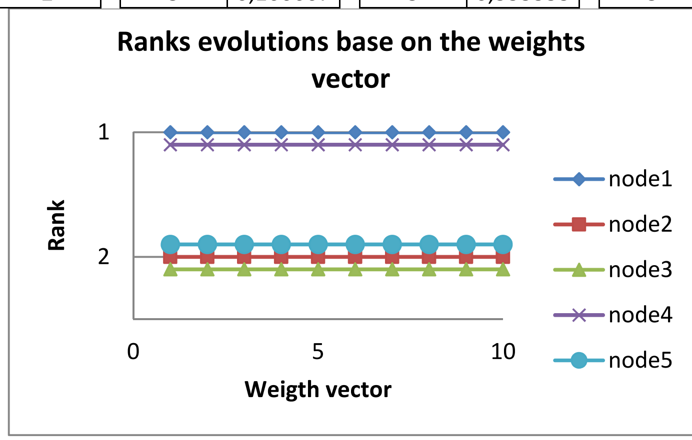
set number	Relative weights base on the normalization value			
1	1	1	1	1
2	4	0	0	0
3	0	4	0	0
4	0	0	4	0
5	0	0	0	4
6	1	1	1,75	0,25
7	1	1	2	0
8	2	0,5	0,5	1
9	0	0,5	3	0,5
10	0,8	1,2	1	1

Node in Rank order	DC value
1	3
4	3
2	2
3	2
5	2

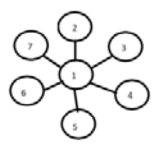
Node in Rank order	CC value
1	0,2
4	0,2
2	0,166667
3	0,166667
5	0,166667

Node in Rank order	BC value
1	1,5
4	1,5
2	0,333333
3	0,333333
5	0,333333

Node in Rank order	EC value
1	0,5
4	0,5
2	0,408248
3	0,408248
5	0,408248



b) Star Network

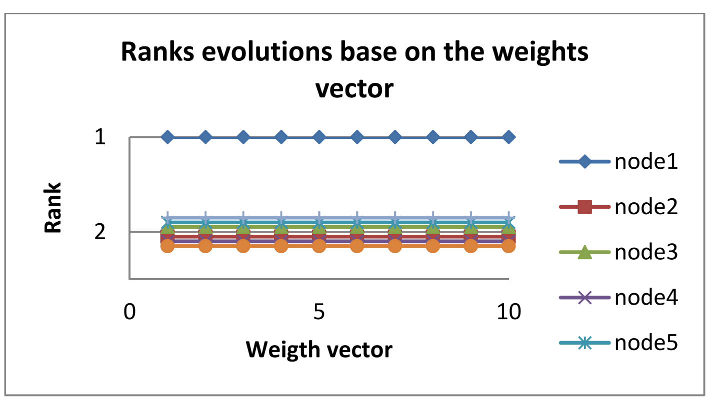


Node in Rank order	DC value
1	6
2	1
3	1
4	1
5	1
6	1
7	1

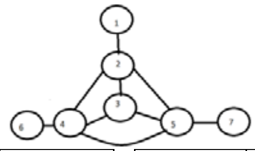
Node in Rank order	CC value
1	0,16666667
2	0,09090909
3	0,09090909
4	0,09090909
5	0,09090909
6	0,09090909
7	0,09090909

