Journal of King Saud University – Computer and Information Sciences 30 (2018) 391–403

Contents lists available at ScienceDirect

Journal of King Saud University – Computer and Information Sciences

journal homepage: www.sciencedirect.com

δ-Space for real-world networks: A correlation analysis of decay centrality vs. degree centrality and closeness centrality

Natarajan Meghanathan

Jackson State University, Jackson, MS 39217, USA

ARTICLE INFO

Article history:
Received 3 January 2017
Revised 9 April 2017
Accepted 21 April 2017
Available online 3 May 2017

Keywords:
Decay centrality
Decay parameter
Closeness centrality
Degree centrality
Correlation
Real-world network graphs

ABSTRACT

We analyze a suite of 48 real-world networks and compute the decay centrality (DEC) of the vertices for the complete range of values for the decay parameter $\delta \in (0, 1)$ as well as determine the Pearson's correlation coefficient (PCC) between the DEC_δ values and degree centrality (DEG) and closeness centrality (CLC). We observe PCC(DEC_δ, DEG) to decrease with increase in δ and PCC(DEC_δ, CLC) to decrease with decrease in δ. We define the δ-$space_r$ for a real-world network with respect to the DEG, DEC, CLC correlation as the difference between the maximum and minimum δ values under which we observe a particular level of correlation (r) between the DEC, DEG and DEC, CLC metrics respectively. We show that the PCC(DEG, CLC) values for the real-world networks exhibit a very strongly positive correlation with the δ-$space_r$ values and demonstrate that one could predict the δ-$space_r$ value for a real-world network using the PCC(DEG, CLC) value for that network. We also analyze the impact of various topological measures on the δ-$space_r$ values for the real-world networks.

1. Introduction

The Decay Centrality (DEC) metric is a parameter-driven centrality metric that has not been explored much in the literature for complex network analysis. Decay centrality is a measure of the closeness of a node to the rest of the nodes in the network (Jackson, 2010). However, unlike closeness centrality (CLC) Freeman, 1979, the importance given to the geodesic distance (typically, in terms of the number of hops if the edges do not have weights) is weighted in terms of a parameter called the decay parameter δ ($0 < \delta < 1$). The formulation for computing the decay centrality of a vertex v_i for a particular value of the decay parameter δ is (Jackson, 2010): $DEC(v_i) = \sum_{v_i \neq v_j} \delta^{d(v_i, v_j)}$ where $d(v_i, v_j)$ is the distance from node v_i to node v_j. The decay parameter δ essentially controls how important is a node v_j to a node v_i ($v_i \neq v_j$) that are at a distance $d(v_i, v_j)$ from each other. If δ is smaller, the distance to the nearby nodes is weighted relatively larger than the distance to the nodes farther away. If δ is larger, the distance to every node is given almost the same importance. As a result, if δ is closer to 0, the decay centrality of the vertices is more likely to exhibit a very strong positive correlation with the degree centrality of the vertices; if δ is closer to 1, the decay centrality of the vertices is more likely to exhibit a very strong positive correlation with the closeness centrality of the vertices. We adopt the ordinal range of values proposed by Evans (1995) and consider two centrality metrics to exhibit strongly positive ($s+$) and very strongly positive ($vs+$) correlation if the Pearson's correlation coefficients (PCC) (Lay et al., 2015) computed on the basis of the values incurred for the two metrics are respectively 0.6 or above and 0.8 or above. As part of our correlation analysis, we analyze a suite of 48 real-world networks whose spectral radius ratio for node degree (Meghanathan, 2014) (a measure of variation in node degree) ranges from 1.01 to 5.51.

The motivation for our research came from the initial results of our correlation study which indicated that the Pearson's correlation coefficient PCC(DEC_δ, DEG) decreases with increase in δ from 0.01 to 0.99 and PCC(DEC_δ, CLC) decreases with decrease in δ. Because of such a trend, we came up with a hypothesis that there could exist a range of δ values (called the δ-*space*) for which we could observe the DEC_δ values to exhibit a particular level of correlation (we focus on strongly and very strongly positive levels of correlation) simultaneously with both the DEG and CLC metrics. In this pursuit, for each real-world network, we identified the maximum δ value (indicated as $\delta^{DEC,DEG}_{\max,r \equiv s+}$ or equivalently as $\delta^{DEC,DEG}_{\max,r \geqslant 0.6}$ and

Peer review under responsibility of King Saud University.

Production and hosting by Elsevier

E-mail address: natarajan.meghanathan@jsums.edu

http://dx.doi.org/10.1016/j.jksuci.2017.04.006
1319-1578/

$\delta^{DEC,DEG}_{\max,r\equiv vs+}$ or equivalently as $\delta^{DEC,DEG}_{\max,r\geqslant 0.8}$) until which PCC($DEC_\delta$, DEG) is 0.6 or above (for strongly positive correlation) or 0.8 or above (for very strongly positive correlation), as well as identified the minimum δ value (indicated as $\delta^{DEC,CLC}_{\min,r\equiv s+}$ or equivalently as $\delta^{DEC,CLC}_{\min,r\geqslant 0.6}$ and $\delta^{DEC,CLC}_{\min,r\equiv vs+}$ or equivalently as $\delta^{DEC,CLC}_{\min,r\geqslant 0.8}$) starting from which PCC(DEC_δ, CLC) continues to be 0.6 or above or 0.8 or above respectively. For real-world networks with $\delta^{DEC,CLC}_{\min,r} \leqslant \delta^{DEC,DEG}_{\max,r}$, there exists a range of δ values $\delta^{DEC,CLC}_{\min,r} \ldots \delta^{DEC,DEG}_{\max,r}$ (quantified and called the *δ-space$_r$*: $\delta^{DEC,DEG}_{\max,r} - \delta^{DEC,CLC}_{\min,r} + \in$; where $\in$ is the level of precision used for δ; for more details see Section 5) under which DEC_δ would exhibit a particular level of correlation ($r \equiv s$+ or *vs*+) with respect to both DEG and CLC. Statistically, the *δ-space$_r$* for a real-world network with respect to a particular level of correlation (r) would correspond to the probability with which the decay centrality metric (computed for a randomly chosen value of δ) could exhibit the particular level of correlation with both the degree centrality and closeness centrality metrics. We hypothesize that the Pearson's correlation coefficient between DEG and CLC for a real-world network is likely to be very strongly correlated to the *δ-space$_r$* values and that one could predict the *δ-space$_r$* value for a network using the PCC(DEC, CLC) observed for that network.

The rest of the paper is organized as follows: Section 2 discusses related work. In Section 3, we review the centrality metrics (DEG, CLC and DEC) and the Pearson's correlation measure as well as explain their computation with an example graph. Section 4 first introduces the notion of *δ-space* for DEC, DEG and DEC, CLC correlation (hereafter referred to as DEG-DEC-CLC correlation) and its computation on the example graph of Section 3. Section 5 introduces the real-world networks that are analyzed in this paper. Section 6 first presents the results of correlation study involving DEC, DEG and CLC and the notion of *δ-space$_r$*. Section 6 then presents the simulation results to corroborate our hypothesis about the relationship between *δ-space$_r$* and the Pearson's correlation coefficient for DEG and CLC. Section 6 also analyzes the impact of various topological measures on the *δ-space$_r$* values (negative values and the highest positive value of 0.99) incurred for the real-world networks. Section 7 concludes the paper.

2. Related work

Decay centrality has not been explored much in the literature for complex network analysis. To the best of our knowledge, ours is the first work to conduct a correlation study focusing on decay centrality. Most of the work (e.g., Li et al., 2015; Meghanathan, 2015) on correlation studies (involving centrality metrics) were focused on the commonly studied centrality metrics such as the neighborhood-based degree centrality and eigenvector centrality (Bonacich, 1987) and shortest path-based betweenness centrality (Freeman, 1977) and closeness centrality. The objective of such correlation studies has been typically to identify computationally-light alternatives (like DEG and its derivatives (Meghanathan, 2017) for computationally-heavy metrics (such as EVC and BWC) for both real-world networks and simulated networks of theoretical models (Renyi, 1959; Barabasi and Albert, 1999). The focus of our paper is different from such typical correlation studies in the literature. We seek to explore the trend of change in the correlation coefficients between a parameter-driven centrality metric (whose values for a node change for different values of the decay parameter) and the degree and closeness centrality metrics whose values are not parameter-driven and remain the same for a particular network.

The most related work to our work is a recent study (Tsakas, 2016) on random networks (Renyi, 1959) for which a single threshold value of the decay parameter (referred here as δ_{thresh}) was observed to exist (for a particular operating condition) such that nodes with high degree centrality also had a high decay centrality computed for δ values less than δ_{thresh} and nodes with high closeness centrality also had a high decay centrality computed for δ values above δ_{thresh}. It was observed in Tsakas (2016) that for random networks: nodes with the largest values for degree centrality and closeness centrality are more likely to be nodes that also incur the largest values for decay centrality for almost all values of δ. In addition, nodes that had the largest decay centrality for a certain value of δ are more likely to be part of the set of nodes that had the largest degree centrality or the largest closeness centrality. The likelihood of all of the above was studied using multinomial logistic regression (Greene, 2011).

In (Dangalchev, 2006), Dangalchev proposed a variant of closeness centrality metric (to quantify the vulnerability of networks to get disconnected) that is essentially the decay centrality of the vertices computed for δ = 0.5. However, there was no correlation analysis reported between Dangalchev's closeness centrality metric and the decay centrality of the vertices for different values of δ. Most of the other works (e.g., Chatterjee and Dutta, 2016; Kang et al., 2012) on decay centrality metric have focused on exploring its suitability for diffusion in socio-economic networks with regards to selecting the seed nodes that could effectively propagate information about a product to putative customers. Nodes that are themselves central and connected to other central nodes (via direct links or shorter paths) in the network are typically preferred for such "agent" roles (Tsakas, 2016; Chatterjee and Dutta, 2016). The use of decay centrality vis-a-vis diffusion centrality (Kang et al., 2012) and eigenvector centrality (Ide et al., 2014; Banerjee et al., 2013) to identify such "agent" nodes for diffusion has been explored in the literature.

3. Review of centrality metrics and Pearson's correlation measure

The centrality metrics that are of interest in this research are degree centrality (DEG), closeness centrality (CLC) and decay centrality (DEC). In this section, we briefly review these three metrics and their computation using a running example graph as well as review the Pearson's correlation measure and its computation with respect to the DEG and CLC metrics for the running example graph.

