

Assortativity Analysis of Real-World Network Graphs Based on Centrality Metrics

Natarajan Meghanathan¹

¹ Department of Computer Science, Jackson State University, USA

Correspondence: Natarajan Meghanathan, Department of Computer Science, Mailbox 18839, Jackson State University, Jackson, MS 39217, USA. Tel: 1-601-979-3661. E-mail: natarajan.meghanathan@jsums.edu

Received: May 8, 2016

Accepted: June 9, 2016

Online Published: July 3, 2016

doi:10.5539/cis.v9n3p7

URL: <http://dx.doi.org/10.5539/cis.v9n3p7>

The research is financed by the NASA EPSCoR sub award (#: NNX14AN38A) from University of Mississippi.

Abstract

Assortativity index (*A. Index*) of real-world network graphs has been traditionally computed based on the degree centrality metric and the networks were classified as assortative, disassortative or neutral if the *A. Index* values are respectively greater than 0, less than 0 or closer to 0. In this paper, we evaluate the *A. Index* of real-world network graphs based on some of the commonly used centrality metrics (betweenness, eigenvector and closeness) in addition to degree centrality and observe that the assortativity classification of real-world network graphs depends on the node-level centrality metric used. We also propose five different levels of assortativity (strongly assortative, weakly assortative, neutral, weakly disassortative and strongly disassortative) for real-world networks and the corresponding range of *A. Index* value for the classification. We analyze a collection of 50 real-world network graphs with respect to each of the above four centrality metrics and estimate the empirical probability of observing a real-world network graph to exhibit a particular level of assortativity. We claim that a real-world network graph is more likely to be neutral with respect to the betweenness and degree centrality metrics and more likely to be assortative with respect to the eigenvector and closeness centrality metrics.

Keywords: Assortativity Index, Centrality, Correlation, Real-World Network Graph

1. Introduction

The assortativity index (*A. Index*) of a network is a measure of the similarity of the end vertices of the edges in the network with respect to a particular node-level metric (Newman, 2010). That is, the *A. Index* of a network is a measure of the extent to which a vertex with a higher value for a particular node-level metric is connected to another vertex that also has a higher value for the node-level metric. Since the *A. Index* is nothing but a correlation coefficient (Pearson's product-moment correlation coefficient) (Newman, 1999) quantifying the extent of similarity of the end vertices of the edges, its value ranges from -1 to 1 (Strang, 2006). Traditionally, in the literature (Newman, 1999), networks with positive values of *A. Index* (closer to 1) are referred to as assortative networks; networks with negative values of *A. Index* (closer to -1) are referred to as disassortative networks and networks with *A. Index* values closer to 0 are classified as neutral. The similarity has been typically evaluated with respect to the degree centrality metric of the vertices, and the classification of networks (as either as assortative, disassortative or neutral) has been so far only based on the degree centrality metric (Newman & Girvan, 2003; Noldus & Van Mieghem, 2015).

Our hypothesis in this paper is that the assortativity classification of a real-world network could depend on the centrality metric used to compute the *A. Index* value of the network. In other words, a network could be classified as assortative with respect to one centrality metric and it could end up being classified as disassortative or neutral with respect to another centrality metric. Also, just having three different levels (assortative, neutral and disassortative) would not be sufficient to accurately assess the extent of assortativity of real-world network graphs whose *A. Index* values are neither close to 0, but nor close to 1 or -1. Until now, a formal range of *A. Index* values has not been defined to assess the level of assortativity of real-world networks. In this paper, we propose to divide the range of values (-1.0 to 1.0) for the *A. Index* fairly even to five levels and setup the following rule: strongly assortative ($0.6 \leq A. Index \leq 1.0$), weakly assortative ($0.2 \leq A. Index < 0.6$), neutral (-0.2

$< A. Index < 0.2$), weakly dissortative ($-0.6 < A. Index \leq -0.2$) and strongly dissortative ($-0.2 < A. Index \leq -1.0$).

We investigate the validity of our hypothesis by analyzing a broader collection of 50 real-world networks whose spectral radius ratio for node degree (a measure of the variation in node degree) ranges from 1.01 to 3.48 (Meghanathan, 2014). We compute the *A. Index* values for these 50 real-world networks with respect to each of the four commonly used centrality metrics: degree centrality (DegC), eigenvector centrality (EVC; Bonacich, 1987), betweenness centrality (BWC; Freeman, 1977) and closeness centrality (CIC; Freeman, 1979), and apply the above proposed range of values to assess the assortativity levels of the real-world networks with respect to each of these four centrality metrics. For 40 of the 50 real-world networks analyzed, we observe that the level of classification of the network (strongly or weakly assortative, neutral, strongly or weakly dissortative) depends on the centrality metric under consideration.

Since we have analyzed a vast collection of networks with varying levels of complexity, we use the results of our assortativity analysis to empirically propose the likelihood of a real-world network being classified neutral or assortative (strongly assortative or weakly assortative) with respect to a particular centrality metric. Based on the results of our assortativity analysis on 50 real-world networks, we claim that any chosen real-world network is more likely (i.e., with a probability of 0.72) to be classified as neutral (neither assortative nor dissortative) with respect to the betweenness centrality metric, and more likely (i.e., with a probability of 0.66) to be classified as assortative (strongly or weakly) with respect to the CIC and EVC metrics. More specifically, we expect a chosen real-world network to be somewhat strongly assortative (with a probability of 0.38) with respect to the CIC metric and somewhat weakly assortative (also with a probability of 0.38) with respect to the EVC metric.

To the best of our knowledge, we have not come across a paper that has conducted a comprehensive assortativity analysis of complex real-world networks with respect to the four commonly used centrality metrics as well as empirically proposed the likelihood of observing a real-world network to be neutral, strongly assortative or weakly assortative with respect to a particular centrality metric. The rest of the paper is organized as follows: Section 2 reviews the four centrality metrics along with an example to illustrate their computation on a sample graph. Section 3 introduces the formulation for Assortativity Index (*A. Index*) and the proposed range of *A. Index* values to classify the assortativity level of a real-world network as well as presents a motivating example to illustrate that the *A. Index* of a network and its classification (as neutral, strongly/weakly assortative or dissortative) could depend on the centrality metric under consideration. Section 4 introduces the 50 complex real-world networks that are analyzed in this paper. Section 5 presents the results of assortativity analysis conducted on the real-world networks with respect to the four centrality metrics. Section 6 discusses related work and highlights the novel contribution of the work done in this paper. Section 7 concludes the paper. Throughout the paper, we use the terms 'node' and 'vertex', 'link' and 'edge', 'network' and 'graph' interchangeably. They mean the same. All the real-world networks analyzed in this paper are modeled as undirected graphs.

2. Centrality Metrics

The four commonly used centrality metrics in complex network analysis are: degree centrality (DegC), eigenvector centrality (EVC; Bonacich, 1987), betweenness centrality (BWC; Freeman, 1977) and closeness centrality (CIC; Freeman, 1979). DegC and EVC are degree-based centrality metrics; whereas BWC and CIC are shortest path-based centrality metrics. Until now, the DegC metric has been typically used for assortativity analysis of real-world networks (Newman & Girvan, 2003; Noldus & Van Mieghem, 2015). In this paper, we are interested in conducting assortativity analysis of real-world networks with respect to all the above four centrality metrics. In this section, we briefly review these four centrality metrics and the procedure to compute them, along with an example for each.

2.1 Degree Centrality

The degree centrality (DegC) of a vertex is the number of edges incident on the vertex. The DegC of the vertices is computed by multiplying the adjacency matrix of the graph with a unit vector of 1s (the number of 1s in the unit vector corresponds to the number of vertices in the graph). Figure 1 illustrates an example to compute the degree centrality of the vertices. As can be noticed from this example, the DegC metric is vulnerable to incurring several ties among the vertices (as the metric values are integers and not real numbers).

2.2 Eigenvector Centrality

The eigenvector centrality (EVC) of a vertex is a measure of the degree of the vertex as well as the degree of its neighbors. The EVC values of the vertices in a graph correspond to the entries in the principal eigenvector of the adjacency matrix of the graph. We use the JAMA: A Java Matrix package

(<http://math.nist.gov/javanumerics/jama/>) to compute the principal eigenvector of the adjacency matrix of a real-world network graph. The entries in the principal eigenvector can also be computed using the Power-Iteration method (Strang, 2006) that is illustrated in Figure 2. The tentative eigenvector X_{i+1} of a network

graph at the end of the $(i+1)^{th}$ iteration is given by: $X_{i+1} = \frac{AX_i}{\|AX_i\|}$, where $\|AX_i\|$ is the normalized value of the

product of the adjacency matrix and the tentative eigenvector X_i at the end of the i^{th} iteration. We continue the iterations until the normalized value of the product vector converges and the tentative eigenvector at that juncture corresponds to the principal eigenvector of the adjacency matrix of the graph. As the EVC values of the vertices are likely to be real numbers and are dependent on the degree of a vertex as well as the degrees of its neighbors, the EVC values of the vertices are more likely to be unique and relatively fewer ties are incurred (compared to degree centrality).

