

# Randomness Index for complex network analysis

Natarajan Meghanathan<sup>1</sup>

Received: 3 April 2017 / Revised: 16 May 2017 / Accepted: 5 June 2017 / Published online: 12 June 2017  
© Springer-Verlag GmbH Austria 2017

**Abstract** The high-level contribution of this paper is a quantitative measure (called Randomness Index) to assess the extent of randomness in the topology of a complex real-world network. We exploit the observation that the local clustering coefficient (LCC) for a node in a truly random network is independent of the degree of the node and is simply the probability for a link to exist between any two nodes in the network. On the other hand, for real-world networks that are not truly random, nodes with a larger degree are more likely to have a lower LCC value and vice versa. For any complex real-world network, we propose to determine the Randomness Index as the Pearson's correlation coefficient (ranging from  $-1$  to  $1$ ) of the degree versus average LCC of the nodes with the particular degree. We evaluate the Randomness Index values for a suite of 48 real-world networks of diverse degree distribution and observe the median value to be  $-0.72$ .

**Keywords** Randomness Index · Complex networks · Degree centrality · Local clustering coefficient · Pearson's correlation coefficient

## 1 Introduction

Complex network analysis is an emerging area of research interest, and it corresponds to analyzing the topology of complex networks (such as biological networks, citation networks, social networks and web) from a graph-theoretic perspective (Newman 2010). Random network models, like

the well-known Erdos–Renyi  $G(N, p_{\text{link}})$  model (Erdos and Renyi 1959): where  $p_{\text{link}}$  is the probability for a link between any two nodes, are used to theoretically model complex real-world networks as truly random networks and form the basis to assess the extent of randomness among the links of the real-world networks. Until now, the assessment has been through a collection of characteristics such as degree distribution, path length and connectedness (Barabasi 2016), but none of these can be unequivocally used to quantify the extent of randomness in any complex real-world network.

In this paper, we investigate one characteristic feature of truly random versus real-world networks (that are not truly random) that has not been explored to its fullest potential. The feature that we seek to explore is the distribution of the degree ( $K$ ) versus average local clustering coefficient (LCC) of the nodes with a certain degree (hereafter, referred to as the  $K$  vs.  $\overline{\text{LCC}}(K)$  distribution). The local clustering coefficient (LCC) for a node (Newman 2010) is the probability that any two neighbors of the node are connected. The LCC (ranges from 0 to 1) for a node is measured as the ratio of the actual number of links between the neighbors of the node to that of the maximum possible number of links between the neighbors of the node. For truly random networks following the  $G(N, p_{\text{link}})$  model, the probability for a link between any two nodes is the same, and hence, the probability that any two neighbors of a node are connected is simply the  $p_{\text{link}}$  value. In other words, the LCC for a node in a truly random network does not depend on the degree of the node. On the other hand, for a real-world network, a node having a smaller degree has a larger chance of having any two of its neighbor nodes to be connected (compared to a node having a larger degree). That is, for a real-world network, a node having a larger

✉ Natarajan Meghanathan  
natarajan.meghanathan@jsums.edu

<sup>1</sup> Jackson State University, Jackson, MS 39217, USA

degree is more likely to have a lower LCC compared to a node having a smaller degree. Hence, for real-world networks, if we average the LCC of the nodes having a certain degree and plot the  $K$  versus  $\overline{LCC}(K)$  distribution, the distribution is more likely to exhibit a strong negative correlation [as already reported in several related works (Barabasi 2016; Foudalis et al. 2011; Pizzuti and Rombo 2014)], whereas for a truly random network generated based on the average degree and number of nodes for a real-world network graph, the  $K$  versus  $\overline{LCC}(K)$  distribution is more likely to be a flat distribution (i.e., no dependence) or sometimes exhibit a positive correlation.

