CORRELATION OF EIGENVECTOR CENTRALITY TO OTHER CENTRALITY MEASURES: RANDOM, SMALL-WORLD AND REAL-WORLD NETWORKS

Xiaojia He[1] and Natarajan Meghanathan[2]

[1]University of Georgia, GA, USA,
[2]Jackson State University, MS, USA
[2]natarajan.meghanathan@jsums.edu

ABSTRACT

In this paper, we thoroughly investigate correlations of eigenvector centrality to five centrality measures, including degree centrality, betweenness centrality, clustering coefficient centrality, closeness centrality, and farness centrality, of various types of network (random network, small-world network, and real-world network). For each network, we compute those six centrality measures, from which the correlation coefficient is determined. Our analysis suggests that the degree centrality and the eigenvector centrality are highly correlated, regardless of the type of network. Furthermore, the eigenvector centrality also highly correlates to betweenness on random and real-world networks. However, it is inconsistent on small-world network, probably owing to its power-law distribution. Finally, it is also revealed that eigenvector centrality is distinct from clustering coefficient centrality, closeness centrality and farness centrality in all tested occasions. The findings in this paper could lead us to further correlation analysis on multiple centrality measures in the near future.

KEYWORDS

Eigenvector Centrality, Correlation Coefficient, Random Network, Small-world Network, Real-world Network

1. INTRODUCTION

Over the past few decades, eigenvector, proposed by Bonacich in 1972 [1-2] is regarded as one of the most popular centrality measures. The general assumption of eigenvector centrality (EVC) is that each node's centrality in a graph is the sum of the centrality values of its neighbors [3]. It considers not only its own degree, but also the degree of the nodes that it is connected to. The nodes are eventually drawn with a radius, also referred as spectral radius [13], proportional to their centrality. Owing to the fact that it is superior to degree centrality intrinsically, EVC has been widely applied to the analysis of social network relations [4-6].

Natarajan Meghanathan et al. (Eds) : NeCoM, CSITEC - 2016
pp. 09–18, 2016. © CS & IT-CSCP 2016 DOI : 10.5121/csit.2016.61202

One often asked question raises along with the application of EVC: how is the EVC correlated to other centrality measures? It is critical to unveil the underlying relationship between EVC and other measures [7]. With this effort, we could reduce the potential redundancy in analyzing network relations with multiple similar centrality measures. It is also interesting to see the importance of EVC if it is highly unrelated to other measures. Recent literature has shown a high correlation between EVC and degree centrality with an average correlation of 0.92 on 58 networks [7]. The correlation between EVC and maximum clique size has also been examined [8]. Some others have also investigated on eigenvector centrality and it continues to be analyzed and developed [9-10]. However, there is still lack of thorough comparison of EVC to other measures on multiple types of networks.

The first part of this paper briefly illustrates how eigenvector centrality is calculated. The second part shows the results with specific comparison between EVC and other centrality measures in random network, small-world network, and multiple real-world networks.

2. EIGENVECTOR CENTRALITY CALCULATION

Adjacency matrix is used to solve the problem of eigenvector centrality measure. On the basis of the work done by Bonacich (1972) [1-2], the centrality of node i follows the form below:

$$\lambda c_i = \sum_{j=1}^{n} R_{ij} c_j$$

where R is an adjacency matrix, and λ is a constant to make the equation with a nonzero solution. The centrality c_i of a node i is thus expressed as positive multiple of the sum of adjacent centralities. In matrix notation, we then have: $\lambda c = Rc$, where c is an eigenvector of R, and λ is its associated eigenvalue. The solution to above equations is already well-known and shown in Figure 1.