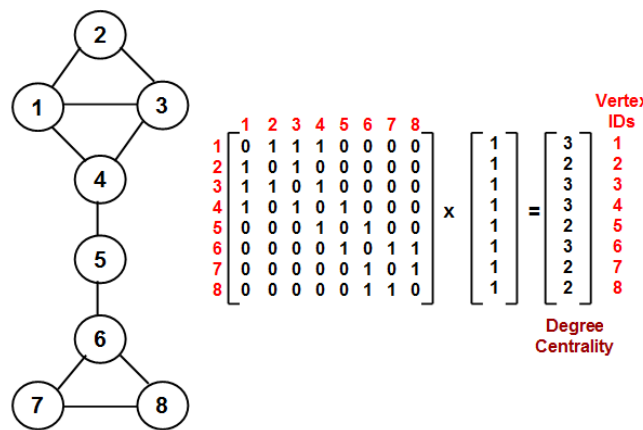


Figure 1. Example to Illustrate the Computation of the Degree Centrality of the Vertices in a Graph

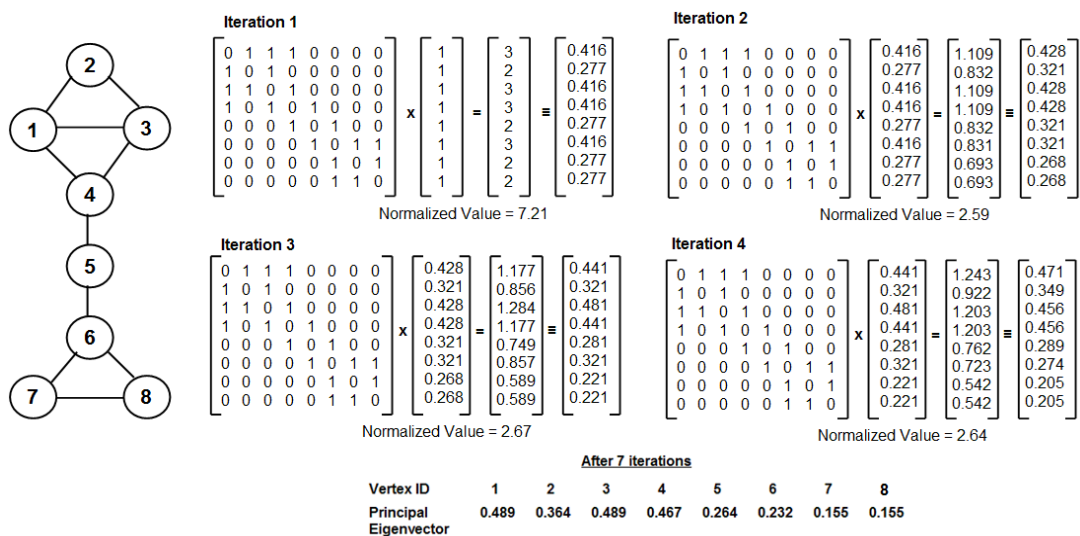


Figure 2. Example to Illustrate the Computation of the Eigenvector Centrality of the Vertices in a Graph

2.3 Betweenness Centrality

The betweenness centrality (BWC) of a vertex is a measure of the fraction of shortest paths going through the vertex when considered across all pairs of vertices in the graph (Freeman, 1977). If $sp_{jk}(i)$ is the number of

shortest paths between vertices j and k that go through vertex i and sp_{jk} is the total number of shortest paths between vertices j and k , then $BWC(i) = \sum_{j \neq k \neq i} \frac{sp_{jk}(i)}{sp_{jk}}$. The BWC of vertices is computed using a Breadth First

Search (BFS; Cormen et al., 2009)-based implementation of the Brandes' algorithm (Brandes, 2001). We use the BFS algorithm to determine the shortest path trees rooted at each vertex and thereby deduce the level numbers of the vertices in the shortest path trees rooted at every vertex in the graph. The level number of a vertex i in the shortest path tree rooted at vertex j is the minimum number of hops from vertex j to i . The root vertex of a shortest path tree is said to be at level 0. The number of shortest paths from a vertex j to itself is 1. The number of shortest paths from a vertex j to a vertex k (at level l in the shortest path tree rooted at vertex j) is the sum of the number of shortest paths from vertex j to each of the vertices that are neighbors of vertex k in the graph and located at level $l-1$ in the shortest path tree rooted at j . The number of shortest paths between vertices j and k that go through vertex i is the maximum of the number of shortest paths from vertex j to vertex i and the number of shortest paths from vertex k to vertex i . Figure 3 illustrates an example for the computation of the BWC of the vertices in the same sample graph used in Figures 1-2. We notice that vertices with a high degree and/or EVC need not have a high BWC and vice-versa. For example, vertices 0 and 2 that had the largest value for the EVC metric have relatively low BWC value; whereas, vertex 4 (with a low degree and low EVC) has the largest value for the BWC.

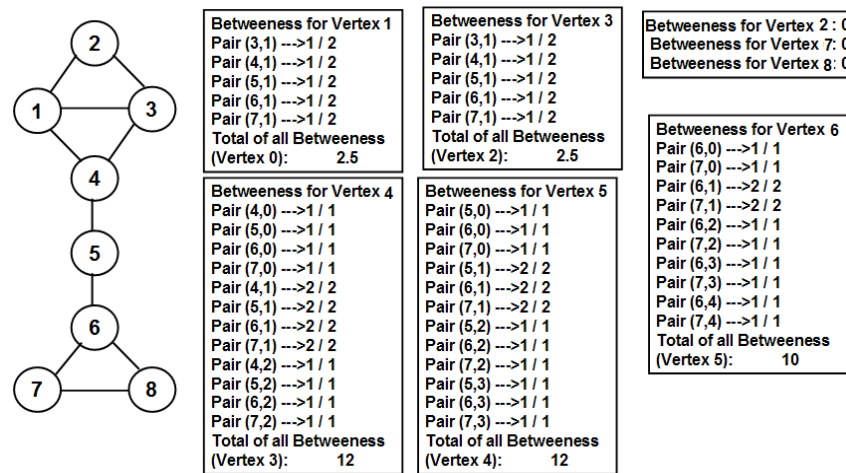


Figure 3. Example to Illustrate the Computation of the Betweenness Centrality of the Vertices in a Graph

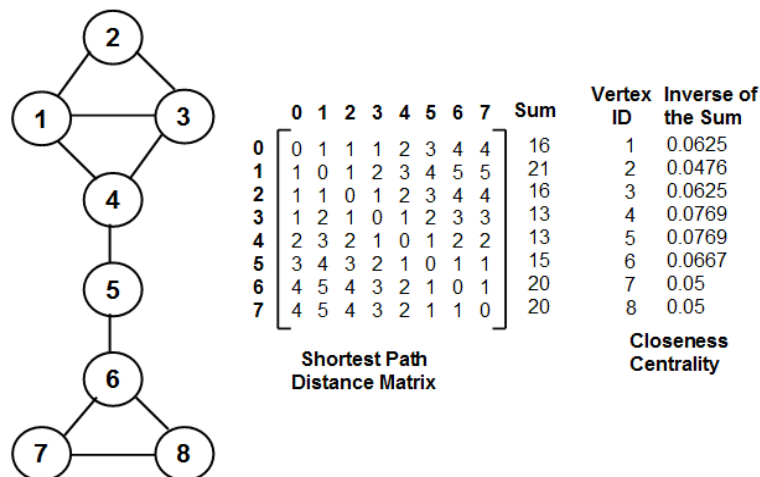


Figure 4. Example to Illustrate the Computation of the Closeness Centrality of the Vertices in a Graph

2.4 Closeness Centrality

The closeness centrality (CIC) of a vertex (Freeman, 1979) is a measure of the relative closeness of the vertex with the rest of the vertices in the graph. The CIC of a vertex is measured by running the BFS algorithm on the vertex and determining the minimum number of hops to each vertex on the shortest path tree rooted at the vertex. The CIC of a vertex is the inverse of the sum of the shortest path lengths (hop counts) to the rest of the vertices in the graph. Figure 4 illustrates an example for the calculation of CIC of the vertices on the same sample graph used in Figures 1-3.

3. Assortativity Analysis

3.1 Network Model

Let $G = (V, E)$ be the set of vertices and edges constituting a real-world network and let $C(i)$ be the value of a centrality metric (C) for any node i in the network. We refer to the first vertex (vertex u) in an edge (u, v) as the upstream vertex and the second vertex (vertex v) in an edge (u, v) as the downstream vertex. As the focus of this research is on undirected graphs, we conveniently adopt the following convention to represent the edges: the ID of the upstream vertex of an edge (u, v) is always less than the ID of the downstream vertex of the edge (i.e., $u < v$). Let U and D be respectively the set of upstream and downstream vertices

constituting the edges of a graph. Let $C_{\bar{U}}$ and $C_{\bar{V}}$ (calculated as in formulation-1 below) be respectively the average values for the centrality metric of interest among the vertices constituting the sets U and V .

$$C_{\bar{U}} = \frac{1}{|U|} \sum_{(u,v) \in E} C(u) \quad C_{\bar{V}} = \frac{1}{|V|} \sum_{(u,v) \in E} C(v) \quad (1)$$

3.2 Assortativity Index

The Assortativity Index (*A. Index*) of a network (Newman, 1999) with respect to a particular node-level centrality metric is a quantitative measure of the extent of similarity of the end vertices of the edges with respect to the chosen centrality metric. The extent of similarity is calculated as the Pearson's Product-Moment Correlation Coefficient (Strang, 2006) between the set of upstream vertices (U) and set of downstream vertices (D) constituting the end vertices of the edges in a real-world network graph. Accordingly, the *A. Index* of a network with respect to a centrality metric C could be formulated as below.

$$A.Index_C = \frac{\sum_{(u,v) \in E} [C(u) - C_{\bar{U}}][C(v) - C_{\bar{V}}]}{\sqrt{\sum_{(u,v) \in E} [C(u) - C_{\bar{U}}]^2} \sqrt{\sum_{(u,v) \in E} [C(v) - C_{\bar{V}}]^2}} \quad (2)$$

3.3 Range of Values for Assortativity Classification

As Assortativity Index is a measure of the level of correlation between the sets of upstream and downstream vertices constituting the edges in a network graph, the values for $A.Index_C$ with respect to any centrality metric (C) would range from -1 to 1. Until now in the literature, a network is considered to be assortative (disassortative) with respect to the chosen node-level metric (C) if the $A.Index_C$ value is closer to 1 (-1). If the $A.Index_C$ value is closer to 0, the network is considered to be neutral with respect to the metric C . However, we do not have a formally defined range of values that clearly indicate how the network should be classified if the $A.Index_C$ values are neither close to 1 or -1 and nor to 0.