Node in Rank order	BC value
1	15
2	0
3	0
4	0
5	0
6	0
7	0

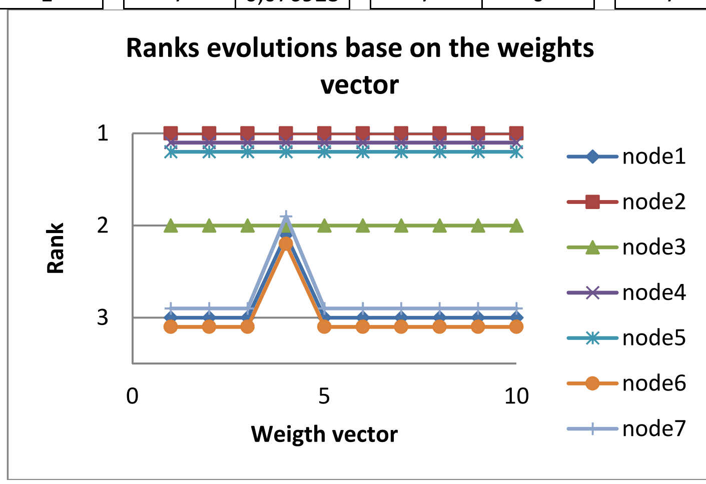
Node in Rank order	EC value
1	0,70710678
2	0,28867513
3	0,28867513
4	0,28867513
5	0,28867513
6	0,28867513
7	0,28867513



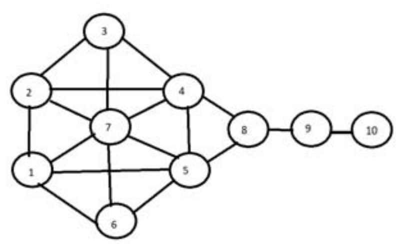
c) Triangular Center Network



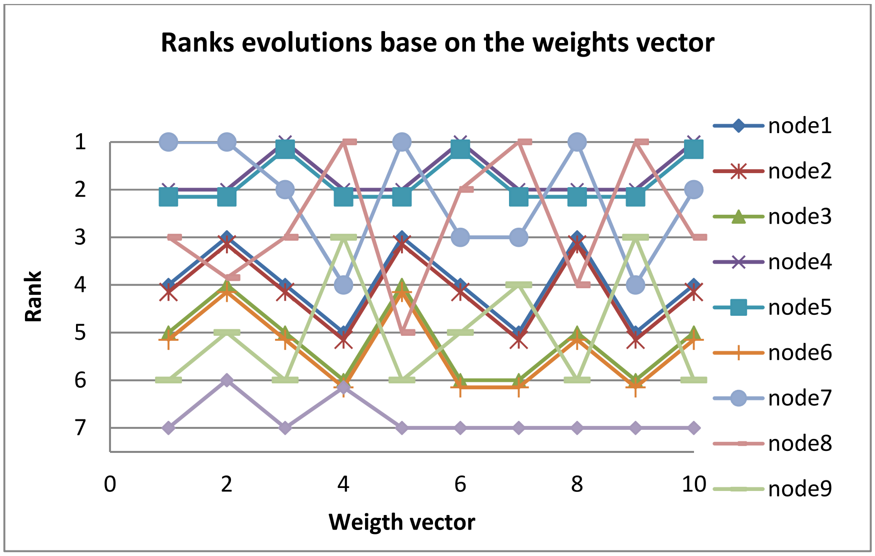
Node in Rank order	DC value	Node in Rank order	CC value	Node in Rank order	BC value	Node in Rank order	EC value
2	4	2	0,125	2	5	2	0,491123
4	4	4	0,125	4	5	4	0,491123
5	4	5	0,125	5	5	5	0,491123
3	3	3	0,111111	1	0	3	0,455296
1	1	1	0,076923	3	0	1	0,151765
6	1	6	0,076923	6	0	6	0,151765
7	1	7	0,076923	7	0	7	0,151765



d) Kite Network



Node in Rank order	DC value	Node in Rank order	CC value	Node in Rank order	BC value	Node in Rank order	EC value
7	6	4	0,071429	8	14	7	0,481021
4	5	5	0,071429	4	8,333333	4	0,397691
5	5	7	0,066667	5	8,333333	5	0,397691
1	4	8	0,066667	9	8	1	0,352209
2	4	1	0,058824	7	3,666667	2	0,352209
3	3	2	0,058824	1	0,833333	3	0,285835
6	3	3	0,055556	2	0,833333	6	0,285835
8	3	6	0,055556	3	0	8	0,195861
9	2	9	0,047619	6	0	9	0,048073
10	1	10	0,034483	10	0	10	0,011163



Conclusions

If a particular node has a lower rank that other node in all of the centralities, that node has a lower rank in any combination of the nodes as long as all the weigh are positive. In the kite example we can observe that the node 8 always has a rank superior the node 9. We can also observe that the nodes 4 and 5 are always on rank 1 or 2, it is ease to notice that these nodes acts a junction between the body and the tale of the kite. We can observe also that there is not a unique rank. The rank is mostly base on the nature of the problem, the same network can have different centers base on different parameters. Another interesting thing is the fact that for different centralities the network can be represented by a vector, matrix or tensor this is mostly because a particular characteristic of the network can be a function of the node, the links or the relationship between the node and the links.