3.1. Degree centrality

The degree centrality (DEG) of a vertex is the number of neighbors incident on the vertex. Fig. 1 illustrates the degree centrality of the vertices (listed above the vertices) in the example graph used in Sections 3 and 4. A key weakness of the degree centrality metric is that the metric can take only integer values (though, weighted degree centrality can take on any real value) and ties among vertices (with same degree) is quite common and unavoidable in network graphs of any size (in the graph of Fig. 1, we

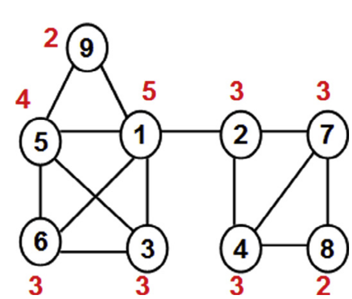


Fig. 1. Degree Centrality of the Vertices in an Example Graph.

observe five of the nine vertices to have a degree of 3). If $d(v_i, v_j)$ is the geodesic distance between two vertices v_i and v_j in a graph, then the degree centrality of a vertex v_i could be computed as the sum of the distances to the one-hop neighbors.

$$DEG(v_i) = \sum_{\substack{v_j \neq v_i \\ d(v_i, v_j)=1}} d(v_i, v_j) \tag{1}$$

3.2. Closeness centrality

The closeness centrality (CLC) of a vertex Freeman, 1979 is a measure of the closeness of the vertex to the rest of the vertices in a graph. The CLC of a vertex is computed as the inverse of the sum of the hop counts (farness) of the shortest paths from the vertex to the rest of the vertices in the graph. If $d(v_i, v_j)$ is the geodesic distance between two vertices v_i and v_j in a graph, then the closeness centrality of a vertex v_i could be computed as the sum of the geodesic distances to the vertices v_j that are in the same component as v_i (a component is the largest set of vertices that are reachable from each other (Cormen et al., 2009). We say that two vertices v_i and v_j are in the same component if the geodesic distance $d(v_i, v_j)$ between them is not infinity.

$$CLC(v_i) = \frac{1}{\sum_{\substack{v_j \neq v_i \\ d(v_i, v_j) \neq \infty}} d(v_i, v_j)} \tag{2}$$

Fig. 2 illustrates the distance matrix (hop counts of the shortest paths between any two vertices) for the example graph of Fig. 1 and also displays the CLC of the vertices. Vertex 1 is the closest vertex to the rest of the vertices (sum of the distances is 12, the minimum) and hence has the largest CLC value of 1/12 = 0.083.

3.3. Decay centrality

Decay centrality (DEC) is a measure of the closeness of a node to the rest of the nodes in the network (Jackson, 2010). However, unlike closeness centrality, the importance given to the distance (typically, in terms of the number of hops if the edges do not have weights) is weighted in terms of a parameter called the decay parameter δ ($0 < \delta < 1$). The formulation for computing the decay centrality of a vertex v_i for a particular value of the decay parameter δ is (Jackson, 2010): $DEC_\delta(v_i) = \sum_{v_i \neq v_j} \delta^{d(v_i, v_j)}$ where $d(v_i, v_j)$ is the distance from node v_i to node v_j. The decay parameter δ essentially controls how important is a node v_j to a node v_i ($v_i \neq v_j$) that are at a distance $d(v_i, v_j)$ from each other. Nodes that have a higher decay centrality are more likely to be nodes that have several neighbors as well as be much closer to the rest of the nodes in the network (Tsakas, 2016). Note that we did not choose the closed interval [0,1] for δ, as $\delta = 1$ would correspond to the component size (one less than the number of vertices in the component to which the vertex belongs to) and not quantify the centrality of the vertices, and $\delta = 0$ would make the decay centrality of the vertices to become zero. Like the case of classical Dijkstra algorithm (Cormen et al., 2009) and other shortest path-based algorithms (Cormen et al., 2009), we adopt the convention that the distance between two vertices v_i and v_j that are in two different components of a graph is not defined and is treated as infinity (∞) for all quantification purposes. As the δ values are in the range (0, 1), $\delta^{d(v_i, v_j)}$ tends to 0 for $d(v_i, v_j) = \infty$. Hence, if a network comprises of two or more components, the decay centrality of a vertex is computed based on the distances to the rest of the vertices within the component.

Fig. 3 presents the decay centrality of the vertices in the example graph of Section 3 for different values of the decay parameter δ. We also illustrate sample calculations of the decay centrality of vertex 1 for three different values of δ. One can observe from Fig. 3, the magnitudes of the values for decay centrality of the vertices are very close to each other as δ approaches 1. The results of our correlation analysis in Section 6 indicate that the subtle differences (as long as $\delta < 1$) in the decay centrality of the vertices are sufficient enough to observe a very strongly positive correlation between decay centrality and closeness centrality for larger values of δ.

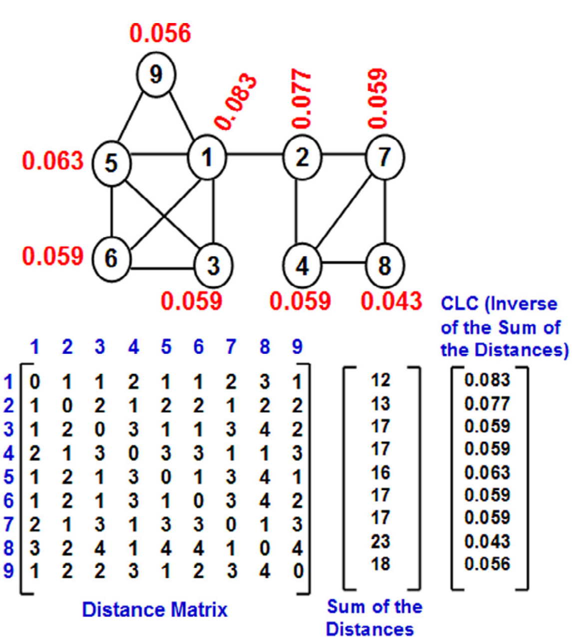


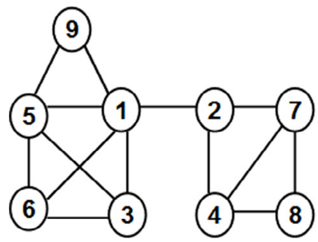
Fig. 2. Closeness Centrality of the Vertices in an Example Graph.

3.4. Pearson's correlation measure

We use the Pearson's correlation coefficient (PCC) Lay et al., 2015 as the measure for analyzing the correlation between the decay centrality (computed for different values of the decay parameter δ) and the degree centrality and closeness centrality. The Pearson's product-moment correlation when applied for centrality metrics is a measure of the linear dependence between any two metrics in consideration (Lay et al., 2015; Meghanathan, 2017). It is referred to as the product-moment based correlation as we calculate the deviation of the data points from their mean value ('mean' is also referred to as 'first moment' in statistics) and use them in the formulation below to calculate the correlation coefficient. If X and Y are the datasets for two centrality metrics: let X_i and Y_i indicate the centrality values for the individual vertices v_i ($1 \le i \le n$, where n is the number of vertices) and $\overline{X}$ and $\overline{Y}$ are the average of the centrality values; PCC(X, Y) is calculated as follows.

$$PCC(X, Y) = \frac{\sum_{i=1}^{n} (X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum_{i=1}^{n} (X_i - \overline{X})^2} \sqrt{\sum_{i=1}^{n} (Y_i - \overline{Y})^2}} \tag{3}$$

Fig. 4 illustrates the computation of the Pearson's correlation coefficient between the degree centrality and closeness centrality. In Table 1, we show the range of correlation coefficient values advocated by Evans (1995) for an ordinal classification of the strength of correlation. We adopt this range in our paper and as shown in Table 1, we consider two centrality metrics to exhibit a strongly positive correlation and very strongly positive correlation if the correlation coefficient values are 0.6 or above and 0.8 or above respectively.



Distance Matrix

	1	2	3	4	5	6	7	8	9
1	0	1	1	2	1	1	2	3	1
2	1	0	2	1	2	2	1	2	2
3	1	2	0	3	1	1	3	4	2
4	2	1	3	0	3	3	1	1	3
5	1	2	1	3	0	1	3	4	1
6	1	2	1	3	1	0	3	4	2
7	2	1	3	1	3	3	0	1	3
8	3	2	4	1	4	4	1	0	4
9	1	2	2	3	1	2	0	3	4

$DEC_{0.05}(1) = 0.05^1 + 0.05^1 + 0.05^2 + 0.05^1 + 0.05^1 + 0.05^2 + 0.05^3 + 0.05^1 = 0.2551$

$DEC_{0.25}(1) = 0.25^1 + 0.25^1 + 0.25^2 + 0.25^1 + 0.25^1 + 0.25^2 + 0.25^3 + 0.25^1 = 1.3906$

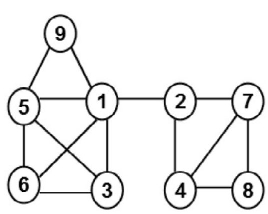
$DEC_{0.95}(1) = 0.95^1 + 0.95^1 + 0.95^2 + 0.95^1 + 0.95^1 + 0.95^2 + 0.95^3 + 0.95^1 = 7.4124$

Vertices

δ	1	2	3	4	5	6	7	8	9
0.05	0.255	0.163	0.155	0.153	0.203	0.155	0.153	0.103	0.108
0.15	0.798	0.563	0.502	0.486	0.629	0.502	0.486	0.328	0.375
0.25	1.391	1.063	0.910	0.875	1.098	0.910	0.875	0.594	0.723
0.35	2.038	1.663	1.396	1.344	1.623	1.396	1.344	0.925	1.168
0.45	2.746	2.363	1.978	1.917	2.226	1.978	1.917	1.358	1.731
0.55	3.521	3.162	2.679	2.618	2.927	2.679	2.618	1.935	2.432
0.65	4.369	4.062	3.522	3.471	3.750	3.523	3.471	2.711	3.295
0.75	5.297	5.062	4.535	4.500	4.723	4.535	4.500	3.750	4.348
0.85	6.309	6.162	5.745	5.729	5.873	5.745	5.729	5.125	5.618
0.95	7.412	7.362	7.184	7.182	7.232	7.184	7.182	6.918	7.137

Decay Centrality Values

Fig. 3. Decay Centrality of the Vertices in an Example Graph.



$$PCC(DEG, CLC) = \frac{0.06800}{\sqrt{6.889} * \sqrt{0.0011}}$$

$$PCC(DEG, CLC) = 0.78$$

Vertex, i	DEG	DEG_i - Avg(DEG)	$(DEG_i$ - Avg(DEG)$)^2$	CLC	CLC_i - Avg(CLC)	$(CLC_i$ - Avg(CLC)$)^2$	(DEG(i) - Avg(DEG)) * (CLC_i - Avg(CLC))
1	5	1.89	3.5721	0.083	0.021	0.000441	0.03969
2	3	-0.11	0.0121	0.077	0.015	0.000225	-0.00165
3	3	-0.11	0.0121	0.059	-0.003	0.000009	0.00033
4	3	-0.11	0.0121	0.059	-0.003	0.000009	0.00033
5	4	0.89	0.7921	0.063	0.001	0.000001	0.00089
6	3	-0.11	0.0121	0.059	-0.003	0.000009	0.00033
7	3	-0.11	0.0121	0.059	-0.003	0.000009	0.00033
8	2	-1.11	1.2321	0.043	-0.019	0.000361	0.02109
9	2	-1.11	1.2321	0.056	-0.006	0.000036	0.00666
	Avg = 3.11		Sum = 6.8889	Avg = 0.062		Sum = 0.001100	Sum = 0.06800

Fig. 4. Sample Illustration of the Computation of the Pearson's Correlation Coefficient between Degree Centrality and Closeness Centrality.