Table 1. Range of Correlation Coefficient Values and the Corresponding Levels of Correlation

Range of Correlation Coefficient Values	Level of Correlation	Range of Correlation Coefficient Values	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

We seek to address this concern as follows: Since $A. Index_C$ is evaluated as a measure of correlation, we will adapt the range of correlation coefficient values (rounded to two decimals) proposed in the literature (Evans, 1995) for the level of correlation (shown in Table 1) and propose the range of assortativity index values (shown in Table 2) for classifying a network with respect to the level of assortativity. We propose only two levels of assortativity and two levels of disassortativity (rather than 5 levels for each) to give enough space for the range of $A. Index$ values to classify a network at a particular level, including neutral (i.e., neither assortative nor disassortative), but still be able to differentiate a strongly assortative (disassortative) network from a weakly assortative (disassortative) network or neutral network with respect to a node-level metric. The color code to be used for the various levels of assortativity are also shown in Table 2.

Table 2. Range of Assortativity Index Values and the Corresponding Levels of Assortativity

Range of Assortativity Index Values	Level of Assortativity	Range of Assortativity Index Values	Level of Assortativity
0.60 to 1.00	Strongly Assortative	-1.00 to -0.60	Strongly Disassortative
0.20 to 0.59	Weakly Assortative	-0.59 to -0.20	Weakly Disassortative
0.00 to 0.19	Neutral	-0.19 to -0.01	Neutral

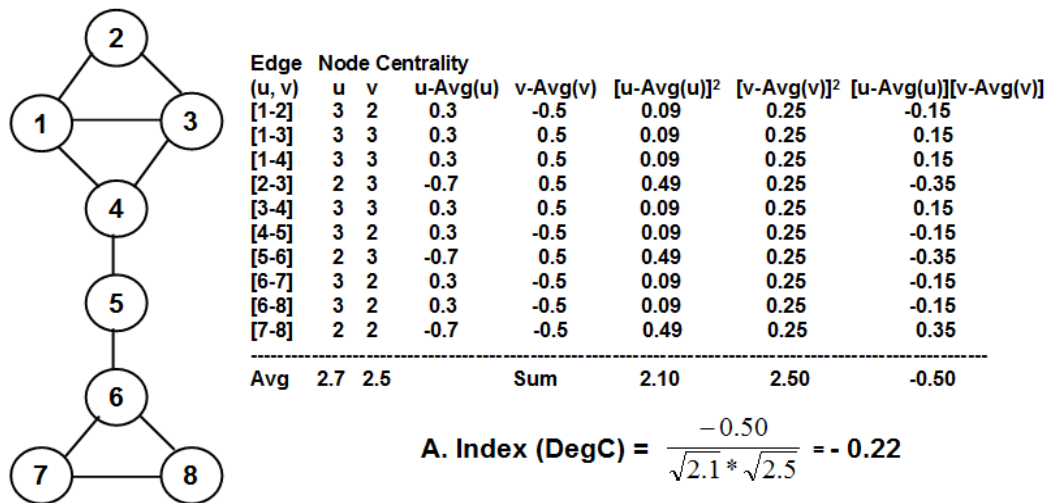


Figure 5. Example to Illustrate the Calculation of the Assortativity Index based on Degree Centrality

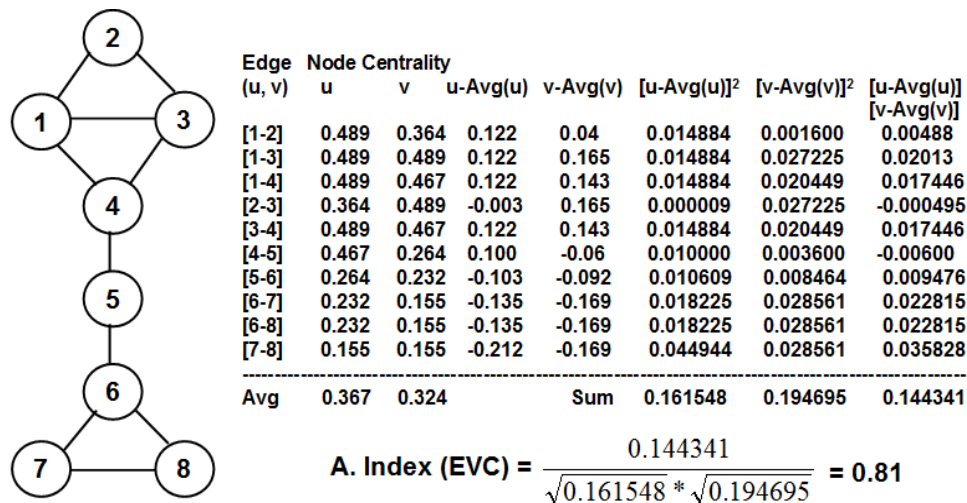


Figure 6. Example to Illustrate the Calculation of the Assortativity Index based on Eigenvector Centrality

3.4 Motivating Example

In this sub section, we illustrate the computation of the assortativity index of the sample graph used in Figures 1-4 with respect to the degree centrality (Figure 5) and eigenvector centrality (Figure 6) metrics. Adopting the proposed range of classification for the level of assortativity, we notice that the sample graph of Figures 1-4 could be classified as "weakly dissortative" (*A. Index* value of -0.22; see Figure 5) with respect to the degree centrality metric and "strongly assortative" (*A. Index* value of 0.81; see Figure 6) with respect to the eigenvector centrality metric. This is a motivating example to vindicate our hypothesis that the assortativity level classification for a network could vary depending on the centrality metric used to compute the *A. Index* values.

4. Real-World Network Graphs

We now present a brief overview of the 50 real-world network graphs analyzed in this paper. We model each network as an undirected graph of nodes and edges. The networks are identified with a unique ID (1, ..., 50) and a three character acronym. We use the spectral radius ratio for node degree (Meghanathan, 2014) to capture the extent of variation in the degree of the nodes: the spectral radius ratio for node degree is the ratio of the principal eigenvalue (Strang, 2006) of the adjacency matrix of the network graph to that of the average node degree. The values for the spectral radius ratio for node degree are 1 or above; the farther is the value from 1, the larger is the variation in node degree. We analyze real-world networks ranging from random networks to scale-free networks as the spectral radius ratio for node degree of the real-world networks analyzed in this graph ranges from 1.01 to 3.48. Table 3 lists the 50 networks along with the values for the number of nodes and edges, the spectral radius ratio for node degree (λ_{sp}) and average degree (k_{avg}). A brief description of the networks is as follows:

- 1) Word Adjacency Network (ADJ; Newman, 2006): This is a network of 112 words (adjectives and nouns, represented as vertices) in the novel *David Copperfield* by Charles Dickens; there exists an edge between two vertices if the corresponding words appeared adjacent to each other at least once in the novel.
- 2) Anna Karenina Network (AKN; Knuth, 1993): This a network of 140 characters (vertices) in the novel *Anna Karenina*; there exists an edge between two vertices if the corresponding characters have appeared together in at least one scene in the novel.
- 3) Jazz Band Network (JBN; Geiser & Danon, 2003): This is a network of 198 Jazz bands (vertices) that recorded between the years 1912 and 1940; there exists an edge between two bands if they shared at least one musician in any of their recordings during this period.
- 4) C. Elegans Neural Network (CEN; White et al., 1986): This is a network of 297 neurons (vertices) in the neural network of the hermaphrodite *Caenorhabditis Elegans*; there is an edge between two vertices if the corresponding neurons interact with each other (in the form of chemical synapses, gap junctions and neuromuscular junctions).
- 5) Centrality Literature Network (CLN; Hummon et al., 1990): This is a network of 118 papers (vertices) published on the topic of centrality in complex networks from 1948 to 1979. There is an edge between two vertices v_i and v_j if one of the corresponding papers has cited the other paper as a reference.
- 6) Citation Graph Drawing Network (CGD; Biedl & Franz, 2001): This is a network of 259 papers (vertices) that were published in the Proceedings of the Graph Drawing (GD) conferences from 1994 to 2000 and cited in the papers published in the GD'2001 conference. There is an edge between two vertices v_i and v_j if one of the corresponding papers has cited the other paper as a reference.
- 7) Copperfield Network (CFN; Knuth, 1993): This is a network of 89 characters in the novel *David Copperfield* by Charles Dickens; there exists an edge between two vertices if the corresponding characters appeared together in at least one scene in the novel.
- 8) Dolphin Network (DON; Lusseau et al., 2003): This is a network of 62 dolphins (vertices) that lived in the Doubtful Sound fiord of New Zealand; there is an edge between two vertices if the corresponding dolphins were seen moving with each other during the observation period.
- 9) Drug Network (DRN; Lee, 2004): This is a network of 212 drug agents (vertices) of different ethnicities. There is a link between two vertices if the corresponding agents know each other.
- 10) Dutch Literature 1976 Network (DLN; Nooy, 1999): This is a network of 37 Dutch literary authors and critics (vertices) in 1976; there exists an edge between two vertices v_i and v_j if the person corresponding to one of them is a critic who made a judgment (through a review or interview) on the literature work of the author corresponding to the other vertex.