The current status of the literature is to visually inspect the  $K$  versus  $\overline{LCC}(K)$  distribution and infer whether a real-world network is truly random or not. We seek to proceed one step further. We propose to quantify the extent of randomness for a complex real-world network in the form of a measure called the *Randomness Index*, and it is the Pearson's correlation coefficient (ranges from  $-1$  to  $1$ ) of the  $K$  vs.  $\overline{LCC}(K)$  values for the vertices of the network. For any complex real-world network, the closer the Randomness Index value to zero (or even positive), the larger the extent of randomness in the distribution of the links in the network, whereas the closer the Randomness Index value to  $-1$ , the lower the extent of randomness. We analyze a suite of 48 real-world networks of diverse degree distribution and observe the median of the Randomness Index values to be  $-0.72$ . Only two of the 48 real-world networks had a positive Randomness Index value, and the rest of the networks had Randomness Index values less than  $-0.15$ .

To the best of our knowledge, we have not come across a single and dimensionless quantitative measure to unequivocally quantify (that is, its range of values  $-1 \dots 1$  remains the same, irrespective of the number of nodes and edges in the network) the extent of randomness of complex real-world networks. The Randomness Index value for a real-world network could form the basis for further exploratory studies conducted on the network. The links in a real-world network with a larger Randomness Index value are more likely to have formed due to random association, and such a network could thence be subjected to further tests that are characteristic for random networks; on the other hand, the links in a real-world network with a lower Randomness Index value are more likely to have not formed due to random association (instead, it could be due to preferential attachment), and such a network could thence be subjected to further tests that are more characteristic for non-random networks (such as scale-free networks). From a structural perspective, the Randomness Index value for a network could be considered as a quantified measure of the significance of the high-degree nodes

in the network. For networks with lower Randomness Index values, the high-degree nodes are more likely to have a lower local clustering coefficient; hence, in such networks, if the high-degree nodes are removed, several structural measures such as neighborhood connectivity, network robustness, path length (for shortest path communication) and feedback capacity for diffusion (Al-Shiridah et al. 2013) are likely to be affected. On the other hand, for networks with larger Randomness Index values, the high-degree nodes are relatively less critical as the probability for any two nodes to be connected does not depend on the degree of the nodes or the degree of their neighbors.

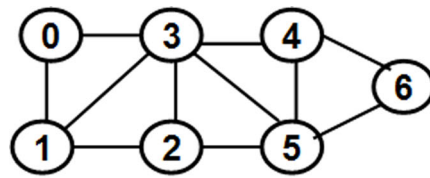
The rest of the paper is organized as follows: In Sect. 2, we present an illustrative example to generate a truly random network for a given real-world network (that is not random), determine the  $K$  versus  $\overline{LCC}(K)$  distribution for these two networks and compute the Randomness Index values for the two network graphs. In Sect. 3, we first present an overview of the 48 real-world networks analyzed in this research and tabulate/discuss the Randomness Index values observed for these networks. We then demonstrate the uniqueness of our proposed Randomness Index metric by comparing the distribution of the values for this metric with those obtained for some of the well-known modeling parameters for complex network analysis as well as with those available to quantify the variation or correlation of node degree. In Sect. 4, we review related work and highlight the unique contributions of the paper. Section 5 concludes the paper. Throughout the paper, the terms “node” and “vertex,” “link” and “edge,” “network” and “graph” are used interchangeably. They mean the same.

## 2 Randomness Index: illustration

The Randomness Index of a complex network is the Pearson's correlation coefficient (Strang 2006) of the degree ( $K$ ) versus the average local clustering coefficient of the vertices with a certain degree,  $\overline{LCC}(K)$ . In this section, we illustrate the computation of the Randomness Index of a real-world network and a truly random network that is generated according to the well-known Erdos–Renyi model (Erdos and Renyi 1959) on the basis of the average degree ( $\langle K \rangle$ ) and number of nodes ( $N$ ) for the real-world network (see Fig. 1 for an illustration of the procedure).