Acknowledgment

This research was funded by the Army Research Office (ARO) - Department of Defense (DOD)–Award No. W911NF-15-1-0481.

References

1. Moskowitz, Ira S. et al. “Network topology and mean infection times.” Social Network Analysis and Mining 6 (2016): 1-14.
2. Andrés Vázquez-Rodas , Luis J. de la Cruz Llopis “A centrality-based topology control protocol for wireless mesh Networks” Ad Hoc Networks Volume 24, Part B, January 2015, Pages 34-54
3. Terrill L. Frantz, Marcelo Cataldo, Kathleen M. Carley “Robustness of centrality measures under uncertainty: Examining the role of network topology” Published online: 15 December 2009 Springer Science+Business Media, LLC 2009
4. Komal Batool, Muaz A. Niazi ” Towards a Methodology for Validation of Centrality Measures in Complex Networks” PLoS ONE 9(4): e90283.doi:10.1371/ journal.pone.0090283
5. Klaus Wehmuth, Artur Ziviani “DACCER: Distributed Assessment of the Closeness Centrality Ranking in complex networks” Computer Networks Volume 57, Issue 13, 9 September 2013, Pages 2536-2548
6. M. Kas, L. R. Carley and K. M. Carley, "Monitoring social centrality for peer-to-peer network protection," in *IEEE Communications Magazine*, vol. 51, no. 12, pp. 155-161, December 2013. doi: 10.1109/MCOM.2013.6685771

7. A. Syarif, A. Abouaissa, L. Idoumghar, P. Lorenz, R. Schott and G. Staples, "New Path Centrality Based on Operator Calculus Approach for Wireless Sensor Network Deployment," in *IEEE Transactions on Emerging Topics in Computing*, vol. PP, no. 99, pp. 1-1.
8. Anne-MarieKermarrec[a]ErwanLe Merrer[b]BrunoSericola[a]GillesTrédan "Second order centrality: Distributed assessment of nodes criticity in complex networks" Computer Communications Volume 34, Issue 5, 15 April 2011, Pages 619-628

Luis Aguinaga

Luis Aguinaga is a Graduate Research Assistant in the Electrical and Computer Engineering Department at the University of the District of Columbia. He is working towards his MSEE thesis in cybersecurity.

Dr. Paul Cotae

Dr. Paul Cotae, Professor of Electrical and Computer Engineering is the Director of the PhD Program at SEAS and Director of the SEAS Research Center. His research is in Digital Communication, Information theory, Statistics and Applied Mathematics and Cybersecurity: Anomaly detection, Detection of Low Rate Denial of Service Attacks, Intrusion Detection, Information Visualization. He published more than 140 conference and journal papers, many of them at IEEE level, authored 2 books and coauthored 3 books in the area of digital communications systems.

During the AY 2014-2015 he spent his sabbatical at the Center for High Assurance Computer Systems Code 5540, Naval Research Laboratory, Washington DC, 20375. Since 2009 he has been selected every summer as ONR Senior Research Fellows for the ASEE Summer Faculty Research Program at NRL.

His research is sponsored by NSF, ONR, AFOSR and USAF. He received in last five years more than $1M for his research from DOD as a sole PI for the following grants: –Army Research Office (ARO) –Award No. W911NF-15-1-0481: "Performance Data-Driven Methods and Tools for Computer Network Defense through Network Science" and Office of Naval Research - Award no. W911NF-11-1-0144 "Information-Driven Blind Doppler Shift Estimation and Compensations Methods for Underwater Wireless Sensor Networks".