Table 3. Fundamental Properties of the Real-World Network Graphs used for Assortativity Analysis

#	Net.	λ_{sp}	#nodes	#edges	k_{avg}	#	Net.	λ_{sp}	#nodes	#edges	k_{avg}
1	ADJ	1.73	112	425	7.589	26	LMN	1.82	77	254	6.597
2	AKN	2.48	140	494	7.057	27	MDN	1.04	62	1167	37.645
3	JBN	1.45	198	2742	27.697	28	MTB	1.95	64	295	9.219
4	CEN	1.68	297	2148	14.465	29	MCE	1.12	77	1549	40.23
5	CLN	2.03	118	613	10.39	30	MSJ	3.48	475	625	2.632
6	CGD	2.24	259	640	4.942	31	AFB	2.29	171	940	10.994
7	CFN	1.83	89	407	9.146	32	MPN	1.23	35	117	6.686
8	DON	1.40	62	159	5.129	33	MMN	1.59	30	61	4.067
9	DRN	2.76	212	284	2.679	34	PBN	1.42	105	441	8.4
10	DLN	1.49	37	81	4.378	35	PSN	1.22	238	5539	46.546
11	ERD	3.00	433	1314	6.069	36	PFN	1.32	67	142	4.239
12	FMH	2.81	147	202	2.748	37	SJN	1.29	75	155	4.133
13	FHT	1.57	33	91	5.515	38	SDI	1.94	230	359	3.122
14	FTC	1.21	48	170	7.083	39	SPR	1.57	92	477	10.37
15	FON	1.01	115	613	10.661	40	SMN	1.05	126	5973	94.81
16	CDF	1.11	58	967	33.345	41	SWC	1.45	35	118	6.743
17	GD96	2.38	180	228	2.533	42	SSM	1.22	24	38	3.167
18	MUN	2.54	167	301	3.605	43	TEN	1.06	22	39	3.545
19	GLN	2.01	67	118	3.522	44	TWF	1.49	47	77	3.277
20	GD01	1.80	101	190	3.762	45	UKF	1.35	83	578	13.928
21	HTN	1.21	115	2164	37.635	46	APN	3.22	332	2126	12.807
22	HCN	1.66	76	302	7.947	47	USS	1.25	49	107	4.367
23	ISP	1.69	309	1924	12.453	48	RHF	1.27	217	1839	16.949
24	KCN	1.47	34	78	4.588	49	WSB	1.22	43	336	15.628
25	KFP	1.70	37	85	4.595	50	WTN	1.38	80	875	21.875

11) Erdos Collaboration Network (ERD; Batagelj & Mrvar, 2006): This is a network of 433 authors (nodes) who have either directly published an article with Paul Erdos or through a chain of collaborators leading to Paul Erdos. There is an edge between two nodes if the corresponding authors have co-authored at least one publication.

12) Faux Mesa High School Friendship Network (FMH; Resnick et al., 1997): This is a network of 147 students (vertices) at a high school community in the rural western part of US; there exists an edge between two vertices if the corresponding students are friends of each other.

13) Friendship Ties in a Hi-Tech Firm (FHT; Krackhardt, 1999): This is a network of 33 employees (vertices) of a small hi-tech computer firm that sells, installs and maintains computer systems; there exists an edge between two vertices v_i and v_j if the employee corresponding to at least one of them considers the employee corresponding to the other vertex as a personal friend.

14) Flying Teams Cadet Network (FTC; Moreno, 1960): This is a network of 48 cadet pilots (vertices) at an US Army Air Forces flying school in 1943 and the cadets were trained in a two-seated aircraft; there exists an edge between two vertices v_i and v_j if the pilot corresponding to at least one of them has indicated the pilot corresponding to the other vertex as a preferred partner with whom s/he likes to fly during the training schedules.

15) US Football Network (FON; Girvan & Newman, 2002): This is a network of 115 football teams (nodes) of US universities that played in the Fall 2000 season; there is an edge between two nodes if the corresponding teams have played against each other in the league games.

16) College Dorm Fraternity Network (CDF; Bernard et al., 1980): This is a network of 58 residents (vertices) in a fraternity college at a West Virginia college; there exists an edge between two vertices if the corresponding residents were seen in a conversation at least once during a five day observation period.

17) GD'96 Network (GD96; Batagelj & Mrvar, 2006): This is a network of 180 AT&T and other WWW websites (vertices) that were cited in the proceedings of the Graph Drawing (GD) conference in 1996; there exists an edge between two vertices if the website corresponding to one of them has a link to the website corresponding to the other vertex.

- 18) Marvel Universe Network (MUN; Gleiser, 2007): This is a collaborative network of 167 characters (vertices) in the comic books published by the Marvel Universe publishing company; there exists an edge between two vertices if the corresponding characters had appeared together in at least one publication.
- 19) Graph and Digraph Glossary Network (GLN; Batagelj & Mrvar, 2006): This is a network of 67 terms (vertices) that appeared in the glossary prepared by Bill Cherowitzo on Graph and Digraph; there appeared an edge between two vertices if the term corresponding to one of them is used to describe the meaning of the term corresponding to the other vertex.
- 20) Graph Drawing 2001 (GD01) Network (Batagelj & Mrvar, 2006): This is a network of 101 papers (vertices) that were cited as references in the papers published in the proceedings of the 2001 Graph Drawing (GD'01) conference; there exists an edge between two vertices if the corresponding papers have been co-cited in at least one paper published in the GD'01 conference.
- 21) Hypertext 2009 Network (HTN; Isella et al., 2011): This is a network of the face-to-face contacts of 115 attendees (vertices) of the ACM Hypertext 2009 conference held in Turin, Italy from June 29 to July 1, 2009. There exists an edge between two vertices if the corresponding conference visitors had face-to-face contact that was active for at least 20 seconds.
- 22) Huckleberry Coappearance Network (HCN; Knuth, 1993): This is a network of 76 characters (vertices) that appeared in the novel Huckleberry Finn by Mark Twain; there is an edge between two vertices if the corresponding characters had a common appearance in at least one scene.
- 23) Infectious Socio-patterns Network (ISP; Isella et al., 2011): This is a network of 309 visitors (vertices) who visited the Science Gallery in Dublin, Ireland during Spring 2009. There existed an edge between two vertices if the corresponding visitors had a continuous face-to-face contact for at least 20 seconds when they participated in the *Infectious Socio-patterns* event (an electronic simulation of the spreading of an epidemic through individuals who are in close proximity) as part of an art science exhibition.
- 24) Karate Club Network (KCN; Zachary, 1977): This is a network of 34 members (nodes) of a Karate Club at a US university in the 1970s; there is an edge between two nodes if the corresponding members were seen interacting with each other during the observation period.
- 25) Korea Family Planning Network (KFP; Rogers & Kincaid, 1980): This is a network of 37 women (vertices) at a Mothers' Club in Korea; there existed an edge between two vertices if the corresponding women were seen discussing family planning methods during an observation period.
- 26) Les Miserables Network (LMN; Knuth, 1993): This is a network of 77 characters (nodes) in the novel *Les Miserables*; there exists an edge between two nodes if the corresponding characters appeared together in at least one of the chapters in the novel.
- 27) Macaque Dominance Network (MDN; Takahata, 1991): This is a network of 62 adult female Japanese macaques (monkeys; vertices) in a colony, known as the "Arashiyama B Group", recorded during the non-mating season from April to early October, 1976. There existed an edge between two vertices if a macaque corresponding to one of them was recorded to have exhibited dominance over the macaque corresponding to the other vertex.
- 28) Madrid Train Bombing Network (MTB; Hayes, 2006): This is a network of 64 suspected individuals and their relatives (vertices) reconstructed by Rodriguez using press accounts in the two major Spanish daily newspapers (El Pais and El Mundo) regarding the bombing of commuter trains in Madrid on March 11, 2004. There existed an edge between two vertices if the corresponding individuals were observed to have a link in the form of friendship, ties to any terrorist organization, co-participation in training camps and/or wars, or co-participation in any previous terrorist attacks.
- 29) Manufacturing Company Employee Network (MFE; Cross et al., 2004): This is a network of 77 employees (nodes) from a research team in a manufacturing company; there exists an edge between two nodes if the two employees are aware of each other's knowledge and skills.
- 30) Social Networks Journal Co-authors (MSJ; McCarty & Freeman, 2008): This is a network of 475 authors (vertices) involved in the production of 295 articles for the Social Networks Journal since its inception until 2008; there is an edge between two vertices if the corresponding authors co-authored at least one paper published in the journal.
- 31) Author Facebook Network (AFB): This is a network of the 171 friends (vertices) of the author in Facebook. There exists an edge between two vertices if the corresponding people are also friends of each other.