For a given real-world network (the top graph in Fig. 1), we first compute the average degree of the nodes in the network. We then determine the probability for a link ( $p_{\text{link}} = \langle K \rangle / N - 1$ ) between any two nodes if we were to generate a truly random network according to the Erdos–

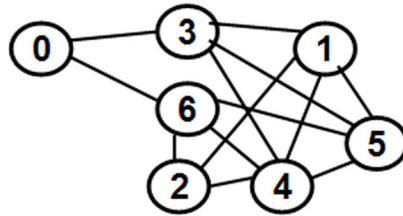
**Fig. 1** Generation of a Truly Random Network according to the Erdos–Renyi Model based on the Average Degree and Number of Nodes for a Real-World Network



**Real-World Network**

Avg. Degree  $\langle K \rangle$   
 $= (2 + 3 + 3 + 5 + 3 + 4 + 2) / 7$   
 $= 3.14$

ER Model  $p_{link} = \langle K \rangle / (N-1)$   
 $p_{link} = 3.14 / 6 = 0.524$



**Truly Random Network**

**Erdos-Renyi Model in Execution**

Index	Pairs	Random Val	Edge
1	0, 1	0.6335	N
2	0, 2	0.7478	N
3	0, 3	0.1721	Y
4	0, 4	0.9234	N
5	0, 5	0.8563	N
6	0, 6	0.3141	Y
7	1, 2	0.1594	Y
8	1, 3	0.2945	Y
9	1, 4	0.2227	Y
10	1, 5	0.0343	Y
11	1, 6	0.7621	N
12	2, 3	0.8595	N
13	2, 4	0.3091	Y
14	2, 5	0.5312	N
15	2, 6	0.1834	Y
16	3, 4	0.4194	Y
17	3, 5	0.2549	Y
18	3, 6	0.6974	N
19	4, 5	0.0968	Y
20	4, 6	0.4486	Y
21	5, 6	0.2983	Y

Renyi model based on  $\langle K \rangle$  and  $N$  for the real-world network. We then consider all possible pairs of vertices (with vertex IDs ranging from 0 to  $N - 1$ ). We generate a random number in the range of  $(0, \dots, 1)$  for each pair of vertices; if the random number generated for a pair is less than or equal to the  $p_{link}$  value, then there is an edge between the vertices forming the pair as part of the truly random network (the bottom graph in Fig. 1); otherwise, not.

Figure 2 presents the calculation of the degree ( $K$ ) versus average LCC for a particular degree ( $\overline{LCC}(K)$ ) distribution for both the real-world network graph and the truly random network graph. For the real-world network

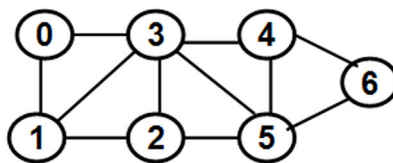
graph, we observe the  $\overline{LCC}(K)$  values to decrease with increase in  $K$ ; on the other hand, for the truly random network, the  $\overline{LCC}(K)$  values remain almost the same for the three  $K$  values of 3, 4 and 5.

Randomness Index

$$= \frac{\sum_K [K - \langle K \rangle] [\overline{LCC}(K) - \langle \overline{LCC}(K) \rangle]}{\sqrt{\sum_K [K - \langle K \rangle]^2} \sqrt{\sum_K [\overline{LCC}(K) - \langle \overline{LCC}(K) \rangle]^2}} \tag{1}$$

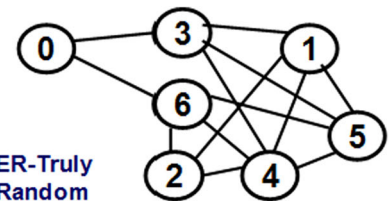
Figures 3 and 4 present the calculation of the Randomness Index (see formulation 1) as the Pearson’s correlation coefficient of the  $K$  versus  $\overline{LCC}(K)$  values for the

**Fig. 2** Computation of the  $K$  versus  $\overline{LCC}(K)$  Values for the Real-World Network and Truly Random Network Graphs of Fig. 1