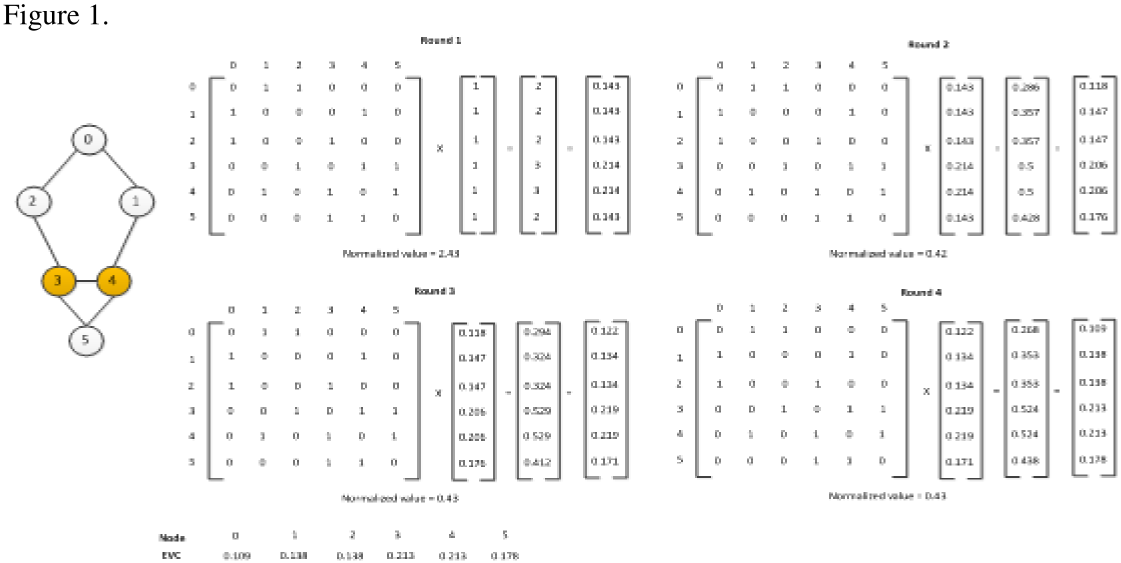

Figure 1. Illustration of the Computation of Eigenvector Centrality (EVC): Nodes 3 and 4 have the Highest EVC Value

3. ANALYSIS OF CORRELATION BETWEEN EVC AND OTHER MEASURES

3.1. Correlation Coefficient Calculation

Correlation coefficient was computed on five centrality measures over EVC on each network to estimate their correlation [10]. The correlation coefficient is a measure of linear correlation between different pairs of data. For instance, with a data pair of (x, y), we can compute its correlation coefficient $R_{x,y}$ as:

$$R_{x,y} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{n\sigma_x\sigma_y}$$

where $\bar{x}$ and $\bar{y}$ are the mean of the measurements of a centrality measure x and y respectively. The values σ_x and σ_y are the standard deviation of a centrality measure x and y respectively. The value of $R_{x,y}$ ranges from -1 to 1. The absolute value close to 1 is regarded as highly correlated, and 0 is regarded as independent.

3.2. Analysis on Random Network

Random networks were simulated to investigate the centrality measures including EVC, degree centrality (DEG), betwenness centrality (BWC), clustering coefficient centrality (CCC), farness centrality (FRC), and closeness centrality (CLC). In this section, networks with 100 nodes were generated. In addition, the probability of linkage between nodes from 0.05 to 0.9 is also involved to evaluate abovementioned centrality measures. The probability of linkage is increased from 0.05 to 0.1 by 0.01; from 0.1 to 0.9 by 0.1. Representative random networks are shown in Figure 2 with a ranking factor of EVC. Correlation between EVC and other four measures, including DEG, BWC, CCC, FRC, and CLC, was then determined. Average correlation coefficient value was calculated based on 100 trials.