- 32) Mexican Political Elite Network (MPN; Gil-Mendieta & Schmidt, 1996): This is a network of 35 Mexican presidents and their close collaborators (vertices); there exists an edge between two vertices if the corresponding two people have ties that could be either political, kinship, friendship or business ties.
- 33) ModMath Network (MMN; Batagelj & Mrvar, 2006): This is a network of 30 school superintendents (vertices) in Allegheny County, Pennsylvania, USA during the 1950s and early 1960s. There exists an edge between two vertices if at least one of the two corresponding superintendents has indicated the other person as a friend in a research survey conducted to see which superintendents (who are in office for at least a year) are more influential to effectively spread around some modern Math methods among the school systems in the county.
- 34) US Politics Books Network (PBN; Krebs, 2003): This is a network of 105 books (vertices) about US politics sold by Amazon.com around the time of the 2004 US presidential election. There exists an edge between two vertices if the corresponding two books were co-purchased by the same buyer (at least one buyer).
- 35) Primary School Contact Network (PSN; Gemmetto et al., 2014): This is a network of children and teachers (238 vertices) used in the study published by an article in BMC Infectious Diseases, 2014 [40]. There exists an edge between two vertices if the corresponding persons were in contact for at least 20 seconds during the observation period.
- 36) Prison Friendship Network (PFN; MacRae, 1960): This is a network of 67 prison inmates (vertices) surveyed by John Gagnon in the 1950s regarding their sociometric choice. There exists an edge between two vertices if an inmate corresponding to at least one of them has listed the inmate corresponding to the other vertex as one of his/her closest friends.
- 37) San Juan Sur Family Network (SJN; Loomis et al., 1953): This is a network of 75 families (vertices) in San Juan Sur, Costa Rica, 1948. There exists an edge between two vertices if at least one of the corresponding families has visited the household of the family corresponding to the other vertex once or more.
- 38) Scotland Corporate Interlocks Network (SDI; Scott, 1980): This is a network of multiple directors (a director who serves on multiple boards) and companies (a total of 230 vertices) during 1904-05 in Scotland. There exists an edge between two vertices v_i and v_j if any of the following are true: (i) both v_i and v_j correspond to two different multiple directors who are in the board of at least one company; (ii) one of the two vertices corresponds to a multiple director and the other vertex corresponds to one of the companies in whose board the person serves.
- 39) Senator Press Release Network (SPR; Grimmer, 2010): This is a network of 92 US senators (vertices) during the period from 2007 to 2010. There exists an edge between two senators if they issued at least one joint press release.
- 40) Slovenian Magazine Network (SMN; Batagelj & Mrvar, 2006): This is a network of 126 different magazines (vertices); there exists an edge between two vertices if at least one reader (among a total of 100,000 readers) indicated that s/he reads the corresponding two magazines as part of a survey conducted in 1999 and 2000.
- 41) Soccer World Cup 1998 Network (SWC; Batagelj & Mrvar, 2006): This is a network of 35 teams (vertices) that participated in the 1998 edition of the Soccer World Cup. A player for a national team could sometimes have contract with one or more other countries. In this network, there is an edge between two vertices if the national team corresponding to at least one of them has contracted players from the country represented by the national team corresponding to the other vertex.
- 42) Sawmill Strike Communication Network (SSM; Michael, 1997): This is a network of 24 employees (vertices) in a sawmill who planned a strike against the new compensation package proposed by their management. There exists an edge between any two vertices if the corresponding employees mutually admitted discussing about the strike with a frequency of three or more (on a 5-point scale).
- 43) Taro Exchange Network (TEN; Schwimmer, 1973): This is a network of 22 families (vertices) in a Papuan village. There exists an edge between two vertices if the corresponding families were seen exchanging gifts during the observation period.
- 44) Teenage Female Friendship Network (TWF; Pearson & Michell, 2000): This is a network of 47 female teenage students (vertices) who studied as a cohort in a school in the West of Scotland from 1995 to 1997. There exists an edge between two vertices if the corresponding students reported (in a survey) that they were best friends of each other.
- 45) UK Faculty Friendship Network (UKF; Nepusz et al., 2008): This is a network of 83 faculty (vertices) at a UK university. There exists an edge between two vertices if the corresponding faculty are friends of each other.
- 46) US Airports 1997 Network (APN; Batagelj & Mrvar, 2006): This is a network of 332 airports (vertices) in

the US in the year 1997. There is an edge between two nodes if there is a direct flight connection between the corresponding airports.

47) US States Network (USS): This is a network of the 48 contiguous states in the US and the District of Columbia (DC). Each of the 48 states and DC is a node and there is an edge involving two nodes if the corresponding states (or DC) have a common border between them.

48) Residence Hall Friendship Network (RHF; Freeman et al., 1998): This is a network of 217 residents (vertices) living at a residence hall located on the Australian National University campus. There exists an edge between two vertices if the corresponding residents are friends of each other.

49) Windsurfers Beach Network (WSB; Freeman et al., 1989): This is a network of 43 windsurfers (vertices) on a beach in southern California during Fall 1986. There exists an edge between two vertices if the corresponding windsurfers were perceived to be close to each other (determined based on a survey).

50) World Trade Metal Network (WTN; Smith & White, 1992): This is a network of 80 countries (vertices) that are involved in trading miscellaneous metals during the period from 1965 to 1980. There exists an edge between two vertices if one of the two corresponding countries imported miscellaneous metals from the country corresponding to the other vertex.

5. Results of Assortativity Analysis

We now present the *A. Index* values obtained for each of the 50 real-world network graphs (listed in Section 4) with respect to each of the four centrality metrics (introduced in Section 2). Table 4 lists the *A. Index* values and the values are color coded as per the range outlined in Table 2. One can easily see that for about 80% of the real-world networks analyzed (i.e., for 40 of the 50 real-world networks analyzed), the level of assortativity is not the same for all the four centrality metrics. For a majority (i.e., 56%) of the real-world networks (i.e., for 28 of the 50 real-world networks), we observe two different levels of assortativity and most of these are the neutral and weakly assortative levels. For very few real-world networks, the two different levels of assortativity represent levels whose ranges of assortativity index values are not contiguous (for example: neutral and strongly assortative). For about 24% of the real-world networks analyzed (i.e., 12 of the 50 real-world networks), we observe three levels of assortativity. For none of the real-world networks, we observe four different levels of assortativity (i.e., one assortativity level per centrality metric). Only 6-14% of the real-world networks are either weakly or strongly disassortative with respect to any centrality metric.

We also plot (Figures 7-10) the distribution of the *A. Index* values for each of the four centrality metrics. We estimate the probability of observing a network to be at a particular level of assortativity with respect to a centrality metric as the fraction of the total number of real-world networks exhibiting the particular level of assortativity with respect to the centrality metric. These empirically estimated probability values are also listed in Figures 7-10. As a high-level conclusion, we could say that there is at least a 50% chance for a real-world network to be neutral (neither assortative nor disassortative) with respect to the degree centrality and betweenness centrality metrics. On the other hand, we observe that there is at least a 50% chance for a real-world network to be assortative (either strongly assortative or weakly assortative) with respect to the closeness centrality and eigenvector centrality metrics.

More specifically: we observe a real-world network to be neutral with respect to the BWC and DegC metrics with a probability of 0.72 and 0.58 respectively. When considered with respect to the EVC metric, we observe a real-world network to be weakly assortative with a probability of 0.38 and strongly assortative with a probability of 0.28. When considered with respect to the CIC metric, we observe a real-world network to be strongly assortative with a probability of 0.38 and weakly assortative with a probability of 0.28. Note that though both BWC and CIC are shortest path-based centrality metrics, we observe that they are poles apart with respect to assortativity. While a real-world network is more likely to be neutral (neither assortative nor disassortative) with respect to the BWC metric, we observe a real-world network to be more likely to be strongly assortative or weakly assortative with respect to the CIC metric. Table 5 summarizes these empirically estimated probability values for all the five levels of assortativity and all the four centrality metrics. Figure 11 presents a pictorial view of the empirically estimated probability values for observing a real-world network at a particular level of assortativity with respect to a centrality metric.

Table 4. Fundamental Properties of the Real-World Network Graphs used for Assortativity Analysis

#	Net.	BWC	DegC	EVC	CIC	#	Net.	BWC	DegC	EVC	CIC
1	ADJ	-0.10	-0.10	0.04	0.13	26	LMN	-0.02	-0.08	0.43	0.21
2	AKN	-0.08	-0.08	0.09	0.12	27	MDN	-0.10	-0.05	-0.02	-0.05
3	JBN	-0.04	0.03	0.35	0.18	28	MTB	0.25	0.73	0.79	1.00
4	CEN	-0.06	-0.09	0.22	0.22	29	MCE	0.05	0.72	0.79	0.94
5	CLN	-0.12	-0.11	0.06	0.06	30	MSJ	0.24	0.35	0.94	1.00
6	CGD	0.07	0.14	0.59	1.00	31	AFB	0.09	0.35	0.89	1.00
7	CFN	-0.07	-0.17	-0.09	-0.09	32	MPN	-0.12	-0.16	0.13	0.09
8	DON	0.12	-0.04	0.64	0.53	33	MMN	0.00	0.10	0.50	0.47
9	DRN	0.30	0.35	0.62	1.00	34	PBN	0.04	-0.02	0.54	0.37
10	DLN	-0.08	0.07	0.34	0.33	35	PSN	0.10	0.22	0.29	0.26
11	ERD	0.05	0.18	0.40	1.00	36	PFN	0.23	0.59	0.67	0.96
12	FMH	0.39	0.65	0.84	1.00	37	SJN	0.03	-0.14	0.51	0.41
13	FHT	0.25	0.60	0.69	1.00	38	SDI	0.22	0.08	0.95	1.00
14	FTC	-0.07	-0.03	0.45	0.23	39	SPR	-0.06	0.02	0.14	0.16
15	FON	0.06	0.19	0.69	0.31	40	SMN	-0.20	-0.23	-0.22	-0.23
16	CDF	-0.10	-0.11	-0.10	-0.12	41	SWC	-0.23	-0.17	-0.02	-0.02
17	GD96	-0.24	-0.32	-0.03	0.47	42	SSM	0.04	-0.02	0.50	0.33
18	MUN	0.04	0.14	0.64	1.00	43	TEN	-0.16	-0.36	0.26	0.23
19	GLN	-0.16	-0.13	0.30	1.00	44	TWF	0.55	0.84	0.93	1.00
20	GD01	-0.92	-0.98	-0.75	-0.54	45	UKF	-0.08	0.00	0.22	0.12
21	HTN	-0.10	-0.12	-0.10	-0.12	46	APN	-0.15	-0.21	-0.02	0.06
22	HCN	0.01	0.03	0.18	1.00	47	USS	0.23	0.23	0.62	0.65
23	ISP	0.14	0.29	0.56	0.77	48	RHF	0.00	0.10	0.38	0.25
24	KCN	-0.36	-0.48	-0.24	-0.08	49	WSB	0.02	0.45	0.50	0.89
25	KFP	0.17	0.24	0.53	0.73	50	WTN	-0.26	-0.39	-0.37	-0.35

An interesting and significant observation from the color-coded Table 3 is that for real-world networks with two or three levels of assortativity with the centrality metrics: the level of assortativity typically exhibited a transition from dissortative to neutral (or) neutral to weakly assortative to strongly assortative when the centrality metrics are considered in this order: BWC, DegC, EVC and CIC. We also notice from Figures 7-10 that the distribution of the *A. Index* values gradually drifts from a predominantly neutral-level distribution (corresponding to the BWC and DegC metrics) to a predominantly assortative-level distribution (corresponding to the EVC and CIC metrics). Such observations further vindicate our conclusions (in the previous paragraphs) regarding the probability of observing a real-world network to be neutral, weakly assortative and strongly assortative with respect to the centrality metrics.