Table 1
Range of Correlation Coefficient Values for the Level of Correlation (adapted from Evans Evans, 1995).

Range of Correlation Coefficient	Level of Correlation	Range of Correlation Coefficient	Level of Correlation
0.80 to 1.00	Very Strong Positive	−1.00 to −0.80	Very Strong Negative
0.60 to 0.79	Strong Positive	−0.79 to −0.60	Strong Negative
0.40 to 0.59	Moderate Positive	−0.59 to −0.40	Moderate Negative
0.20 to 0.39	Weak Positive	−0.39 to −0.20	Weak Negative
0.00 to 0.19	Very Weak Positive	−0.19 to −0.01	Very Weak Negative

4. δ-space for DEG-DEC–CLC correlation and hypothesis

Fig. 5 displays the Pearson's correlation coefficient values between each of the two centrality metrics (DEG and CLC) and the decay centrality values (DEC_δ) computed for different values of the decay parameter δ ranging from 0.01 to 0.99 for the example graph of Figs. 1–4. We see PCC(DEC_δ, DEG) to monotonically decrease with increase in δ and PCC(DEC_δ, CLC) to monotonically decrease with decrease in δ. A similar trend is also noticed for all the 48 real-world network graphs analyzed in Section 6. Using this as the basis, we define the δ-$space_r$ for a real-world network with respect to a particular level of correlation (r) as the difference between the maximum δ value for which we observe the DEC, DEG correlation at the particular level and the minimum δ value for which we observe the DEC, CLC correlation at the same level.

δ-$space_r$ for a particular level of correlation (r) basically quantifies the range of values in the closed interval [$\delta^{DEC,CLC}_{\min,r}$, $\delta^{DEC,DEG}_{\max,r}$] that could be chosen from the open interval (0, 1) to determine decay centrality values that exhibit the particular level of correlation (r) with both degree and closeness centralities. Quantitatively, δ-$space_r$ is defined as follows (see formulations 4 and 5 below), where ε corresponds to the level of precision used for δ in the range (0, 1). Note that δ-$space_r$ could be determined for any value of the Pearson's correlation coefficient or level of correlation of interest. In this paper, ε = 0.01 and $r \in$ ($vs + \equiv \delta \geq 0.8$ or $s + \equiv \delta \geq 0.6$).

$$\delta - space_r = \delta^{DEC,DEG}_{\max,r} - \delta^{DEC,CLC}_{\min,r} + \varepsilon \quad \text{if } \delta^{DEC,CLC}_{\min,r} \leqslant \delta^{DEC,DEG}_{\max,r} \tag{4}$$

$$\delta - space_r = \delta^{DEC,DEG}_{\max,r} - \delta^{DEC,CLC}_{\min,r} \quad \text{if } \delta^{DEC,CLC}_{\min,r} > \delta^{DEC,DEG}_{\max,r} \tag{5}$$

Note that if $\delta^{DEC,CLC}_{\min,r} = \delta^{DEC,DEG}_{\max,r}$, it implies there is one δ value (δ = $\delta^{DEC,CLC}_{\min,r} = \delta^{DEC,DEG}_{\max,r}$) for which DEC_δ would exhibit the targeted level of correlation (r) with both DEG and CLC. If $\delta^{DEC,CLC}_{\min,r} < \delta^{DEC,DEG}_{\max,r}$, there

δ	PCC(DEC$_δ$, DEG)	PCC(DEC$_δ$, CLC)
0.01	0.9999	0.7910
0.02	0.9996	0.7990
0.03	0.9991	0.8068
0.05	0.9975	0.8215
0.15	0.9809	0.8838
0.25	0.9549	0.9262
0.35	0.9295	0.9543
0.45	0.9037	0.9713
0.55	0.8694	0.9804
0.65	0.8372	0.9843
0.75	0.8128	0.9845
0.80	0.8014	0.9852
0.81	0.7992	0.9854
0.85	0.7909	0.9864
0.95	0.7713	0.9886
0.99	0.7639	0.9892

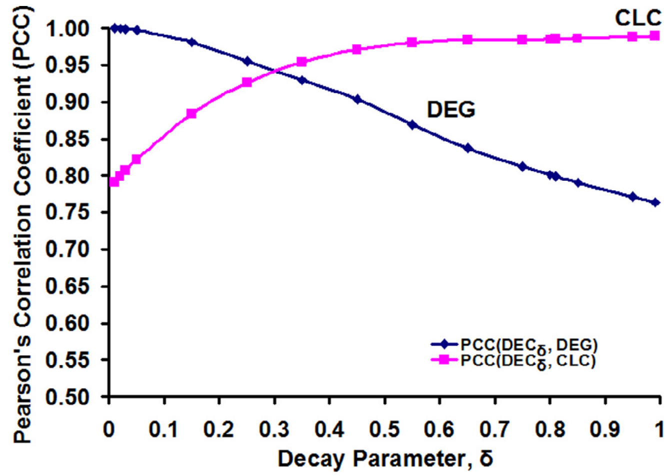


Fig. 5. Distribution of the Pearson's Correlation Coefficient Values between Decay Centrality and the Two Centrality Metrics (Degree Centrality and Closeness Centrality) vs. the Decay Parameter δ for the Example Graph of Figs. 1–4.

might be more than one δ value that could be chosen from the closed interval [$\delta^{DEC,CLC}_{\min,r}$, $\delta^{DEC,DEG}_{\max,r}$]; hence, we add the precision level ε in the formulation for δ-*space*$_r$ when $\delta^{DEC,CLC}_{\min,r} \leqslant \delta^{DEC,DEG}_{\max,r}$. On the other hand, if $\delta^{DEC,CLC}_{\min,r} > \delta^{DEC,DEG}_{\max,r}$, it implies there is not even one single δ value for which DEC$_\delta$ would exhibit the particular level of correlation (r) with both DEG and CLC. Hence, we do not add the precision level ε in the δ-*space*$_r$ formulation for $\delta^{DEC,CLC}_{\min,r} > \delta^{DEC,DEG}_{\max,r}$.

Among the δ values shown in Fig. 5, the largest δ value for which PCC(DEC$_\delta$, DEG) ≥ 0.80 is $\delta^{DEC,DEG}_{\max,r\geqslant 0.8} = 0.80$ and the smallest δ value for which PCC(DEC$_\delta$, CLC) ≥ 0.80 is $\delta^{DEC,CLC}_{\min,r\geqslant 0.8} = 0.03$. Hence, the δ-*space*$_r$ for a very strongly positive DEG-DEC-CLC correlation for the example graph is $\delta^{DEC,DEG}_{\max,r\geqslant 0.8} - \delta^{DEC,CLC}_{\min,r\geqslant 0.8} + \varepsilon$ = 0.80–0.03 + 0.01 = 0.78. Statistically, 0.78 is the probability for observing the decay centrality of the vertices (computed for a randomly chosen value of δ) to exhibit a very strongly positive correlation with both degree centrality and closeness centrality. Note that 0.78 is also PCC(DEG, CLC) for the example graph (computed in Fig. 4).

As the largest value for δ-*space*$_r$ is based on the closed interval [0+ε, 1−ε] and the δ values are from the open interval (0, 1), the δ-*space*$_r$ value for a real-world network would directly correspond to the probability with which the decay centrality of the vertices computed for a randomly chosen δ value from the range (0, 1) would exhibit the particular level of correlation (r) with both degree centrality and closeness centrality.

Hypothesis: As larger the δ-*space*$_r$ value for a real-world network, larger the probability for observing a correlation level r between DEC, DEG and DEC, CLC, we hypothesize that the Pearson's correlation coefficient between degree centrality and closeness centrality PCC(DEG, CLC) for a particular real-world network would be very strongly correlated with the δ-*space*$_r$ for that network (for both the strongly and very strongly positive levels of correlation) and claim that one could use the PCC(DEG, CLC) value of a real-world network to predict the δ-*space*$_r$ value for the real-world network.

5. Real-world networks

In this section, we introduce the 48 real-world networks (see Table 2) analyzed in this paper (the three character code acronym, name and the network type) and tabulate the values for the number of nodes and edges as well as the average node degree (k_{avg}) and spectral radius ratio for node degree (λ_{sp}). All the real-world networks are modeled as undirected graphs. The spectral radius ratio for node degree (Meghanathan, 2014) is a measure of the variation in node degree and is calculated as the ratio of the principal eigenvalue (Bonacich, 1987) of the adjacency matrix of the graph to that of the average node degree. The spectral radius ratio for node degree is independent of the number of vertices and the actual degree values for the vertices in the graph. The spectral radius ratio for node degree is always greater than or equal to 1; the farther is the ratio from the value of 1, the larger the variation in node degree. The spectral radius ratio for node degree for the real-world network graphs analyzed in this paper ranges from 1.01 to 5.51 (indicating that the real-world network graphs analyzed range from random networks (Renyi, 1959) with smaller variation in node degree to scale-free networks (Barabasi and Albert, 1999) of larger variation in node degree).

The networks considered cover a broad range of categories (as listed below along with the number of networks in each category): Acquaintance network (12), Friendship network (9), Co-appearance network (6), Employment network (4), Citation network (3), Collaboration network (3), Literature network (3), Political network (2), Biological network (2), Game network (2), Transportation network and Trade network (1 each). A brief description about each category of networks is as follows: An *acquaintance network* is a kind of social network in which the participant nodes slightly (not closely) know each other, as observed typically during an observation period. A *friendship network* is a kind of social network in which the participant nodes closely know each other and the relationship is not captured over an observation period. A *co-appearance network* is a network typically extracted from novels/books in such a way that two characters or words (modeled as nodes) are connected if they appear alongside each other. An *employment network* is a network in which the interaction/relationship between people is primarily due to their employment requirements and not due to any personal liking. A *citation network* is a network in which two papers (nodes) are connected if one paper cites the other paper as reference. A *collaboration network* is a network of researchers/authors who are listed as co-authors in at least one publication. A *biological network* is a network that models the interactions between genes, proteins, animals of a particular species, etc. A *political network* is a network of entities (typically politicians) involved in politics. A *game network* is a network of teams or players playing for different teams and their associations. A *literature network* is a network of books/papers/terminologies/authors (other than collaboration, citation or co-authorship) involved in a particular area of literature. A *transportation network* is a network of entities (like airports and

Table 2
Real-World Networks used in the Correlation Analysis.