As shown in Figure 3, EVC is highly correlated to BWC and DEG. Our data suggests a strong correlation between EVC and DEG, ranging from 0.8754 to 0.9995. The result is similar to the paper from Valente et al (2008) [7], which also suggested a high correlation between EVC and DEG. Additionally, there also exists a strong correlation between EVC and BWC, ranging from 0.7605 to 0.9661. Notably, it is rarely papered on such high correlation between EVC and BWC. Although there is a strong correlation between EVC, BWC and DEG, it is not the same case for CCC, CLC and FRC. It clearly shows an extremely low correlation between EVC and other three measures, with an absolute value smaller than 0.01. The result is consistent over all tested probability of linkage value.

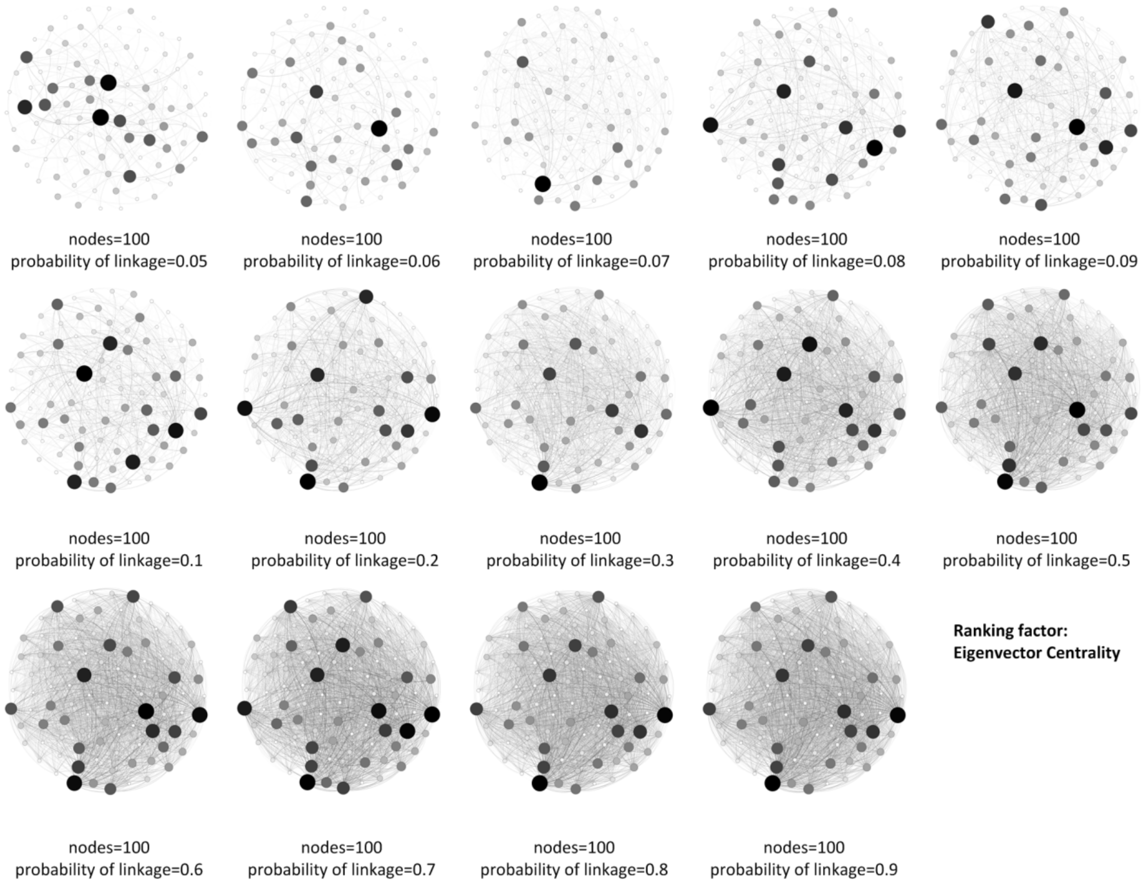

Figure 2. Simulation of Random Networks with Varying Probability of Linkage: Ranking is based on Eigenvector Centrality