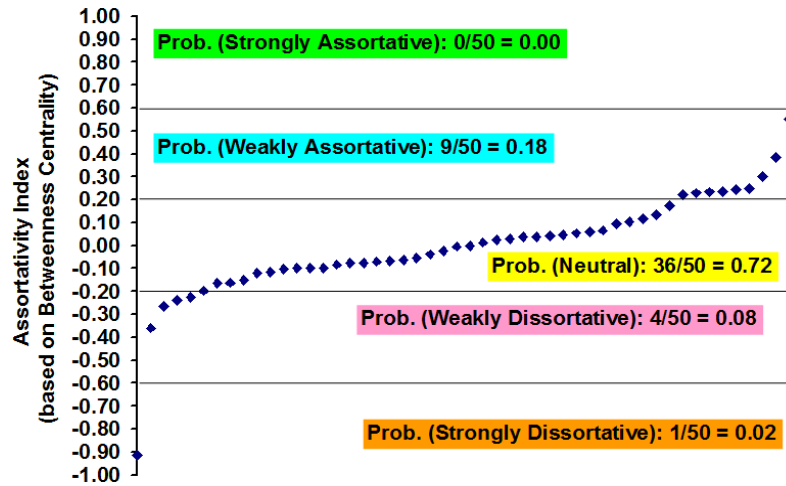


Figure 7. Distribution of Assortativity Index Values for Real-World Networks (based on Betweenness Centrality)

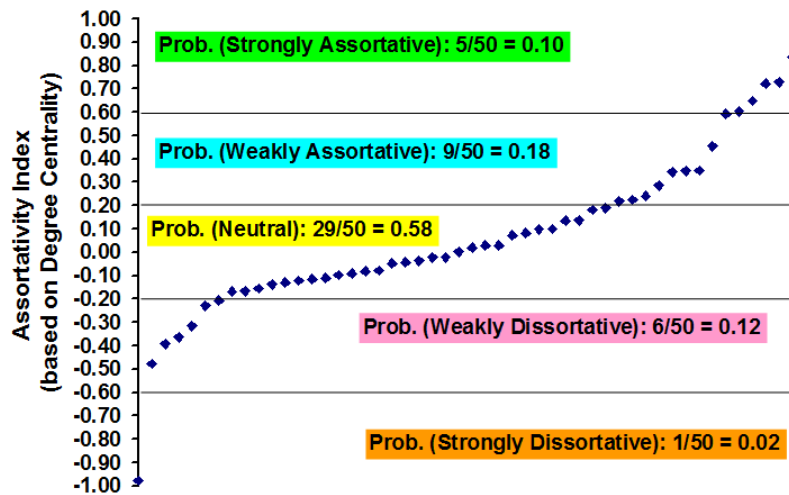


Figure 8. Distribution of Assortativity Index Values for Real-World Networks (based on Degree Centrality)

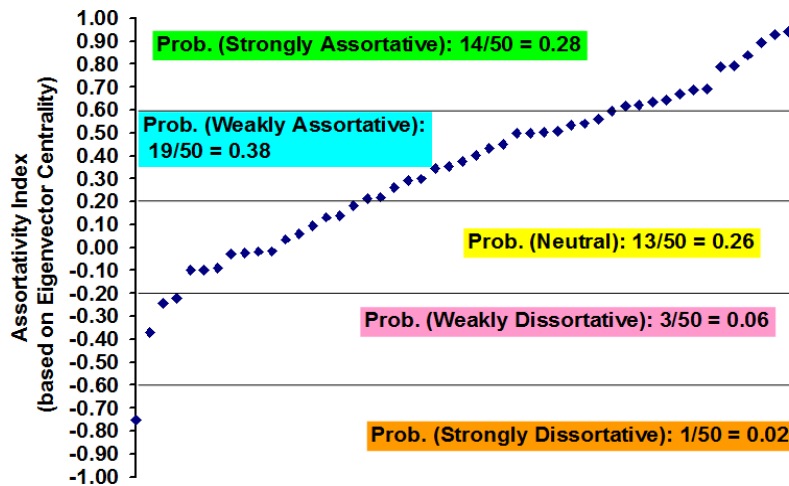


Figure 9. Distribution of Assortativity Index Values for Real-World Networks (based on Eigenvector Centrality)

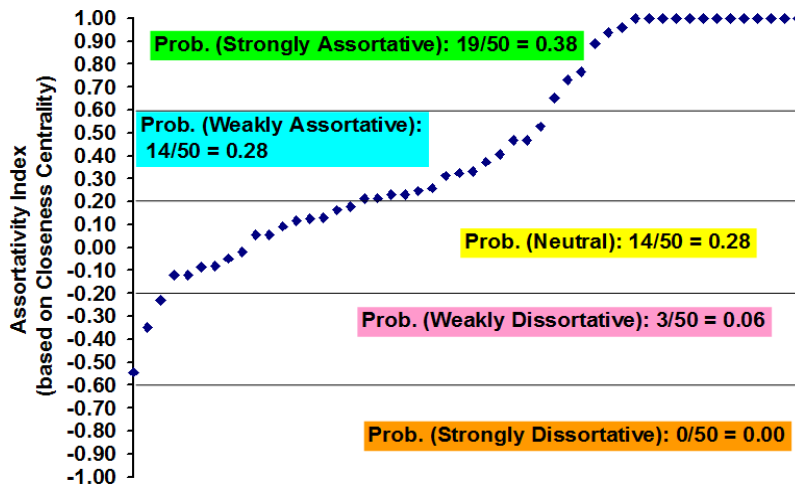


Figure 10. Distribution of Assortativity Index Values for Real-World Networks (based on Closeness Centrality)

Table 5. Empirically Estimated Probability Values for the Assortative Level of a Real-World Network with respect to the Centrality Metrics

Centrality Metric	<i>Strongly Dissortative</i>	<i>Weakly Dissortative</i>	<i>Neutral</i>	<i>Weakly Assortative</i>	<i>Strongly Assortative</i>
Betweenness	0.02	0.08	0.72	0.18	0.00
Degree	0.02	0.12	0.58	0.18	0.10
Eigenvector	0.02	0.06	0.26	0.38	0.28
Closeness	0.00	0.06	0.28	0.28	0.38

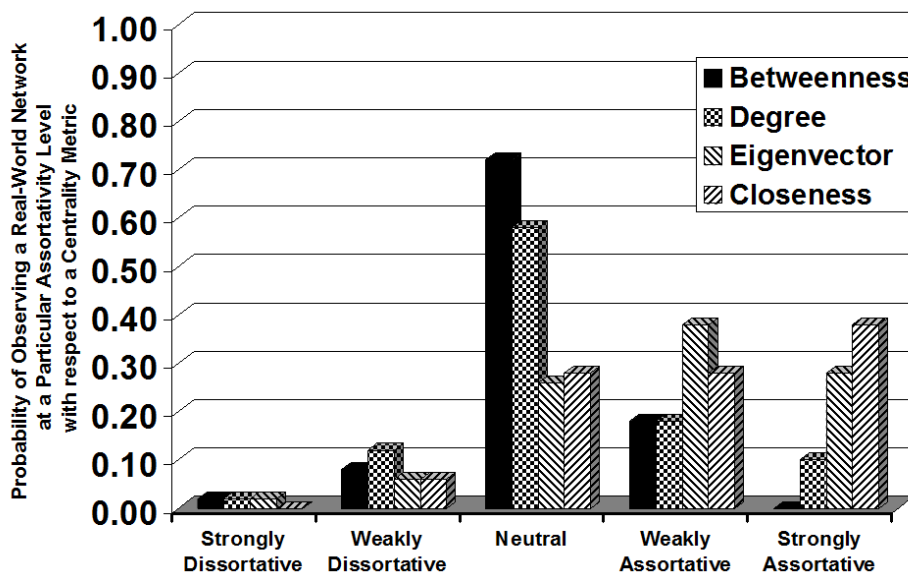


Figure 11. Empirically Estimated Probability for Observing a Real-World Network at a Particular Level of Assortativity with respect to a Centrality Metric

6. Related Work

To the best of our knowledge, all the results reported in the literature (e.g., Newman, 1999; Newman & Girvan, 2003; Noldus & Van Mieghem, 2015) on assortativity of real-world network graphs is based on the degree centrality metric. Ours is the first effort to study the assortativity of real-world network graphs based on the other

commonly used centrality metrics such as betweenness centrality, eigenvector centrality and closeness centrality. We analyze the assortativity of a large collection of real-world network graphs (with a broad range of variation in node degree) and empirically propose the likelihood of observing a real-world network graph to be neutral or assortative with respect to a centrality metric. In this section, we discuss results from the most related work on assortativity and centrality metrics in the literature.

Traditionally, based on the degree centrality metric, social networks have been found to be assortative (high degree nodes tend to attach to high degree nodes); whereas, the technological and biological networks have been observed to be disassortative (i.e., low degree nodes tend to attach to high degree nodes and vice-versa; Newman, 2003). The networks generated from theoretical models such as the Erdos-Renyi random networks (Erdos & Renyi, 1959), Barabasi-Albert scale-free networks (Barabasi & Albert, 1999) and the Watts-Strogatz small-world networks (Watts & Strogatz, 1998) have also been observed to be neutral (neither assortative nor disassortative) with respect to the degree centrality metric (Newman, 1999). In addition, networks that evolve with time without any constraints have been observed to reach a maximum entropy state (entropy is a quantitative measure of robustness; Demetrius & Manke, 2005) with heterogeneous connectivity distribution and in such a state, networks have been usually disassortative (Johnson et al., 2010) with respect to the degree centrality metric. On the other hand, networks that evolved with constraints (with respect to the number of links a node can maintain) tend to transition from being disassortative to assortative with time (Konig et al., 2010). Also, synthetic social network graphs generated using the Monte Carlo Metropolis-Hastings type algorithms (Chib & Greenberg, 1995) were observed to quickly evolve to a giant component if edge distribution (based on *remaining degree*: one less than the degree centrality; Newman, 1999) follows assortative matching (rather than disassortative matching; Newman, 2003).