**Real-World Network**

Degree (K)	Degree Nodes	LCC	$\overline{LCC}(K)$
2	0, 6	1, 1	1.0
3	1, 2, 4	2/3, 2/3, 2/3	0.67
4	5	3/6	0.5
5	3	4/10	0.4



**ER-Truly Random Network**

Degree (K)	Degree Nodes	LCC	$\overline{LCC}(K)$
2	0	0.0	0.0
3	2	2/3	0.67
4	1, 3, 5, 6	4/6, 3/6	0.542
5	4	4/6, 2/6	0.6

**Fig. 3** Computation of the Randomness Index for the  $K$  versus  $\overline{LCC}(K)$  Values for the Real-World Network Graph of Fig. 1

$K$	$\overline{LCC}(K)$	$K - \langle K \rangle$	$\frac{\overline{LCC}(K)}{-\langle \overline{LCC}(K) \rangle}$	$[K - \langle K \rangle]^2$	$\left[ \frac{\overline{LCC}(K)}{-\langle \overline{LCC}(K) \rangle} \right]^2$	$\left[ \frac{K - \langle K \rangle}{\left[ \frac{\overline{LCC}(K)}{-\langle \overline{LCC}(K) \rangle} \right]} \right]^2$
2	1.00	-1.5	0.36	2.25	0.1296	-0.540
3	0.67	-0.5	0.03	0.25	0.0009	-0.015
4	0.50	0.5	-0.14	0.25	0.0196	-0.070
5	0.40	1.5	-0.24	2.25	0.0576	-0.360
3.5	0.64			Sum 5.00	0.2077	-0.985
$\langle K \rangle$	$\langle \overline{LCC}(K) \rangle$					

$$\text{Randomness Index} = \frac{-0.985}{\sqrt{5.00} * \sqrt{0.2077}} = -0.9665$$

**Fig. 4** Computation of the Randomness Index for the  $K$  versus  $\overline{LCC}(K)$  Values for the Truly Random Graph of Fig. 1

$K$	$\overline{LCC}(K)$	$K - \langle K \rangle$	$\frac{\overline{LCC}(K)}{-\langle \overline{LCC}(K) \rangle}$	$[K - \langle K \rangle]^2$	$\left[ \frac{\overline{LCC}(K)}{-\langle \overline{LCC}(K) \rangle} \right]^2$	$\left[ \frac{K - \langle K \rangle}{\left[ \frac{\overline{LCC}(K)}{-\langle \overline{LCC}(K) \rangle} \right]} \right]^2$
2	0.00	-1.5	-0.45	2.25	0.2025	0.675
3	0.67	-0.5	0.22	0.25	0.0484	-0.110
4	0.54	0.5	0.09	0.25	0.0081	0.045
5	0.60	1.5	0.15	2.25	0.0225	0.225
3.5	0.45			Sum 5.00	0.2815	0.835
$\langle K \rangle$	$\langle \overline{LCC}(K) \rangle$					

$$\text{Randomness Index} = \frac{0.835}{\sqrt{5.00} * \sqrt{0.2815}} = 0.7038$$

real-world network graph and truly random graph, respectively. Let  $\langle K \rangle$  and  $\langle \overline{LCC}(K) \rangle$ , respectively, denote the average of the  $K$  values (represented by  $\langle K \rangle$ ) and average of the  $\overline{LCC}(K)$  values (represented by  $\langle \overline{LCC}(K) \rangle$ ). We determine the terms  $[K - \langle K \rangle]$  and  $[\overline{LCC}(K) - \langle \overline{LCC}(K) \rangle]$  as well as their squares and product. We observe the Randomness Index (Fig. 3) for the real-world network graph to be  $-0.9665$ , clearly vindicating our earlier assertion that the  $K$  and  $\overline{LCC}(K)$  values typically exhibit a strong negative correlation. On the other hand, the Randomness Index (Fig. 4) for the truly random graph (generated based on the  $\langle K \rangle$  and  $N$  for the real-world network graph) is  $0.7038$ , again vindicating our assertion that the Randomness Index for the truly random graphs generated based on the parameters for the real-world network graph could be close to zero or positive.