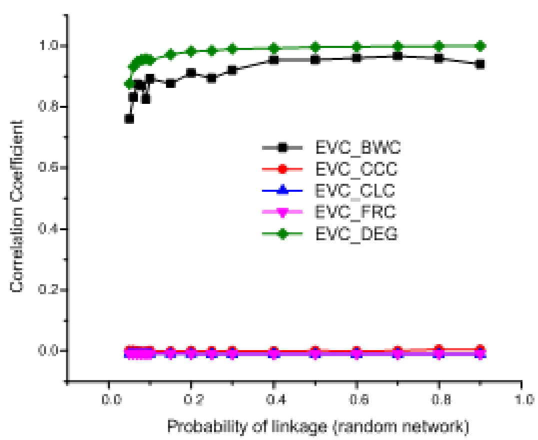

Figure 3. Correlation Coefficient between EVC and other Four Measures, including DEG, BWC, CCC, FRC, and CLC, on Random Networks with various Probability of Linkage

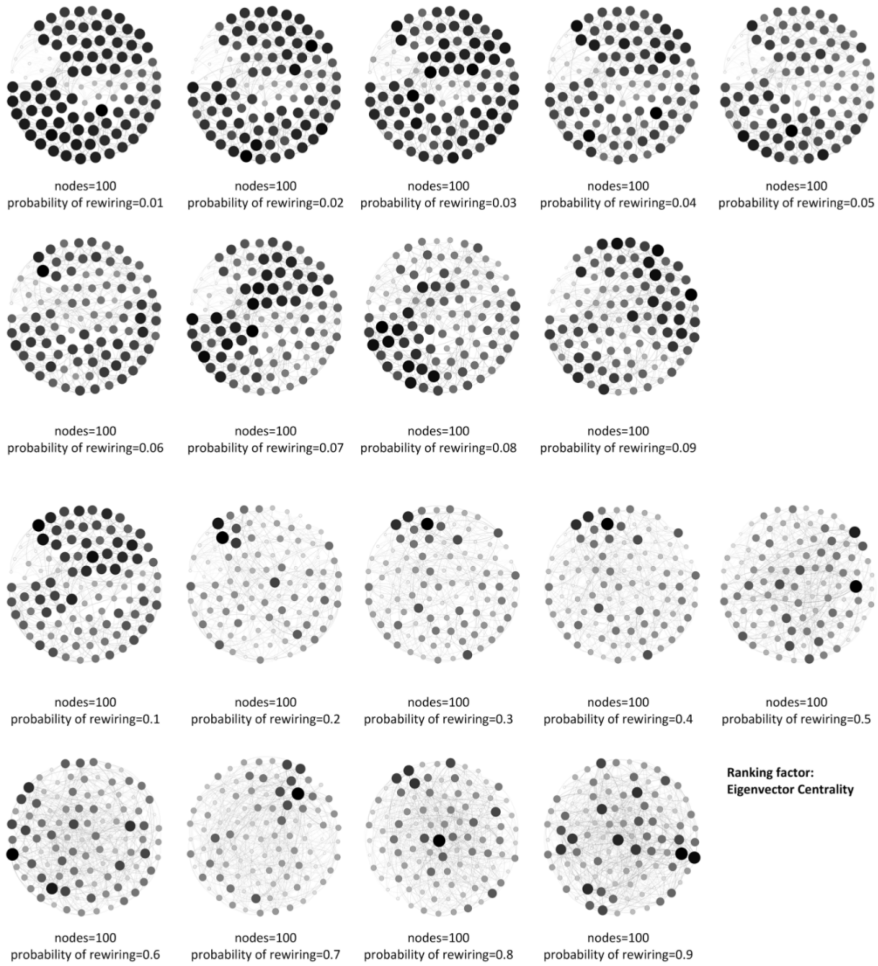

Figure 4. Simulation of Small-World Networks with various Probability of Rewiring. Ranking Factor is Eigenvector Centrality