#	Net.	Net. Description	Ref.	Network Type	λ_{sp}	#nodes	#edges	k_{avg}
1	ADJ	Word Adjacency Network	Newman, September 2006	Co-appearance Net.	1.73	112	425	7.589
2	AKN	Anna Karnenina Network	Knuth, 1993	Co-appearance Net.	2.48	140	494	7.057
3	JBN	Jazz Band Network	Geiser and Danon, 2003	Employment Net.	1.45	198	2742	27.697
4	CEN	C. Elegans Neural Network	White et al., 1986	Biological Net.	1.68	297	2148	14.465
5	CLN	Centrality Literature Net.	Hummon et al., 1990	Citation Net.	2.03	118	613	10.39
6	CGD	Citation Graph Drawing Net	Biedl and Franz, 2001	Citation Net.	2.24	259	640	4.942
7	CFN	Copperfield Network	Knuth, December 1993	Co-appearance Net.	1.83	89	407	9.146
8	DON	Dolphin Network	Lusseau et al., 2003	Acquaintance Net.	1.40	62	159	5.129
9	DRN	Drug Network	Lee, 2004	Acquaintance Net.	2.76	212	284	2.679
10	DLN	Dutch Literature 1976 Net.	de Nooy, 1999	Literature Net.	1.49	37	81	4.378
11	ERD	Erdos Collaboration Net.	Batagelj and Mrvar, 2006	Collaboration Net.	3.00	433	1314	6.069
12	FMH	Faux Mesa High School Net	Resnick et al., 1997	Friendship Net.	2.81	147	202	2.748
13	FHT	Friendship in Hi-Tech Firm	Krackhardt, 1999	Friendship Net.	1.57	33	91	5.515
14	FTC	Flying Teams Cade Net.	Moreno, 1960	Employment Net.	1.21	48	170	7.083
15	FON	US Football Network	Girvan and Newman, 2002	Game Net.	1.01	115	613	10.661
16	CDF	College Dorm Fraternity Net	Bernard et al., 1980	Acquaintance Net.	1.11	58	967	33.345
17	GD96	Graph Drawing 1996 Net	Batagelj and Mrvar, 2006	Citation Net.	2.38	180	228	2.533
18	MUN	Marvel Universe Network	Gleiser, 2007	Co-appearance Net.	2.54	167	301	3.605
19	GLN	Graph Glossary Network	Batagelj and Mrvar, 2006	Literature Net.	2.01	67	118	3.522
20	HTN	Hypertext 2009 Network	Isella et al., 2011	Acquaintance Net.	1.21	115	2164	37.635
21	HCN	Huckleberry Coappear. Net.	Knuth, 1993	Co-appearance Net.	1.66	76	302	7.947
22	ISP	Infectious Socio-Patterns Net	Isella et al., 2011	Acquaintance Net.	1.69	309	1924	12.453
23	KCN	Karate Club Network	Zachary, 1977	Acquaintance Net.	1.47	34	78	4.588
24	KFP	Korea Family Planning Net.	Rogers and Kincaid, 1980	Acquaintance Net.	1.70	37	85	4.595
25	LMN	Les Miserables Network	Knuth, 1993	Co-appearance Net.	1.82	77	254	6.597
26	MDN	Macaque Dominance Net.	Takahata, 1991	Biological Net.	1.04	62	1167	37.645
27	MTB	Madrid Train Bombing Net.	Hayes, 2006	Acquaintance Net.	1.95	64	295	9.219
28	MCE	Manufact. Comp. Empl. Net.	Cross et al., 2004	Employment Net.	1.12	77	1549	40.23
29	MSJ	Soc. Net. Journal Co-authors	McCarty and Freeman, 2008	Co-author Net.	3.48	475	625	2.632
30	AFB	Author Facebook Network	-	Friendship Net.	2.29	171	940	10.994
31	MPN	Mexican Political Elite Net.	Gil-Mendieta and Schmidt, 1996	Political Net.	1.23	35	117	6.686
32	MMN	ModMath Network	Batagelj and Mrvar, 2006	Friendship Net.	1.59	30	61	4.067
33	NSC	Net. Science Co-author Net.	Newman, September 2006	Co-author Net.	5.51	1,589	2,743	3.45
34	PBN	US Politics Books Network	Krebs, 2003	Literature Net.	1.42	105	441	8.4
35	PSN	Primary School Contact Net.	Gemmetto et al., 2014	Acquaintance Net.	1.22	238	5539	46.546
36	PFN	Prison Friendship Network	MacRae, 1960	Friendship Net.	1.32	67	142	4.239
37	SJN	San Juan Sur Family Net.	Loomis et al., 1953	Acquaintance Net.	1.29	75	155	4.133
38	SDI	Scotland Corp. Interlock Net	Scott, 1980	Employment Net.	1.94	230	359	3.122
39	SPR	Senator Press Release Net.	Grimmer, 2010	Political Net.	1.47	92	477	10.37
40	SWC	Soccer World Cup 1998 Net	Batagelj and Mrvar, 2006	Game Net.	1.45	35	118	6.743
41	SSM	Sawmill Strike Comm. Net.	Michael, 1997	Acquaintance Net.	1.22	24	38	3.167
42	TEN	Taro Exchange Network	Schwimmer, 1973	Acquaintance Net.	1.06	22	39	3.545
43	TWF	Teenage Female Friend Net.	Pearson and Michell, 2000	Friendship Net.	1.59	47	77	3.277
44	UKF	UK Faculty Friendship Net.	Nepusz et al., 2008	Friendship Net.	1.35	83	578	13.928
45	APN	US Airports 1997 Network	Batagelj and Mrvar, 2006	Transportation Net.	3.22	332	2126	12.807
46	RHF	Residence Hall Friend Net.	Freeman et al., 1998	Friendship Net.	1.27	217	1839	16.949
47	WSB	Windsurfers Beach Network	Freeman et al., 1989	Friendship Net.	1.22	43	336	15.628
48	WTN	World Trade Metal Network	Smith and White, 1992	Trade Net.	1.38	80	875	21.875

their flight connections) involved in public transportation. A *trade network* is a network of countries/people involved in certain trade.

6. Results of correlation analysis

We now present in detail the results of correlation analysis conducted between decay centrality (DEC$_\delta$ computed for $\delta = 0.01, \ldots, 0.99$, in increments of 0.01) and the degree centrality and closeness centrality metrics for each of the 48 real-world networks. In pursuit of validating our hypothesis, we first determine the two values $\delta^{DEC,DEG}_{\max,r}$ and $\delta^{DEC,CLC}_{\min,r}$ for each real-world network (see Table 3) with respect to the two levels of correlation (strongly positive: $r \geq 0.6$ and very strongly positive: $r \geq 0.8$). Table 3 also lists the δ-$space_r$ values (calculated as per the discussion in Section 4) for each real-world network with respect to the two levels of correlation.

Fig. 6 plots the sorted values of δ-$space_r$ for the two levels of correlation. The median value of δ-$space_r$ is 0.52 for $r \equiv vs+$(very strongly positive correlation) and 0.99 for $r \equiv s+$(strongly positive correlation). Fig. 7 plots the distribution of the δ-$space_{r\equiv vs+}$ values and the δ-$space_{r\equiv s+}$ values: we observe real-world networks that had a δ-$space_{r\equiv s+}$ of 0.99 end up having δ-$space_{r\equiv vs+}$ ranging from 0 to 0.99. With respect to the strongly positive level of correlation, about 30 of the 48 real-world networks (close to about 2/3rds of the networks) had a δ-$space_{r\equiv s+}$ of 0.99 and only one of the 48 real-world networks had a negative value for δ-$space_{r\equiv s+}$. On the other hand, with respect to the very strongly positive level of correlation, only 14 of the 48 real-world networks had a δ-$space_{r\equiv vs+}$ of 0.99, and 12 of the 48 real-world networks (25% of the networks) had a negative value for δ-$space_{r\equiv vs+}$.

Fig. 8 plots the δ-$space_{r\equiv s+}$ values of the real-world networks and the difference between the δ-$space_{r\equiv s+}$ and δ-$space_{r\equiv vs+}$ values (δ-$space_{r\equiv s+}$ − δ-$space_{r\equiv vs+}$), in the decreasing order of the difference (i.e., networks for which the difference in the δ-$space_r$ values is larger appears on the left and networks for which the difference in the δ-$space_r$ values is closer to 0 or equal to 0 appear on the right). Though real-world networks with δ-$space_{r\equiv s+}$ values of 0.99 also suffered a larger drop in the case of the δ-$space_{r\equiv vs+}$ values, we observe a majority of the real-world networks that suffered a larger drop in the δ-$space_{r\equiv vs+}$ values are those that already had a

Table 3
δ-*space*$_r$ for Strongly Positive and Very Strongly Positive DEG-DEC-CLC Correlation.

#	Net.	$\delta_{\max,r}^{DEC-DEG}$ for r values below		$\delta_{\min,r}^{DEC-CLC}$ for r values below		δ-*space*$_{r\equiv s+}$: Strongly Positive Correlation	δ-*space*$_{r\equiv vs+}$: Very Strongly Positive Correlation
		≥ 0.6	≥ 0.8	≥ 0.6	≥ 0.8	Median: 0.99	Median: 0.52
1	ADJ	0.99	0.65	0.01	0.01	0.99	0.65
2	AKN	0.99	0.59	0.01	0.01	0.99	0.59
3	JBN	0.99	0.75	0.01	0.01	0.99	0.75
4	CEN	0.99	0.32	0.01	0.05	0.99	0.28
5	CLN	0.67	0.42	0.31	0.54	0.37	-0.12
6	CGD	0.88	0.59	0.24	0.56	0.65	0.04
7	CFN	0.99	0.89	0.01	0.01	0.99	0.89
8	DON	0.99	0.63	0.01	0.11	0.99	0.53
9	DRN	0.99	0.61	0.01	0.61	0.99	0.01
10	DLN	0.99	0.99	0.01	0.01	0.99	0.99
11	ERD	0.74	0.47	0.52	0.72	0.23	-0.25
12	FMH	0.99	0.66	0.01	0.47	0.99	0.20
13	FHT	0.83	0.60	0.29	0.57	0.55	0.04
14	FTC	0.99	0.96	0.01	0.01	0.99	0.96
15	FON	0.22	0.12	0.06	0.12	0.17	0.01
16	CDF	0.99	0.99	0.01	0.01	0.99	0.99
17	GD96	0.53	0.27	0.05	0.19	0.49	0.09
18	MUN	0.54	0.30	0.25	0.50	0.30	-0.20
19	GLN	0.69	0.40	0.25	0.50	0.45	-0.10
20	HTN	0.99	0.99	0.01	0.01	0.99	0.99
21	HCN	0.60	0.37	0.31	0.53	0.30	-0.16
22	ISP	0.99	0.53	0.01	0.08	0.99	0.46
23	KCN	0.99	0.66	0.01	0.04	0.99	0.63
24	KFP	0.91	0.68	0.34	0.70	0.58	-0.02
25	LMN	0.99	0.62	0.01	0.01	0.99	0.62
26	MDN	0.99	0.99	0.01	0.01	0.99	0.99
27	MTB	0.79	0.57	0.38	0.62	0.42	-0.05
28	MCE	0.99	0.99	0.01	0.01	0.99	0.99
29	MSJ	0.52	0.34	0.37	0.59	0.16	-0.25
30	AFB	0.43	0.29	0.47	0.62	-0.03	-0.33
31	MPN	0.99	0.99	0.01	0.01	0.99	0.99
32	MMN	0.99	0.83	0.01	0.18	0.99	0.66
33	NSC	0.42	0.26	0.24	0.43	0.19	-0.17
34	PBN	0.90	0.53	0.03	0.31	0.88	0.23
35	PSN	0.99	0.99	0.01	0.01	0.99	0.99
36	PFN	0.99	0.99	0.01	0.01	0.99	0.99
37	SJN	0.99	0.44	0.01	0.11	0.99	0.34
38	SDI	0.57	0.31	0.42	0.71	0.16	-0.40
39	SPR	0.99	0.99	0.01	0.01	0.99	0.99
40	SWC	0.99	0.99	0.01	0.01	0.99	0.99
41	SSM	0.99	0.65	0.01	0.02	0.99	0.64
42	TEN	0.93	0.46	0.01	0.18	0.93	0.29
43	TWF	0.85	0.64	0.50	0.73	0.36	-0.09
44	UKF	0.99	0.99	0.01	0.01	0.99	0.99
45	APN	0.99	0.50	0.01	0.01	0.99	0.50
46	RHF	0.99	0.99	0.01	0.01	0.99	0.99
47	WSB	0.99	0.99	0.01	0.01	0.99	0.99
48	WTN	0.99	0.99	0.01	0.01	0.99	0.99

The color shade is to illustrate the networks for which the delta-space is negative.

lower δ-*space*$_{r\equiv s+}$ value. It is evident from Table 3 that 11 of the 12 real-world networks for which δ-*space*$_{r\equiv vs+}$ is negative are those networks for which δ-*space*$_{r\equiv s+}$ is less than 0.5.