### 3 Real-world network graphs and their analysis

In this section, we first introduce the 48 real-world networks analyzed in this paper. Table 1 lists the three-character code acronym, name and the network type, 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 ( $\lambda_{sp}$ ). 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 1.0; the farther is the ratio from the value of 1.0, 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 3.48 (indicating that the real-world network graphs analyzed range from networks with smaller variation in node degree to networks with larger variation in node degree).

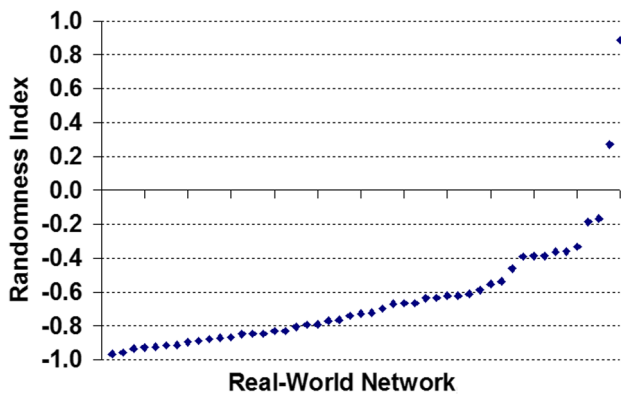
The networks considered cover a broad range of categories (as listed below along with the number of networks in each category): Acquaintance network (12), Friendship network (9), Co-appearance network (6), Employment network (4), Citation network (3), Literature network (3),

**Table 1** Real-World Networks used in the correlation analysis

#	Net.	Net. description	References	Network Type	$\lambda_{sp}$	#nodes	#edges	$k_{avg}$
1	ADJ	Word Adjacency Network	Newman (2006)	Co-appearance Net.	1.73	112	425	7.589
2	AKN	Anna Karenina 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 (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 Co-appear. 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 et al. (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	PBN	US Politics Books Network	Krebs (2003)	Literature Net.	1.42	105	441	8.4
34	PSN	Primary School Contact Net.	Gemmetto et al. (2014)	Acquaintance Net.	1.22	238	5539	46.546
35	PFN	Prison Friendship Network	MacRae (1960)	Friendship Net.	1.32	67	142	4.239
36	SJN	San Juan Sur Family Net.	Loomis et al. (1953)	Acquaintance Net.	1.29	75	155	4.133
37	SDI	Scotland Corp. Interlock Net	Scott (1900)	Employment Net.	1.94	230	359	3.122
38	SPR	Senator Press Release Net.	Grimmer (2010)	Political Net.	1.47	92	477	10.37
39	SWC	Soccer World Cup 1998 Net	Batagelj and Mrvar (2006)	Game Net.	1.45	35	118	6.743
40	SSM	Sawmill Strike Comm. Net.	Michael (1997)	Acquaintance Net.	1.22	24	38	3.167
41	TEN	Taro Exchange Network	Schwimmer (1973)	Acquaintance Net.	1.06	22	39	3.545
42	TWF	Teenage Female Friend Net.	Pearson and Michell (2000)	Friendship Net.	1.59	47	77	3.277
43	UKF	UK Faculty Friendship Net.	Nepusz et al. (2008)	Friendship Net.	1.35	83	578	13.928
44	APN	US Airports 1997 Network	Batagelj and Mrvar (2006)	Transportation Net.	3.22	332	2126	12.807
45	USS	US States Network	Meghanathan (2017)	Geographical Net.	1.24	49	107	4.37
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