3.3. Analysis on Small-World Network

Moreover, we also investigated on small-world networks evolved from regular network. Similar to random network simulation, 100 nodes with a k-regular value (initial number of links per node) of 10 are set for small-world network simulation. In this section, the probability of rewiring was from 0.01 to 0.09 with increment of 0.01; and from 0.1 to 0.9 with increment of 0.1. Representative random networks are shown in Figure 4 with a ranking factor of EVC. Correlation between EVC and other four measures, including DEG, BWC, CCC, FRC, and CLC, was then determined. Average correlation coefficient value was calculated based on 100 trials.

On small-world networks, there still presents a strong correlation between DEG and EVC. The correlation coefficient was larger than 0.71 when the probability of rewiring reaches 0.1. On our previous paper, a transformation between small-world network and random network was revealed [11]. It was found that simulated network from a regular network would be small-world network when the probability of rewiring is from 0.01 to 0.1; however, it changes to random network when the probability of rewiring is between 0.1 and 1.0 [13]. On the basis of this fact, the high correlation between DEG and EVC with a probability of rewiring value of 0.1 to 0.9 in Figure 5 can be explained and is in agreement with Figure 3. However, the correlation between EVC and BWC is not consistent on small-world network. Overall, it is relatively low in terms of correlation between EVC and BWC with a value less than 0.1 when probability of rewiring reaches 0.5. It is noted that there is a clear consistency on the low correlation between EVC and other three measures, including CCC, CLC and FRC, which is similar to random network.

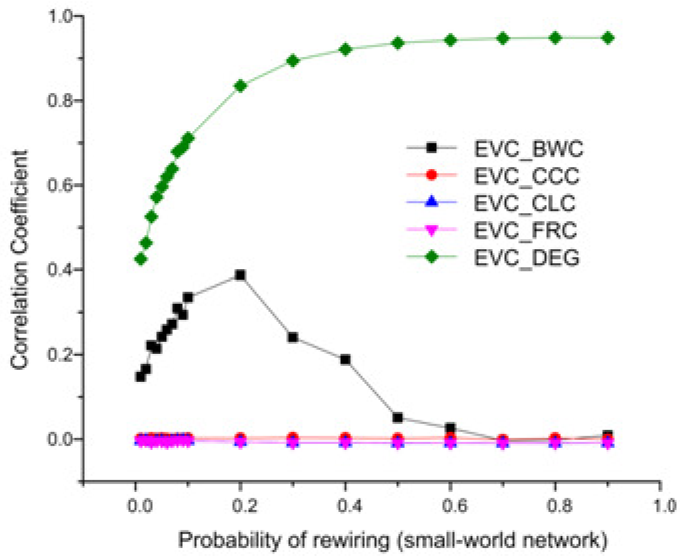

Figure 5. Correlation Coefficient between EVC and other Four Measures, including DEG, BWC, CCC, FRC, and CLC, on Small-World Networks with various Probability of Rewiring

3.4. Analysis on Real-World Network

Finally, multiple real-world networks are involved in our analysis for further investigation. Analysis on real-world networks is crucial to understanding how EVC relates to other measures in real world. Here we selected nine real-world networks (see Figure 6), including dolphins social network (Dolphins) [14], WordAdj Adjacency network of common adjectives and nouns in the novel David Copperfield by Charles Dickens (WordAdj) [15], Celegensmetabolic Network representing the metabolic network of C. elegans (Celegm), Celegensneural Network representing the neural network of C. elegans (Celegn) [16], American football games network between Division IA colleges during regular season Fall 2000 (Football) [17], Karate Social network of friendships between 34 members of a karate club at a US university in the 1970 (Karate) [18], LesMis Coappearance network of characters in the novel Les Miserables (LesMis) [19], US Airports network (AirNet) [20], and political books network (BookNet) [21]. Average correlation between EVC and other four measures, including DEG, BWC, CCC, FRC, and CLC, was determined on 100 trials.