6.1. Statistical property: Actual δ-space vs. probability of observing a particular level of DEG-DEC-CLC correlation

In this sub section, we show that the δ-*space*$_r$ for a real-world network with respect to a particular level of correlation (r) corresponds to the probability that the real-world network would exhibit the particular level of correlation between DEC$_\delta$ and DEG as well as between DEC$_\delta$ and CLC for the DEC values determined using a randomly chosen value of δ in the range (0, 1). For each real-world network, we computed the DEC$_\delta$ values using hundred randomly chosen values of δ in the range (0, 1). For each of these hundred sets of DEC$_\delta$ values, we computed the Pearson's correlation coefficient between the DEC$_\delta$ values and the DEG values of the vertices as well as between the DEC$_\delta$ values and the CLC values of the vertices. We determined the fraction of these hundred sets of DEC$_\delta$ values that exhibited strongly positive and very strongly positive correlations with each of DEG and CLC metrics. Fig. 9 presents a plot of the actual δ-*space*$_r$ values for the real-world networks vs. the probability for observing the particular level of DEG-DEC-CLC correlation using a randomly chosen value of the decay parameter

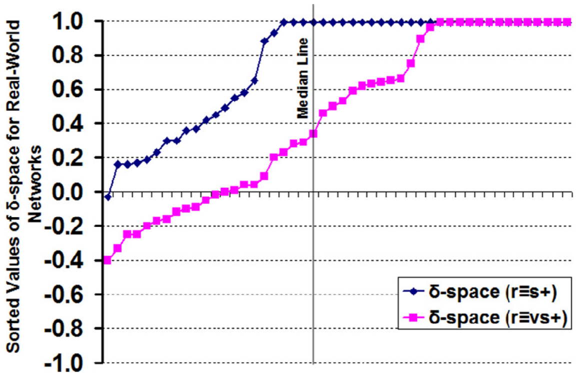


Fig. 6. Plot of the Sorted Values of δ-$space_r$ for Strongly Positive and Very Strongly Positive DEG-DEC-CLC Correlation.

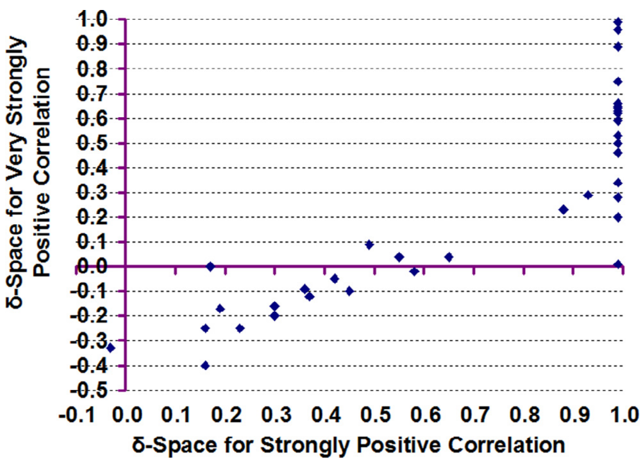


Fig. 7. Distribution of the δ-$space_r$ Values for Strongly Positive DEG-DEC-CLC Correlation vs. and δ-$space_r$ Values for Very Strongly Positive DEG-DEC-CLC Correlation.

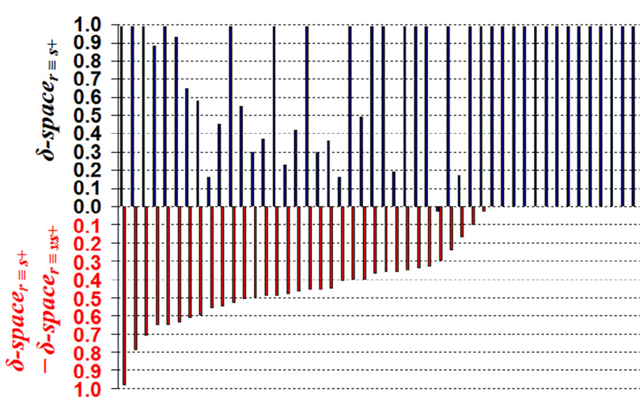


Fig. 8. δ-$space_{r\equiv s+}$ Values and Difference between the δ-$space_{r\equiv s+}$ Values and δ-$space_{r\equiv vs+}$ Values (in the decreasing order of the difference).

δ. We observe a very close coincidence between the two values for all the real-world networks as all the data points lie close to the diagonal line for both the strongly positive and very strongly positive levels of correlation. The small difference (between the actual δ-$space_r$ and the probability of observing a particular level of correlation, r) could be attributed to the randomness associated with the hundred randomly chosen δ values. For real-world networks with a negative δ-$space_r$, we notice the probability for observing the particular level of correlation (r) is zero. The median of the probability for observing strongly positive level of DEG-DEC-CLC correlation is 1.0 (closer to the actual median value of 0.99) and the probability for observing very strongly positive level of correlation is 0.53 (closer to the actual median value of 0.52).

6.2. Validation of hypothesis: Actual DEG, CLC Pearson's correlation coefficient vs. predicted δ-$space_r$ for DEG-DEC-CLC correlation

Since larger δ-$space_r$ corresponds to a larger probability for observing the particular level of correlation (r) between DEC, DEG and DEC, CLC metrics, we hypothesize that larger δ-$space_r$ would also correspond to a larger value for the Pearson's correlation coefficient between DEG and CLC metrics (a kind of transitive relationship) and vice versa. Hence, we claim that the δ-$space_r$ value for a real-world network could be predicted using the Pearson's correlation coefficient between DEG and CLC. Figs. 10–11 corroborate our hypothesis. We build individual linear models between the actual PCC(DEG, CLC) values (as the predictor variable) and the actual δ-$space_{r\equiv s+}$ values as well as the actual δ-$space_{r\equiv vs+}$ values (both as the predicted variables). The linear models are shown in Eqs. (6) and (7). Notice the R^2 values for the linear correlations (6) and (7) are significantly high, validating our hypothesis. Fig. 11 presents the distribution of the actual δ-$space_{r\equiv s+}$ and δ-$space_{r\equiv vs+}$ values vs. the predicted δ-$space_{r\equiv s+}$ and δ-$space_{r\equiv vs+}$ values using the actual PCC(DEG, CLC) values for the real-world networks substituted in the linear correlations (6) and (7).

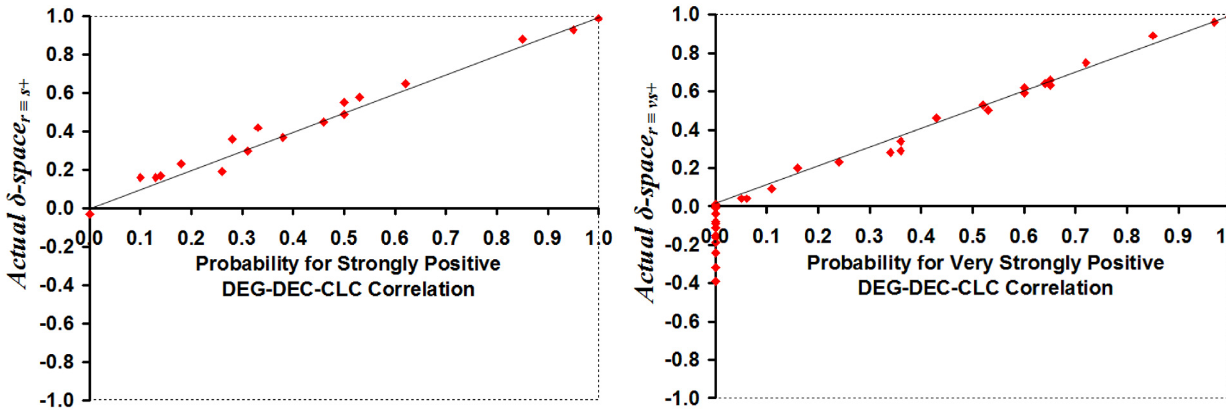
$$\delta - space_{r\equiv s+} = 1.0617^{*}\mathrm{PCC(DEG, CLC)} + 0.0693; \quad R^2 = 0.8499 \tag{6}$$

$$\delta - space_{r\equiv vs+} = 1.5780^{*}\mathrm{PCC(DEG, CLC)} - 0.5959; \quad R^2 = 0.8916 \tag{7}$$

In Fig. 11, we observe a majority of the predicted δ-$space_{r\equiv vs+}$ values to be slightly larger than the actual δ-$space_{r\equiv vs+}$ values; this could be also inferred from the median of the predicted δ-$space_{r\equiv vs+}$ values (0.55) being slightly larger than the median of the actual δ-$space_{r\equiv +}$ values (0.52). In the case of strongly positive correlation, we observe that for 14 of the 30 real-world networks (for which the actual δ-$space_{r\equiv s+}$ values are 0.99), the predicted δ-$space_{r\equiv s+}$ values are greater than 0.99 (the values are as large as 1.12); however, the median of the predicted δ-$space_{r\equiv s+}$ values is still 0.99 (same as the median of the actual δ-$space_{r\equiv s+}$ values).