Iyer et al. (2013) analyzed the robustness of networks due to targeted removal of vertices that are ranked higher with respect to centrality metrics. It has been observed that disassortative networks degrade more rapidly due to the removal of vertices with higher degree; whereas, assortative networks degrade more rapidly due to the removal of vertices with higher betweenness (at least for the first 25% of the vertices) as the high degree vertices in assortative networks tend to form a concentrated interconnected core that would be difficult to break due to the removal of few vertices. For neutral networks (with assortativity index close to 0), targeted node removal based on degree has been observed to be the most effective method to degrade the network and targeted removal based on eigenvector centrality has been observed to be the least effective (Iyer et al., 2013). The findings from this paper could be considered complementary to the above research results as we observe real-world network graphs to be more likely assortative with respect to the EVC metric; hence, removal of vertices with higher EVC is more likely to have a relatively less degrading effect on the assortativity of networks.

Zhang et al. (2012) argued that assortativity level of the different communities with their neighborhood need not be the same as the assortativity level of the entire network. This could be attributed to the differences in the connectivity of the vertices in the various communities to their respective outside world. In this regard, Zhang et al. (2012) proposed an alternate metric called the Universal Assortativity Coefficient (UAC) defined for a community (sub graph) of vertices as the sum of the local assortativity indices of edges (Newman, 1999) emanating from the vertices that are part of the community. The local assortativity index of an edge is calculated as per the *remaining degree* based formulation proposed by Newman (1999): Edges with positive local assortativity index are referred to as assortative and edges with negative local assortativity index are referred to as disassortative. Zhang et al. (2012) claimed that a globally assortative network could still have majority of its edges to be locally disassortative and vice-versa. Similar to local edge assortativity, a measure called local node assortativity (Piraveenan et al., 2008) based on the remaining degree of a node has also been proposed in the literature; the sum of local node assortativity values is equal to the network assortativity. It has been shown by Piraveenan et al (2009) that distribution profiles of the local assortativity of nodes vs. their degrees could be used to identify assortative hubs in social and biological networks and disassortative hubs in scale-free networks such as the Internet. All of the above analyses has been based on only the degree centrality metric and is heavily based on the concept of remaining degree.

Joyce et al. (2010) proposed the notion of *leverage centrality* to capture the assortative or disassortative neighborhood of a node. The range of values for leverage centrality is $(-1, \dots, 1)$: positive values indicating an assortative neighborhood and negative values indicating a disassortative neighborhood. A node has a positive leverage centrality if it is connected to more nodes than its neighbors (assortative neighborhood); a node connected to fewer nodes than its neighbors has a negative leverage centrality (disassortative neighborhood). Nodes having higher leverage centrality are perceived to be important for facilitating information flow to and from its neighbors. Leverage centrality of a node is estimated simply based on the degree of the node and that of

its neighbors; the centrality metric has been observed to be strongly correlated with betweenness centrality and weakly correlated with eigenvector centrality metric. Per the trends observed in this paper, we expect a real-world network to be more likely to be neutral (neither assortative nor disassortative) with respect to the leverage centrality metric (as is also observed for the betweenness centrality metric).

Meghanathan (2015) observed the degree centrality and betweenness centrality metrics to be highly correlated for real-world network graphs: we opine such a correlation is justified with the real-world network graphs exhibiting almost similar levels of assortativity with respect to both these centrality metrics in this paper. Meghanathan (2016) showed that a maximal assortative matching of vertices (with the objective of maximizing the assortativity index) in real-world network graphs (with respect to the degree centrality metric) cannot maximize the number of matched vertices and vice-versa. We attribute such a phenomenon to the relatively neutral levels of assortativity of the edges with respect to the degree centrality metric and its closely correlated betweenness centrality metric. As we observe the closeness centrality metric to exhibit stronger levels of assortativity, we opine that a maximal assortative matching of vertices based on closeness centrality would relatively increase the number of vertices matched (compared to the degree centrality metric).

7. Conclusions and Future Work

We have shown that the assortativity classification of real-world network graphs is dependent on the node-level centrality metric used to compute the assortativity index values of the edges. As part of this analysis, we formally propose five levels of assortativity and their associated ranges in the space of assortativity index values from -1 to 1. We computed the assortativity index values for a suite of 50 real-world network graphs (with spectral radius ratio for node degree ranging from 1.01 to 3.48) with respect to each of the four commonly used centrality metrics: degree centrality (DegC), eigenvector centrality (EVC), betweenness centrality (BWC) and closeness centrality (CIC). We observe about 80% of the real-world network graphs to exhibit more than one assortativity level (depending on the centrality metric used to compute the assortativity index values): 56% exhibiting two assortativity levels and 24% exhibiting three assortativity levels. We notice for a majority of these real-world network graphs, the level of assortativity exhibited a transition from disassortative to neutral (or) neutral to weakly assortative to strongly assortative when the centrality metrics are considered in this order: BWC, DegC, EVC and CIC. Using the results of the assortativity analysis, we also estimated the empirical probability for a real-world network graph to exhibit a particular level of assortativity: We claim that a real-network graph is more likely (probability of 0.72) to be neutral (neither assortative nor disassortative) with respect to the BWC metric and is more likely to be assortative (strongly or weakly assortative: probability of $0.38 + 0.28 = 0.66$) with respect to the EVC and CIC metrics.

We have thus unraveled significant information about the assortativity of real-world network graphs with respect to the other commonly used centrality metrics such as betweenness, closeness and eigenvector centrality. As part of future work, we plan to analyze the centrality-based assortativity of complex network graphs generated from theoretical models (such as the Erdos-Renyi random network model (Erdos & Renyi, 1959), Barabasi-Albert scale-free network model (Barabasi & Albert, 1999) and the Watts-Strogatz small-world network model; Watts & Strogatz, 1998). We also plan to investigate the use of centrality metrics (other than degree centrality) to compute maximal assortative matching and maximal disassortative matching (Meghanathan, 2016) for real-world network graphs.

Acknowledgments

The research is financed by the NASA EPSCoR sub award (#: NNX14AN38A) from University of Mississippi.

References

- Barabasi, A. L., & Albert, R. (1999). Emergence of Scaling in Random Networks. *Science*, 286(5439), 509-512. <http://dx.doi.org/10.1126/science.286.5439.509>
- Batagelj, V., & Mrvar, A. (2006). *Pajek Datasets*. Retrieved from <http://vlado.fmf.uni-lj.si/pub/networks/data/>
- 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. [http://dx.doi.org/10.1016/0378-8733\(79\)90014-5](http://dx.doi.org/10.1016/0378-8733(79)90014-5)
- Biedl, T., & Franz, B. J. (2001). *Graph-Drawing Contest Report*. Paper presented at the 9th International Symposium on Graph Drawing, Vienna, Austria. http://dx.doi.org/10.1007/3-540-45848-4_60
- Bonacich, P. (1987). Power and Centrality: A Family of Measures. *American Journal of Sociology*, 92(5), 1170-1182. <http://dx.doi.org/10.1086/228631>