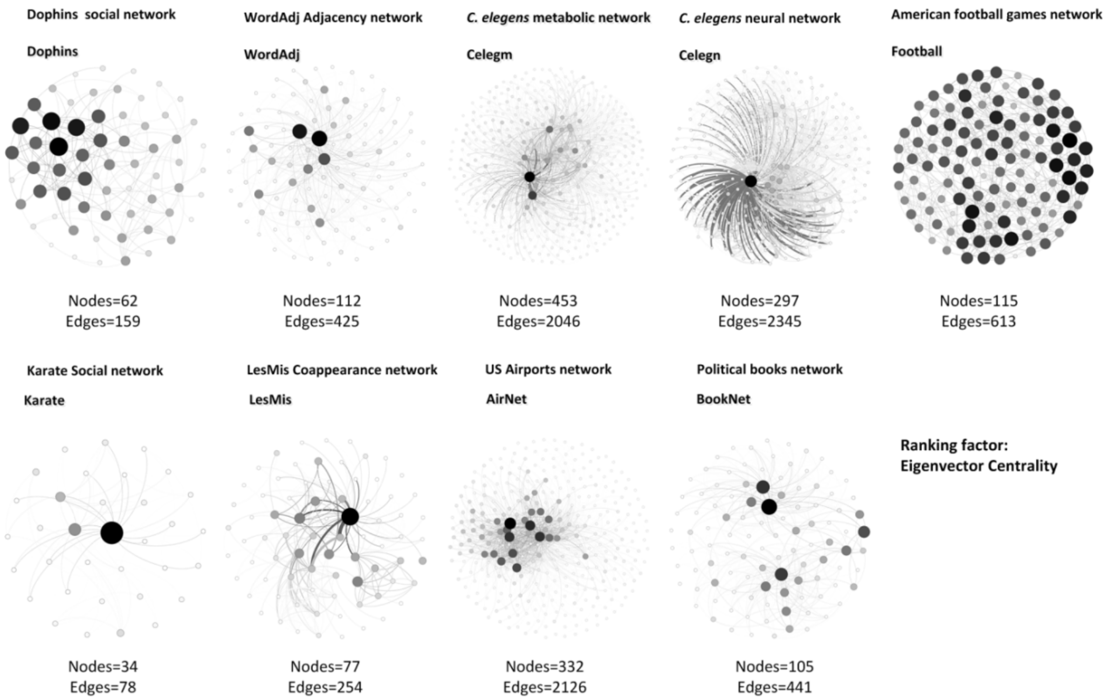

Figure 6. Real-World Networks Distribution with Ranking Factor of Eigenvector Centrality