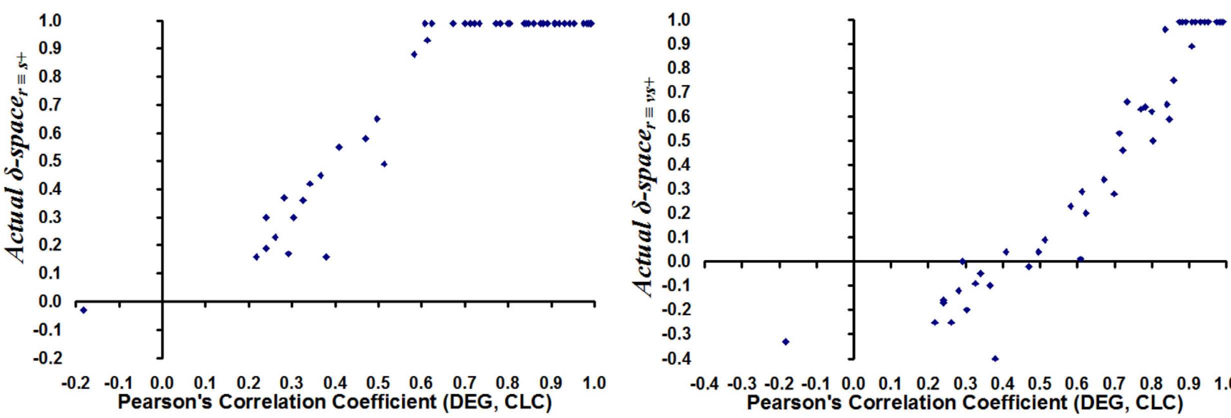
6.3. Degree centrality: Bottleneck centrality metric for lower positive δ-$space_{r\equiv vs+}$ values

Fig. 12 presents the real-world networks in the decreasing order of the δ-$space_{r\equiv vs+}$ values, wherein 0 < δ-$space_{r\equiv vs+}$ < 0.99 (there are 22 such real-world networks out of the total of 48). With this figure, we are able to analyze which of the two centrality metrics (DEG or CLC) is a bottleneck metric that prevents the δ-$space_{r\equiv vs+}$ values from being the maximum value of 0.99. We observe $\delta^{DEC,CLC}_{\min,r\geqslant 0.8}$ to be less than 0.20 for 17 of the 22 real-world networks (implying the closeness centrality metric provided the opportunity for δ-$space_{r\equiv vs+}$ to be 0.80 or above for 17 of the 22 real-world networks for which δ-$space_{r\equiv vs+}$ <0.99) whereas $\delta^{DEC,DEG}_{\max,r\geqslant 0.8}$ is greater than 0.80 for only 3 of the 22 real-world networks that are of interest in this case (implying the degree centrality metric provided the opportunity for δ-$space_{r\equiv vs+}$ to be 0.80 or above for only 3 of the 22 real-world networks for which δ- $space_{r\equiv vs+}$ <0.99). Hence, it is obvious that the relatively lower δ- $space_{r\equiv vs+}$ values (vis-a-vis the δ-$space_{r\equiv s+}$ values) is due to the reduced range of δ values for which the decay centrality metric exhibits a very strong positive correlation with the degree centrality metric.



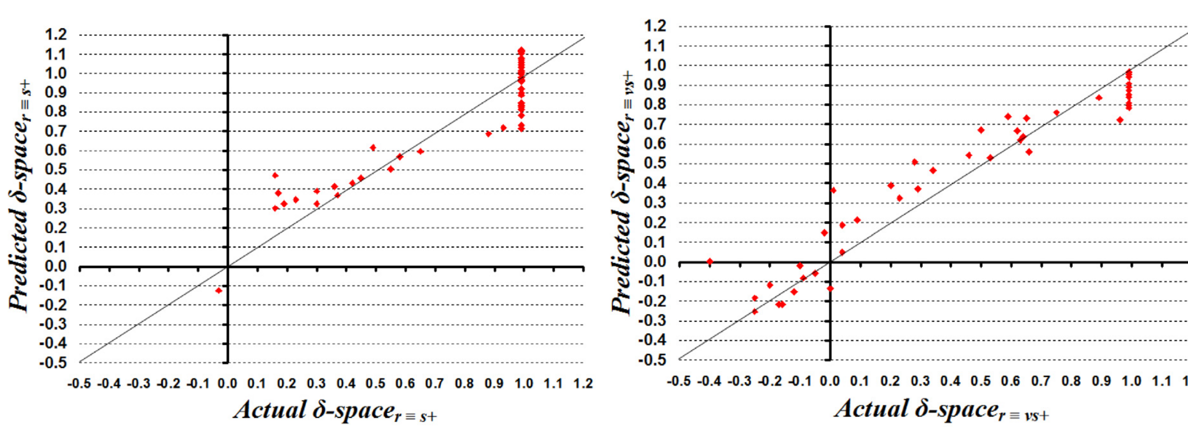
(a) Strongly Positive Correlation (b) Very Strongly Positive Correlation

Fig. 9. Probability for Observing Strongly Positive or Very Strongly Positive DEG-DEC-CLC Correlation vs. the Actual $\delta\text{-}space_{r\equiv s+}$ Values and $\delta\text{-}space_{r\equiv vs+}$ Values.



(a) Strongly Positive Correlation (b) Very Strongly Positive Correlation

Fig. 10. Pearson's Correlation Coefficient between DEG and CLC vs. the Actual $\delta\text{-}space_{r\equiv s+}$ Values and the Actual $\delta\text{-}space_{r\equiv vs+}$ Values.



(a) Strongly Positive Correlation (b) Very Strongly Positive Correlation

Fig. 11. Actual {$\delta\text{-}space_{r\equiv s+}$, $\delta\text{-}space_{r\equiv vs+}$} Values vs. Predicted {$\delta\text{-}space_{r\equiv +}$, $\delta\text{-}space_{r\equiv vs+}$} Values.

6.4. Impact of topological measures on $\delta\text{-}space_{r\equiv vs+}$ values

We identified a total of six topological measures that could be potential drivers for the $\delta\text{-}space_{r\equiv vs+}$ values of the real-world networks. The six topological measures considered are: Spectral radius ratio for node degree ($\lambda_{sp}(k)$), Average component size, Algebraic connectivity (ALGC), Graph modularity (G_m), Graph density (G_d) and Degree-based Assortativity index ($A.\ Index_k$). We identified a total of 26 real-world networks (among the 48 real-world networks) that either had $\delta\text{-}space_{r\equiv vs+}$ values of 0.99 (there are 14 such

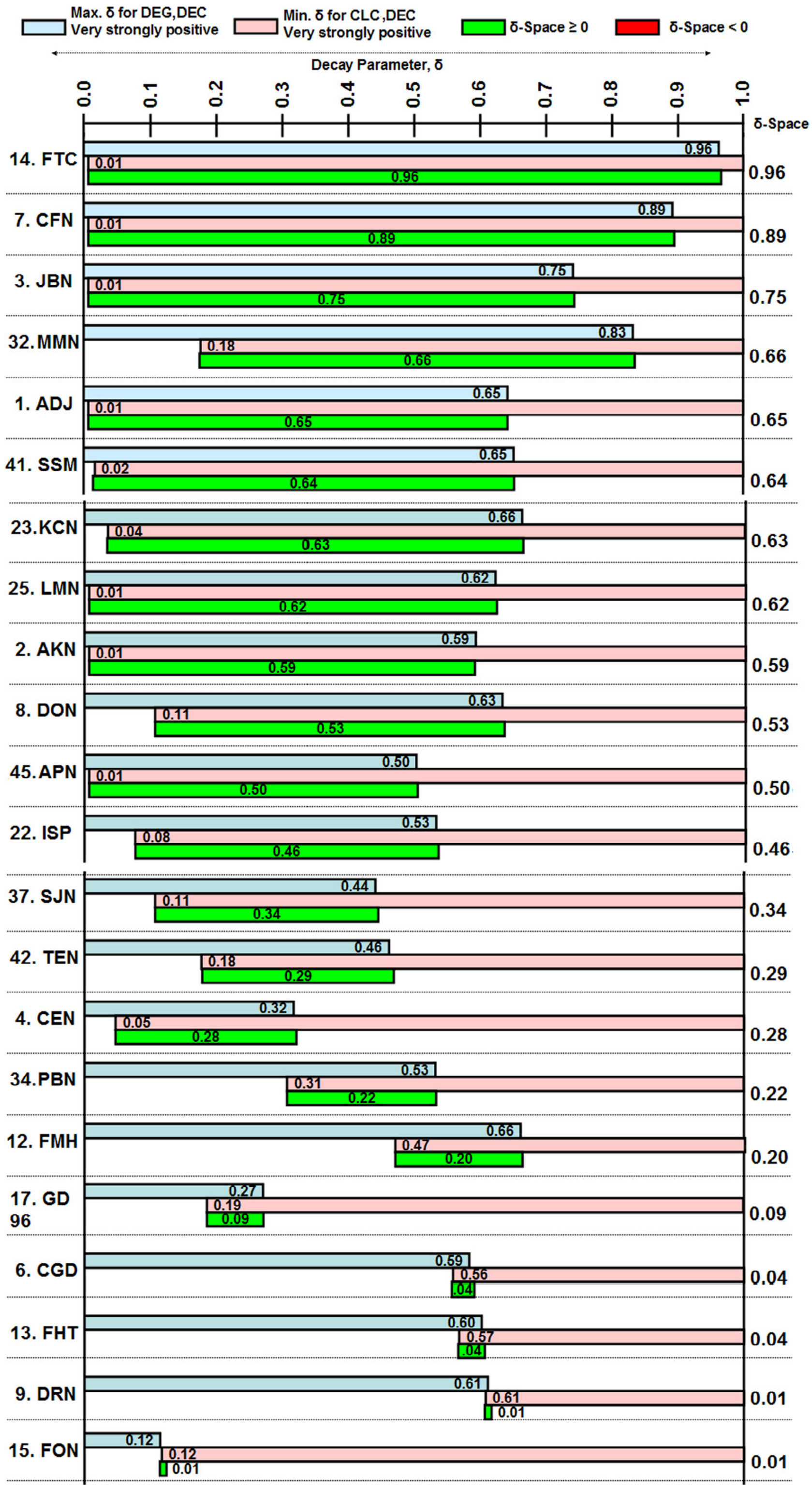


Fig. 12. Real-World Networks in the Decreasing Order of $\delta\text{-}space_{r\equiv vs+}$ Values ($0 < \delta\text{-}space_{r\equiv vs+} < 0.99$).

networks) or negative $\delta\text{-}space_{r\equiv vs+}$ values (there are 12 such networks). We plotted (in Fig. 13) the values incurred for each of the above six topological measures vs. the actual $\delta\text{-}space_{r\equiv vs+}$ values and use these plots to analyze the impact of the topological measures on the $\delta\text{-}space_{r\equiv vs+}$ values.