- Brandes, U. (2001). A Faster Algorithm for Betweenness Centrality. *The Journal of Mathematical Sociology*, 25(2), 163-177. <http://dx.doi.org/10.1080/0022250X.2001.9990249>
- Chib, S., & Greenberg, E. (1995). Understanding the Metropolis-Hastings Algorithm. *American Statistician*, 49(4), 327-335. <http://dx.doi.org/10.1080/00031305.1995.10476177>
- Cormen, T. H., Leiserson, C. E., Rivest, R. L., & Stein, C. (2009). *Introduction to Algorithms*. (3rd Ed.) MIT Press.
- Cross, R. L., Parker, A., & Cross, R. (2004). *The Hidden Power of Social Networks: Understanding How Work Really Gets Done in Organizations*. (1st Ed.) Harvard Business Review Press.
- Demetrius, L., & Manke, T. (2005). Robustness and Network Evolution - An Entropic Principle. *Physica A*, 346, 682-696. <http://dx.doi.org/10.1098%2Frsif.2006.0140>
- Erdos, P., & Renyi, A. (1959). On Random Graphs I. *Publicationes Mathematicae*, 6, 290-297.
- Evans, J. D. (1995). *Straightforward Statistics for the Behavioral Sciences*. (1st Ed.) Brooks Cole Publishing Company.
- Freeman, L. (1977). A Set of Measures of Centrality based on Betweenness. *Sociometry*, 40(1), 35-41. <http://dx.doi.org/10.2307/3033543>
- Freeman, L. (1979). Centrality in Social Networks Conceptual Clarification. *Social Networks*, 1(3), 215-239. [http://dx.doi.org/10.1016/0378-8733\(78\)90021-7](http://dx.doi.org/10.1016/0378-8733(78)90021-7)
- Freeman, L., Webster, C. M., & Kirke, D. M. (1998). Exploring Social Structure using Dynamic Three-Dimensional Color Images. *Social Networks*, 20(2), 109-118. [http://dx.doi.org/10.1016/S0378-8733\(97\)00016-6](http://dx.doi.org/10.1016/S0378-8733(97)00016-6)
- Freeman, L., Freeman, S. C., & Michaelson, A. G. (1989). How Humans See Social Groups: A Test of the Sailer-Gaulin Models. *Journal of Quantitative Anthropology*, 1, 229-238.
- Geiser, P., & Danon, L. (2003). Community Structure in Jazz. *Advances in Complex Systems*, 6(4), 563-573. <http://dx.doi.org/10.1142/S0219525903001067>
- Gemmetto, V., Barrat, A., & Cattuto, C. (2014). Mitigation of Infectious Disease at School: Targeted Class Closure vs. School Closure. *BMC Infectious Diseases*, 14(695), 1-10. <http://dx.doi.org/10.1186/s12879-014-0695-9>
- Gil-Mendieta, J., & Schmidt, S. (1996). The Political Network in Mexico. *Social Networks*, 18(4), 355-381. [http://dx.doi.org/10.1016/0378-8733\(95\)00281-2](http://dx.doi.org/10.1016/0378-8733(95)00281-2)
- Girvan, M., & Newman, M. (2002). Community Structure in Social and Biological Networks. *Proceedings of the National Academy of Sciences USA*, 99(12), 7821-7826. <http://dx.doi.org/10.1073/pnas.122653799>
- Gleiser, P. M. (2007). How to become a Superhero. *Journal of Statistical Mechanics: Theory and Experiments*, P09020. <http://dx.doi.org/10.1088/1742-5468/2007/09/P09020>
- Grimmer, J. (2010). A Bayesian Hierarchical Topic Model for Political Texts: Measuring Expressed Agendas in Senate Press Releases. *Political Analysis*, 18(1), 1-35. <http://dx.doi.org/10.1093/pan/mpp034>
- Hayes, B. (2006). Connecting the Dots. *American Scientist*, 94(5), 400-404. <http://dx.doi.org/10.1511/2006.61.3495>
- Hummon, N. P., Doreian, P., & Freeman, L. C. (1990). Analyzing the Structure of the Centrality-Productivity Literature created between 1948 and 1979. *Science Communication*, 11(4), 459-480. <http://dx.doi.org/10.1177/107554709001100405>
- 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. *Journal of Theoretical Biology*, 271(1), 161-180. <http://dx.doi.org/10.1016/j.jtbi.2010.11.033>
- Iyer, S., Killingback, T., Sundaram, B., & Wang, Z. (2013). Attack Robustness and Centrality of Complex Networks. *PLoS One*, 8(4), e59613. <http://dx.doi.org/10.1371/journal.pone.0059613>
- Johnson, S., Torres, J. J., Marro, J., & Munoz, M. A. (2010). Entropic Origin of Dissortativity in Complex Networks. *Physical Review Letters*, 104, 108702.
- Joyce, K. E., Laurienti, P. J., Burdette, J. H., & Hayasaka, S. (2010). A New Measure of Centrality for Brain Networks. *PLoS One*, 5(8), e12200. <http://dx.doi.org/10.1371/journal.pone.0012200>

- Knuth, D. E. (1993). *The Stanford GraphBase: A Platform for Combinatorial Computing*. (1st Ed.) Addison-Wesley.
- Konig, M. D., Tessone, C. J., & Zenou, Y. (2010). From Assortative to Dissortative Networks: The Role of Capacity Constraints. *Advances in Complex Systems*, 13(4), 483-500. <http://dx.doi.org/10.1142/S0219525910002700>
- Krackhardt, D. (1999). The ties that torture: Simmelian tie analysis in organizations. *Research in the Sociology of Organizations*, 16(1), 183-210.
- Krebs, V. (2003). Proxy Networks: Analyzing One Network to Reveal Another. *Bulletin de Méthodologie Sociologique*, 79, 61-70.
- Lee, J. S. (2004). Generating Networks of Illegal Drug Users using Large Samples of Partial Ego-Network Data. *Intelligence and Security Informatics, Lecture Notes in Computer Science*, 3073, 390-402. http://dx.doi.org/10.1007/978-3-540-25952-7_29
- Loomis, C. P., Morales, J. O., Clifford, R. A., & Leonard, O. E. (1953). *Turrialba Social Systems and the Introduction of Change*. (1st Ed.) The Free Press.
- Lusseau, D., Schneider, K., Boisseau, O. J., Hasse, P., Slooten, E., & Dawson, S. M. (2003). The Bottleneck Dolphin Community of Doubtful Sound Features a Large Proportion of Long-lasting Associations. *Behavioral Ecology and Sociobiology*, 54(3), 396-405. <http://dx.doi.org/10.1007/s00265-003-0651-y>
- MacRae, D. (1960). Direct Factor Analysis of Sociometric Data. *Sociometry*, 23(4), 360-371.
- McCarty, C., & Freeman, L. (2008). *Network Datasets*. Retrieved from <http://moreno.ss.uci.edu/data.html>
- Meghanathan, N. (2014). *Spectral Radius as a Measure of Variation in Node Degree for Complex Network Graphs*. Paper presented at the 3rd International Conference on Digital Contents and Applications, Hainan, China. <http://dx.doi.org/10.1109/UNESST.2014.8>
- Meghanathan, N. (2015). Correlation Coefficient Analysis of Centrality Metrics for Complex Network Graphs. Paper presented at the 4th Computer Science Online Conference. http://dx.doi.org/10.1007/978-3-319-18503-3_2
- Meghanathan, N. (2016). Maximal Assortative Matching for Complex Network Graphs. *Journal of King Saud University – Computer and Information Sciences*, 28(2), 230-246. <http://dx.doi.org/10.1016/j.jksuci.2015.10.004>
- Michael, J. H. (1997). Labor Dispute Reconciliation in a Forest Products Manufacturing Facility. *Forest Products Journal*, 47(11-12), 41-45.
- Moreno, J. L. (1960). *The Sociometry Reader*. (1st Ed.) The Free Press.
- Nepusz, T., Petroczi, A., Negyessy, L., & Bazso, F. (2008). Fuzzy Communities and the Concept of Bridgeness in Complex Networks. *Physical Review E*, 77(1), 016107.
- Newman, M. (2010). *Networks: An Introduction*. (1st Ed.) Oxford University Press.
- Newman, M. (1999). Assortative Mixing in Networks. *Physical Review Letters*, 89(20), 208701. <http://dx.doi.org/10.1103/PhysRevLett.89.208701>
- Newman, M. (2006). Finding Community Structure in Networks using the Eigenvectors of Matrices. *Physical Review E*, 74(3), 036104. <http://dx.doi.org/10.1103/PhysRevE.74.036104>
- Newman, M. (2003). Mixing Patterns in Networks. *Physical Review E*, 67(1), 026126, <http://dx.doi.org/10.1103/PhysRevE.67.026126>
- Newman, M., & Girvan, M. (2003). Mixing Patterns and Community Structure in Networks. *Statistical Mechanics of Complex Networks: Lecture Notes in Physics*, 625, 66-87. http://dx.doi.org/10.1007/978-3-540-44943-0_5
- Noldus, R., & Van Mieghem, P. (2015). Assortativity in Complex Networks. *Journal of Complex Networks*, 3(4), 507-542. <http://dx.doi.org/10.1093/comnet/cnv005>
- Nooy, W. (1999). A Literary Playground: Literary Criticism and Balance Theory. *Poetics*, 26(5-6), 385-404. [http://dx.doi.org/10.1016/S0304-422X\(99\)00009-1](http://dx.doi.org/10.1016/S0304-422X(99)00009-1)
- Pearson, M., & Michell, L. (2000). Smoke Rings: Social Network Analysis of Friendship Groups, Smoking and Drug-taking. *Drugs: Education, Prevention and Policy*, 7(1), 21-37. <http://dx.doi.org/>

10.1080/dep.7.1.21.37

- Piraveenan, M., Prokopenko, M., & Zomaya, A. Y. (2008). Local Assortativeness in Scale-Free Networks. *Europhysics Letters*, 84(2), 28002. <http://dx.doi.org/10.1209/0295-5075/84/28002>
- Piraveenan, M., Prokopenko, M., & Zomaya, A. Y. (2009). Local Assortativity and Growth of Internet. *Interdisciplinary Physics: The European Physical Journal B*, 70(2), 275-285. <http://dx.doi.org/10.1140/epjb/e2009-00219-y>
- 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. *Journal of the American Medical Association*, 278(10), 823-832. <http://dx.doi.org/10.1001/jama.1997.03550100049038>
- Rogers, E. M., & Kincaid, D. L. (1980). *Communication Networks: Toward a New Paradigm for Research*. (1st Ed.) The Free Press.
- Schwimmer, E. (1973). *Exchange in the Social Structure of the Orokaiva: Traditional and Emergent Ideologies in the Northern District of Papua*. (1st Ed.) C Hurst and Co-Publishers Ltd.
- Scott, J. P. (1980). *The Anatomy of Scottish Capital: Scottish Companies and Scottish Capital*. (1st ed.) Croom Helm.
- Smith, D. A., & White, D. R. (1992). Structure and Dynamics of the Global Economy: Network Analysis of International Trade 1965-1980. *Social Forces*, 70(4), 857-893. <http://dx.doi.org/10.1093/sf/70.4.857>
- Strang, G. (2006). *Linear Algebra and its Applications*. (4th Ed.) Brooks Cole.
- Takahata, Y. (1991). *Diachronic Changes in the Dominance Relations of Adult Female Japanese Monkeys of the Arashiyama B Group*. (1st Ed.) Albany: State University of New York Press.
- Watts, D. J., & Strogatz, S. H. (1998). Collective Dynamics of Small-World Networks. *Nature*, 393, 440-442. 509-512.
- White, J. G., Southgate, E., Thomson, J. N., & Brenner, S. (1986). The Structure of the Nervous System of the Nematode *Caenorhabditis Elegans*. *Philosophical Transactions B*, 314(1165), 1-340. <http://dx.doi.org/10.1098/rstb.1986.0056>
- Zachary, W. W. (1977). An Information Flow Model for Conflict and Fission in Small Groups. *Journal of Anthropological Research*, 33(4), 452-473.
- Zhang, G. Q., Cheng, S. Q., Zhang, G. Q. (2012). A Universal Assortativity Measure for Network Analysis. *arXiv: 1212.6456* [physics.soc-ph].

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).