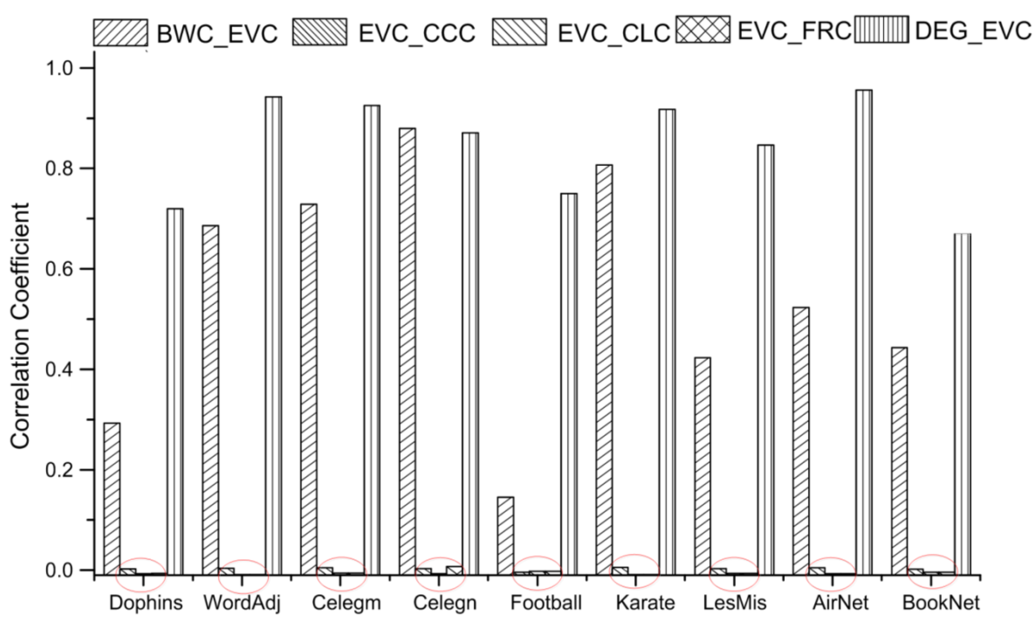

Figure 7. Correlation Coefficient between EVC and other Four Measures, including DEG, BWC, CCC, FRC, and CLC, on Real-World Networks

Similar to random and small-world network, the correlation of EVC to CCC, CLC, and FRC is close to zero. This result supports our previous statement that EVC is independent from CCC, CLC and FRC on any tested network. It is noticed that EVC is highly correlated to DEG with a correlation coefficient over 0.66. In particular, the correlation coefficient is over 0.91 on four real-world networks out of nine, including WordAdj, Celegm, Karate, and AirNet; it is over 0.71 on eight real-world networks out of nine. Furthermore, similar to small-world network, the correlation between EVC and BWC is not consistent on all real-world networks. Data shows a high correlation on Celegm, Celegn, and Karate networks with a value over 0.72; however it also presents a low correlation on Dophins and Football networks with a value less than 0.3. Intriguingly, we find that the ones with low correlation coefficient are not directed (Dophins and Football); however, all directed networks shows a high correlation between EVC and BWC. This suggestion is also supported by Valente et al (2008) [7].

3.5. Overall Discussion

Our data suggest that EVC is highly correlated to DEG, probably owing to the fact that both measures are symmetric. We also find relatively strong but varied correlation between EVC and BWC on random network and real-world networks. The high correlation between EVC and DEG revealed on all tested networks would suggest redundancy over EVC to DEG. In addition, there could also be a redundancy over EVC to BWC on undirected networks. However, the inconsistent result of the correlation between EVC and BWC on small-world network could be due to its intrinsic power-law distribution [12], differ from those of the regular and random networks [10]. It is also found that the BWC follows a power-law distribution [10]. Thus, the inconsistency could be well explained. Lastly, the analysis on the correlation of EVC to CCC, CLC and FRC indicates that they are distinct to EVC.

4. CONCLUSIONS

In this paper, in order to investigate correlations of EVC to other five measures, we applied correlation coefficient analysis on various types of networks, including random network, small-world network, as well as multiple real-world networks. We found that EVC was strongly correlated with DEG, and the correlation was robust in the sense that the extent of correlation was little affected by the types of the network, particularly directed network. The finding on the correlation between EVC and BWC suggests they are independent on a network with power-law distribution. With all tested networks, EVC is independent from CCC, CLC and FRC. This finding has not been papered so far and could be helpful in understanding different characteristics of networks. All findings in this paper can be used to guide our future research on correlation analysis among centrality measures on various networks.

REFERENCES

[1] P. Bonacich, "Technique for Analyzing Overlapping Memberships," Sociological Methodology, vol. 4, pp. 176-185, 1972.

[2] P. Bonacich, "Factoring and Weighting Approaches to Status Scores and Clique Identification," Journal of Mathematical Sociology, vol. 2, no. 1, pp. 113-120, 1972.