From Fig. 13-a, we observe that all the 14 real-world networks with $\delta\text{-}space_{r\equiv vs+}$ values of 0.99 have $\lambda_{sp}(k)$ values less than 1.5. On the other hand, all the 12 real-world networks with negative $\delta\text{-}space_{r\equiv vs+}$ values had $\lambda_{sp}(k)$ values above 1.5, and 7 of these 12 networks had $\lambda_{sp}(k)$ values above 2.0. Note that λ_{sp} is a measure of variation in node degree and is greater than or equal to 1.0. The closer the λ_{sp} value to 1.0: the degree centrality of the nodes are comparable to each other (similar to the case of random networks). On the other hand, farther the λ_{sp} value from 1.0, there could exist

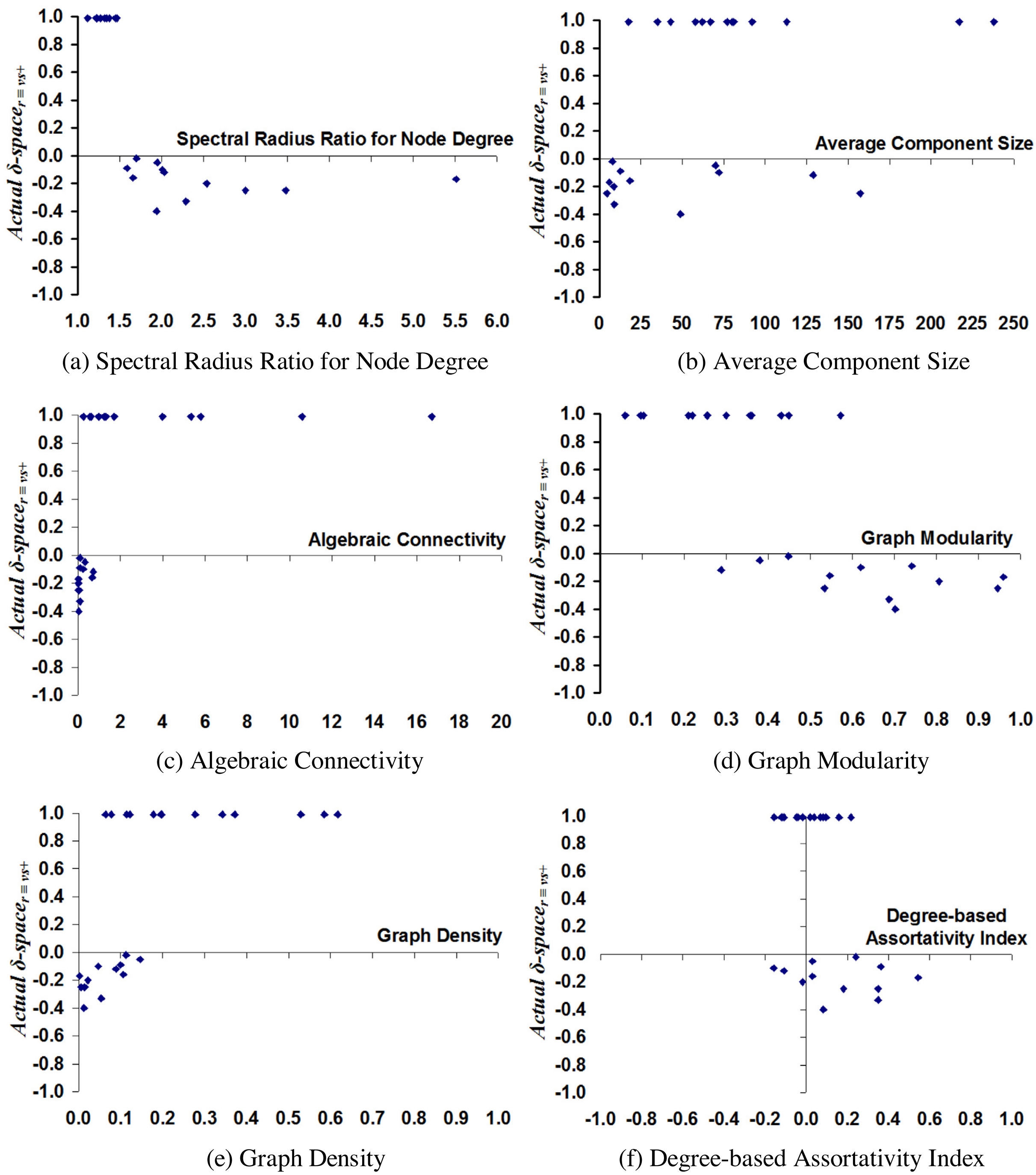


Fig. 13. Impact of Various Topological Measures for Real-World Networks with Negative δ-$space_{r\equiv vs+}$ Values and δ-$space_{r\equiv vs+}$ Values of 0.99.

one or more nodes (like hubs) whose degree centrality might be appreciably larger than the rest of the nodes (characteristic of scale-free networks with preferential attachment). From the above observations, we could conclude that real-world networks with larger variation in node degree (i.e., real-world networks that exhibit scale-free properties) are more likely to exhibit a negative δ-$space_{r\equiv vs+}$ value (i.e., there would not be even a single value for the decay parameter δ for which the DEC_δ metric would exhibit a very strongly positive correlation with the DEG and CLC metrics).

We measure the average component size for a real-world network as the total number of nodes in the network divided by the number of components (a component is the largest connected set of vertices) in the network. From Fig. 13-b, we observe real-world networks with negative δ-$space_{r\equiv vs+}$ values to more likely have a lower average component size (median: 15.5) compared to real-world networks with δ-$space_{r\equiv vs+}$ values of 0.99 (median: 72). Thus, for networks with smaller component size, the correlation between DEG and DEC_δ and the correlation between CLC and DEC_δ appear to be relatively more sensitive to the δ values. In other words, for networks with smaller component size, it is difficult to find a range of δ values for which both the DEG- DEC_δ and CLC-DEC_δ correlations would be very strong.

The algebraic connectivity (ALGC) Fiedler, 1973 of a network is the second smallest eigenvalue of the Laplacian matrix (Mohar, 1992) of the network graph. For any two different vertices i and j, an entry (i, j) in the Laplacian matrix is 0 if there is no edge between i and j and -1 if there is an edge between i and j. For a vertex i, an entry (i, i) in the Laplacian matrix is -k_i, where k_i is the degree of vertex i. The algebraic connectivity of a network is a quantitative measure of the strength of connectivity of the network with respect to link removals (Meghanathan, 2016). Fig. 13-c indicates that real-world networks with negative δ-$space_{r\equiv vs+}$ values have significantly smaller ALGC values (median: 0.10) compared to the ALGC values (median: 1.38) of networks with δ-$space_{r\equiv vs+}$ values of 0.99. None of the 12 real-world networks with negative δ-$space_{r\equiv vs+}$ values have ALGC above 0.75; whereas, 10 of the 14 real-world networks with δ-$space_{r\equiv vs+}$ values of 0.99 have ALGC above 0.75.

The modularity of a network is a quantitative measure of the strength of the division of the network into communities or modules (Newman, 2006). The modularity score for a network ranges from [−0.5, ..., 1); networks with larger modularity have dense connections among the nodes within a community and sparse connections between the nodes across communities. We use the Gephi software (Bastian et al., 2009) to compute the modularity scores of the 48 real-world networks studied in this research. The Gephi software uses the well-known randomization method (Newman, 2006) to compute the modularity score for a network. From Fig. 13-d, we observe that 9 of the 12 real-world networks with negative δ-$space_{r\equiv vs+}$ values had modularity above 0.5 (with the median being 0.65), whereas 13 of the 14 real-world networks with δ-$space_{r\equiv vs+}$ values of 0.99 had modularity less than 0.5 (with the median being 0.25). The negative δ-$space_{r\equiv vs+}$ values for real-world networks with larger modularity scores vindicate the scarcity of edges connecting two different communities in such networks and highlights the difficulty in observing stronger correlation between the three centrality metrics (DEG, DEC and CLC).

Fig. 13-e clearly indicates that the density of the real-world networks (median: 0.05) with negative δ-$space_{r\equiv vs+}$ values are likely to be lower than the density of the networks (median: 0.19) with δ-$space_{r\equiv vs+}$ values of 0.99. This observation coincides with the earlier observations made for graphs with lower algebraic connectivity as well as for those with higher modularity (note that networks with high modularity need not be dense overall due to the scarcity of links connecting any two communities). We measure the density of a graph as the ratio of the actual number of edges to that of the maximum possible number of edges (considering that there could be an edge between any two vertices) in the graph.

The assortativity index (A. Index) of a network (Newman, 2002) is a quantitative measure of the extent of similarity between the end vertices of the edges with respect to a node-level metric (typically measured on the basis of the degree of the vertices). The degree-based A. Index ($A.\ Index_k$) Newman, 2002 of a network is computed as the Pearson's correlation coefficient (ranges from −1 to 1) of the degree values of the end vertices of the edges in the network. Networks with A. Index values closer to 1 (or −1) are considered to be assortative (or dissortative). Networks with A. Index values closer to 0 are considered to be neutral with respect to the node-level metric used to compute the correlation coefficient. We observe real-world networks with negative δ-$space_{r\equiv vs+}$ values to more likely have a larger $A.\ Index_k$ value (median: 0.13) compared to real-world networks with δ-$space_{r\equiv vs+}$ values of 0.99 (median: 0.00). In other words, we opine that assortative real-world networks are more likely to have lower δ-$space_{r\equiv vs+}$ values compared to networks that are either neutral or dissortative.

7. Conclusions and future work

Our paper is innovative on the following lines: We analyze real-world networks rather than the simulated random networks. We use the Pearson's correlation measure to study the correlation between the actual centrality values rather than multinomial logistic regression (Greene, 2011) to study the sets of vertices that had the largest values of centrality. For all the 48 real-world networks (of spectral radius ratio for node degree ranging from 1.01 to 5.51) analyzed, we observe the Pearson's correlation coefficient (PCC) between the decay centrality (DEC) and the degree centrality (DEG) to decrease with increase in the decay parameter δ and the Pearson's correlation coefficient between DEC and closeness centrality (CLC) to decrease with decrease in δ. Such a trend of variation of the Pearson's correlation coefficient involving DEC, DEG and CLC on the basis of the decay parameter has not been reported in the literature until now.

Unlike the observation for random networks in Tsakas (2016), for each of the 48 real-world networks studied in this paper: we observe two different δ values below which or above which the decay centrality metric exhibits a particular level of correlation (r) with the DEG or CLC metrics respectively. We make use of this observation and propose the notion of δ-$space_r$ ($r \geq 0.6$ for strongly positive correlation and $r \geq 0.8$ for very strongly positive correlation) as the difference between the maximum δ value for which DEC_δ exhibits the particular level of correlation (r) with DEG and the minimum δ value for which DEC_δ exhibits the same level of correlation with CLC. We hypothesize the existence of a transitive relationship (i.e., if PCC(DEG, DEC_δ) $\geq r$ and PCC(DEC_δ, CLC) $\geq r$, then PCC(DEG, CLC) is more likely to be $\geq r$) between DEG and CLC and show that we could use the PCC(DEG, CLC) value observed for a real-world network to predict the δ-$space_r$ value for the network. We also analyze the impact of various topological measures on the δ-$space_{r\equiv vs+}$ values of the real-world networks. We observe real-world networks are more likely to incur negative δ-$space_{r\equiv vs+}$ values (rather than incur δ-$space_{r\equiv vs+}$ values of 0.99) under one or more of the following topological conditions: larger variation in node degree, more assortativeness with respect to node degree, lower connectivity, higher graph modularity, lower graph density, or lower component size.