[3] P. Bonacich, "Some Unique Properties of Eigenvector Centrality," Social Networks, vol. 29, no. 4, pp. 555-564, 2007.

[4] B. Ruhnau, "Eigenvector Centrality: A Node Centrality?," Social Networks, vol. 22, no. 4, pp. 357-365, October 2000.

[5] P. Bonacich, and P. Lloyd, "Eigenvector-like Measures of Centrality for Asymmetric Relations," Social Networks, vol. 23, no. 3, pp. 191-201, 2001.

[6] N. Meghanathan, "Correlation Coefficient Analysis: Centrality vs. Maximal Clique Size for Complex Real-World Network Graphs," International Journal of Network Science, vol. 1, no. 1, pp. 3-27, 2016.

[7] T. W. Valente, K. Coronges, C. Lakon and E. Costenbader, "How Correlated are Network Centrality Measures?," Connect (Tor), vol. 28, no. 1, pp. 16-26, January 2008.

[8] N. Meghanathan, "Correlation Analysis between Maximal Clique Size and Centrality Metrics for Random Networks and Scale-Free Networks," Computer and Information Science, vol. 9, no. 2, pp. 41-57, May 2016.

[9] J. M. Pappas and B. Wooldridge, "Middle Managers' Divergent Strategic Activity: An Investigation of Multiple Measures of Network Centrality," Journal of Management Studies, vol. 44, no. 3, pp. 323-341, May 2007.

[10] C.-Y. Lee, "Correlations among Centrality Measures for Complex Networks," arXiv:physics/ 0605220 [physics.soc-ph], pp. 1-18, May 2006.

[11] N. Meghanathan, "Exploiting the Discriminating Power of the Eigenvector Centrality Measure to Detect Graph Isomorphism," International Journal in Foundations of Computer Science and Technology, vol. 5, no. 6, pp. 1-13, November 2015.

[12] D. J. Watts and S. H. Strogatz, "Collective Dynamics of 'Small-World' Networks," Nature, vol. 393, pp. 440-442, June 1998.

[13] N. Meghanathan, "Using Spectral Radius Ratio for Node Degree to Analyze the Evolution of Complex Networks," International Journal of Computer Networks and Communications, vol. 7, no. 3, pp. 1-12, May 2015.

[14] D. Lusseau, D, K. Schneider, O. J. Boisseau, P. Hasse, E. Slooten, and S. M. Dawson, "The Bottlenose Dolphin Community of Doubtful Sound Features a Large Proportion of Long-lasting Associations," Behavioral Ecology and Sociobiology, vol. 54, no. 3, pp. 396-405, 2003.

[15] M. Newman, "Finding Community Structure in Networks using the Eigenvectors of Matrices," Physical Review E, vol. 74, no. 3, 036104, 2006.

[16] J. G. White, E. Southgate, J. N. Thomson and S. Brenner, "The Structure of the Nervous System of the Nematode Caenorhabditis Elegans," Philosophical Transactions B, vol. 314, no. 1165, pp. 1-340, 1986.

[17] M. Newman, M and M. Girvan, "Mixing Patterns and Community Structure in Networks," Statistical Mechanics of Complex Networks: Lecture Notes in Physics, vol. 625, pp. 66-87, 2003.

[18] Zachary, W. W. (1977). An Information Flow Model for Conflict and Fission in Small Groups. Journal of Anthropological Research, vol. 33, no. 4, pp. 452-473, 1977.

[19] D. E. Knuth, The Stanford GraphBase: A Platform for Combinatorial Computing, 1st Edition, Addison-Wesley, 1993.

[20] V. Batagelj and A. Mrvar, Pajek Datasets. http://vlado.fmf.uni-lj.si/pub/networks/data/, 2006.

[21] V. Krebs, "Proxy Networks: Analyzing One Network to Reveal Another," Bulletin de Méthodologie Sociologique, vol. 79, pp. 61-70, 2003.