As part of future work, we will investigate the relationship between δ-$space_r$ and the Pearson's correlation coefficient observed between decay centrality and the computationally-heavy metrics such as eigenvector centrality and betweenness centrality. We also plan to investigate whether the monotonically decreasing/increasing trend (observed in real-world networks) for the Pearson's correlation coefficient involving decay centrality and the degree centrality/closeness centrality metrics is also observed for networks simulated from theoretical models.

References

Jackson, M.O., 2010. Social and Economic Networks. Princeton University Press.
Freeman, L.C., 1979. Centrality in social networks: conceptual clarification. Social Networks 1 (3), 215–239.
Evans, J.D., 1995. Straightforward Statistics for the Behavioral Sciences,. Brooks Cole Publishing Company.
Lay, D.C., Lay, S.R., McDonald, J.J., 2015. Linear Algebra and its Applications,. Pearson.
Meghanathan, N., Spectral Radius as a Measure of Variation in Node Degree for Complex Network Graphs. In: Proceedings of the 3rd International Conference on Digital Contents and Applications, pp. 30–33, Hainan, China, December 20–23, 2014.
Cormen, T.H., Leiserson, C.E., Rivest, R.L., Stein, C., 2009. Introduction to Algorithms. MIT Press, Cambridge, USA.
Tsakas, N., "On Decay Centrality," arXiv: 1604.05582, April 2016.
Bonacich, P., 1987. Power and centrality: a family of measures. Am. J. Sociol. 92 (5), 1170–1182.
Renyi, E., 1959. On Random Graphs I. Publicationes Mathematicae 6, 290–297.
Barabasi, A.-L., Albert, R., 1999. Emergence of scaling in random networks. Science 286 (5439), 509–512.
Newman, M.E.J., 2006. Finding Community Structure in Networks using the Eigenvectors of Matrices. Phys. Rev. E 74 (3), 036104.
Knuth, D.E., 1993. The Stanford GraphBase: A Platform for Combinatorial Computing. Addison-Wesley, Reading, MA.

Geiser, P., Danon, L., 2003. Community structure in Jazz. Adv. Complex Systems 6 (4), 563–573.
White, J.G., Southgate, E., Thomson, J.N., Brenner, S., 1986. The structure of the nervous system of the nematode *Caenorhabditis Elegans*. Philos. Trans. B 314 (1165), 1–340.
Hummon, N.P., Doreian, P., Freeman, L.C., 1990. Analyzing the structure of the centrality-productivity literature created between 1948 and 1979. Sci. Commun. 11 (4), 459–480. http://dx.doi.org/10.1177/107554709001100405.
Biedl, T., Franz, B.J., Graph-Drawing Contest Report. In: Proceedings of the 9th International Symposium on Graph Drawing, pp. 513–521, September 2001.
Lusseau, D., Schneider, K., Boisseau, O.J., Haase, P., Slooten, E., Dawson, S.M., 2003. The Bottlenose Dolphin Community of Doubtful Sound Features a Large Proportion of Long-lasting Associations. Behav. Ecol. Sociobiol. 54 (3), 396–405.
Lee, J.-S., 2004. Generating networks of illegal drug users using large samples of partial ego-network data. Intelligence and Security Informatics, Lecture Nodes in Computer Science 3073, 390–402.
de Nooy, W., 1999. A literary playground: literary criticism and balance theory. Poetics 26 (5–6), 385–404.
Batagelj, V., Mrvar, A., Pajek Datasets, <http://vlado.fmf.uni-lj.si/pub/networks/data/>, 2006.
Resnick, M.D., Bearman, P.S., Blum, R.W., Bauman, K.E., Harris, K.M., Jones, J., Tabor, J., Beuhring, T., Sieving, R.E., Shew, M., Ireland, M., Bearinger, L.H., Udry, J.R., 1997. Protecting adolescents from harm. Findings from the National Longitudinal Study on Adolescent Health. J. Am. Med. Assoc. 278 (10), 823–832.
Krackhardt, D., 1999. The ties that torture: Simmelian Tie analysis in organizations. Research Sociol. Organiz. 16, 183–210.
Moreno, J.L., 1960. The Sociometry Reader. The Free Press, Glencoe, IL, USA, pp. 534–547.
Girvan, M., Newman, M.E.J., 2002. Community structure in social and biological networks. Proc. Natl. Acad. Sci. U.S.A. 99 (12), 7821–7826.
Bernard, H.R., Killworth, P.D., Sailer, L., 1980. Informant accuracy in social network data IV: a comparison of clique-level structure in behavioral and cognitive network data. Social Networks 2 (3), 191–218.
Gleiser, P.M., How to become a Superhero, J. Statistical Mechanics: Theory and Experiments, P09020, September 2007.
Isella, L., Stehle, J., Barrat, A., Cattuto, C., Pinton, J.F., Van den Broeck, W., 2011. What's in a Crowd? Analysis of face-to-face behavioral networks. J. Theor. Biol. 271 (1), 166–180. http://dx.doi.org/10.1016/j.jtbi.2010.11.033.
Zachary, W.W., 1977. An information flow model for conflict and fission in small groups. J. Anthropol. Res. 33 (4), 452–473.
Rogers, E.M., Kincaid, D.L., 1980. Communication Networks: Toward a New Paradigm for Research. Free Press.
Takahata, Y., 1991. In: Diachronic Changes in the Dominance Relations of Adult Female Japanese Monkeys of the Arashiyama B Group: The Monkeys of Arashiyama. State University of New York Press, Albany, pp. 124–139.
Hayes, B., 2006. Connecting the Dots. Am. Sci. 94 (5), 400–404.
Cross, R.L., Parker, A., Cross, R., 2004. The Hidden Power of Social Networks: Understanding How Work Really Gets Done in Organizations,. Harvard Business Review Press.
McCarty, C., Freeman, L., <http://moreno.ss.uci.edu/data.html>, 2008.
Gil-Mendieta, J., Schmidt, S., 1996. The political network in Mexico. Social Networks 18 (4), 355–381.
Krebs, V., 2003. Proxy networks: analyzing one network to reveal another. Bulletin de Méthodologie Sociologique 79. 61-40.
Gemmetto, V., Barrat, A., Cattuto, C., 2014. Mitigation of infectious disease at school: targeted class closure vs. school closure. BMC Infect. Dis. 14 (695), 1–10.
MacRae, D., 1960. Direct factor analysis of sociometric data. Sociometry 23 (4), 360–371.
Loomis, C.P., Morales, J.O., Clifford, R.A., Leonard, O.E., 1953. Turrialba Social Systems and the Introduction of Change. The Free Press, Glencoe, IL, USA, pp. 45–78.
Scott, J.P., 1980. The Anatomy of Scottish Capital: Scottish Companies and Scottish Capital, 1900–1979,. Croom Helm.
Grimmer, J., 2010. A Bayesian hierarchical topic mode for political texts: measuring expressed agendas in senate press releases. Political Analysis 18 (1), 1–35.
Michael, J.H., 1997. Labor dispute reconciliation in a forest products manufacturing facility. Forest Prod. J. 47 (11–12), 41–45.
Schwimmer, E., 1973. Exchange in the Social Structure of the Orokaiva: Traditional and Emergent Ideologies in the Northern District of Papua. C Hurst and Co-Publishers Ltd.
Pearson, M., Michell, L., 2000. Smoke rings: social network analysis of friendship groups, smoking and drug-taking. Drugs: Educ. Prev. Policy 7 (1), 21–37.
Nepusz, T., Petroczi, A., Negyessy, L., Bazso, F., 2008. Fuzzy communities and the concept of bridgeness in complex networks. Phys. Rev. E 77 (1), 016107.
Freeman, L.C., Webster, C.M., Kirke, D.M., 1998. Exploring social structure using dynamic three-dimensional color images. Social Networks 20 (2), 109–118.
Freeman, L.C., Freeman, S.C., Michaelson, A.G., 1989. How humans see social groups: a test of the Sailer-Gaulin models. J. Quant. Anthropol. 1, 229–238.
Smith, D.A., White, D.R., 1992. Structure and dynamics of the global economy: network analysis of international trade 1965–1980. Soc. Forces 70 (4), 857–893.
Li, C., Li, Q., Van Mieghem, P., Stanley, H.E., Wang, H., 2015. Correlation between centrality metrics and their application to the opinion model. Eur. Phys. J. B 88 (65), 1–13.
Meghanathan, N., Correlation coefficient analysis of centrality metrics for complex network graphs. In: Proceedings of the 4th Computer Science Online Conference, (CSOC-2015), Intelligent Systems in Cybernetics and Automation Theory: Advances in Intelligent Systems and Computing, Vol. 348, pp. 11–20, April 27–30, 2015.
Freeman, L., 1977. A set of measures of centrality based on betweenness. Sociometry 40 (1), 35–41.
Meghanathan, N., 2017. A computationally-lightweight and localized centrality metric in lieu of betweenness centrality for complex network analysis. Springer Vietnam J. Comput. Sci. 4 (1), 23–38.
Greene, W.H., 2011. Econometric Analysis. Pearson.
Chatterjee, K., Dutta, B., 2016. Credibility and strategic learning in networks. Int. Econ. Rev. 57 (3), 759–786.
Kang, C., Molinaro, C., Kraus, S., Shavitt, Y., Subrahmanian, V.S., Diffusion centrality in social networks. In: Proceedings of the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, pp. 558–564, Istanbul, Turkey, August 2012.
Ide, K., Namatame, A., Ponnambalam, L., Xiuju, F., Goh, R.S.M., 2014. A New centrality measure for probabilistic diffusion in network. Adv. Comput. Sci. Int. J. 3 (5), 115–121.
Banerjee, A., Chandrasekhar, A.G., Duflo, E., Jackson, M.O., 2013. The diffusion of microfinance. Science 341 (6144), 1–49.
Dangalchev, C., 2006. Residual closeness in networks. Phys. A 365, 556–564.
Fiedler, M., 1973. Algebraic connectivity of graphs. Czechoslovak Math. J. 23 (98), 298–305.
Mohar, B., 1992. Laplace Eigenvalues of graphs – a survey. Discrete Math. 109 (1–3), 171–183.
Meghanathan, N., Algebraic connectivity based stable data gathering trees for mobile sensor networks. In: Proceedings of the International Conference on Computational Science and Computational Intelligence (CSCI'16), pp. 947–952, Las Vegas, USA, Dec 15–17, 2016.
Newman, M.E.J., 2006. Modularity and community structure in networks. Proc. Natl. Acad. Sci. U.S.A. 103 (23), 8577–8696.
Bastian, M., Heymann, S., Jacomy, M., Gephi: an open source software for exploring and manipulating networks. In: Proceedings of the 3rd International Conference on Weblogs and Social Media, pp. 361–362, San Jose, CA, USA, May 2009.
Newman, M.E.J., 2002. Assortative mixing in networks. Phys. Rev. Lett. 89, 208701.