**Table 2** Randomness Index of the real-world network graphs

#	Net.	Randomness Index	#	Net.	Randomness Index	#	Net.	Randomness Index
1	ADJ	-0.3614	17	GD96	-0.3320	33	PBN	-0.8048
2	AKN	-0.8962	18	MUN	-0.9662	34	PSN	-0.9139
3	JBN	-0.7923	19	GLN	-0.6649	35	PFN	-0.6210
4	CEN	-0.6221	20	HTN	-0.7648	36	SJN	-0.7275
5	CLN	-0.6977	21	HCN	-0.8457	37	SDI	-0.4612
6	CGD	-0.6650	22	ISP	-0.7218	38	SPR	-0.3867
7	CFN	-0.8447	23	KCN	-0.8779	39	SWC	-0.5376
8	DON	-0.3899	24	KFP	-0.1853	40	SSM	-0.8676
9	DRN	-0.5884	25	LMN	-0.8476	41	TEN	-0.8712
10	DLN	-0.3860	26	MDN	-0.8296	42	TWF	-0.6113
11	ERD	-0.6337	27	MTB	-0.6354	43	UKF	-0.7918
12	FMH	-0.7700	28	MCE	-0.9144	44	APN	-0.8887
13	FHT	-0.7403	29	MSJ	-0.8286	45	USS	-0.9234
14	FTC	-0.6694	30	AFB	-0.1667	46	RHF	-0.5527
15	FON	0.8855	31	MPN	-0.3598	47	WSB	-0.9266
16	CDF	-0.9336	32	MMN	0.2703	48	WTN	-0.9564



**Fig. 5** Sorted Order of the Randomness Index of the Real-World Networks

Collaboration network (2), Political network (2), Biological network (2), Game network (2), Geographical Network, 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 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 their flight connections) involved in public transportation. A *trade network* is a network of countries/people involved in certain trade. The reader is referred to Meghanathan (2017) for a more detailed description of the individual real-world networks.

Table 2 lists the Randomness Index values for the 48 real-world network graphs. Only two networks (# 15—Football Network and # 32—ModMath Network) incur a positive Randomness Index value, indicating that their edges are formed due to random association. This could be justified due to the fact that the Football Network comprises of edges whose end vertices are teams that have played against each other during the round-robin league. Since the round-robin league games are arbitrarily scheduled between any two teams, it is convincing to observe a neutral or positive correlation between degree and local clustering coefficient values. Likewise, the ModMath Network comprises of school superintendents who have an edge between them if at least either of them considers the

other person as someone who can effectively spread around modern Math methods among the schools in a county. In such a network, the links among the neighbor nodes of a node are more likely to be formed independent of the degree of the node.

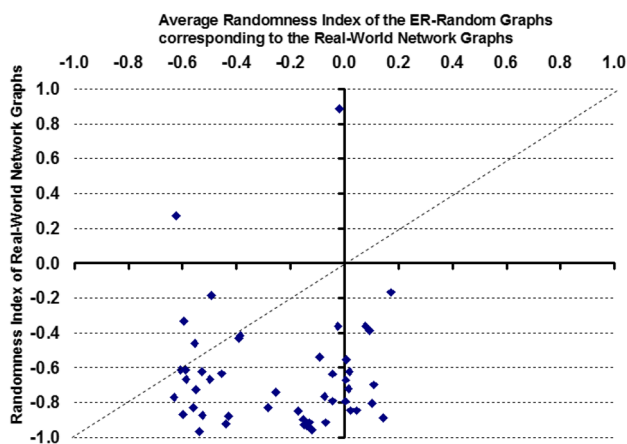
Figure 5 presents the Randomness Index values of the networks in the sorted order. We observe the median of the Randomness Index values to be  $-0.72$ , and 35 of the 48 real-world networks (i.e., about 70% of the networks) exhibit Randomness Index values less than  $-0.60$ , all of which indicate a strong negative correlation, as per the ordinal scale proposed by Evans (Evans 1995). Thus, the extent of randomness in real-world networks is minimal, and links are more likely not to get arbitrarily created. However, if links are to get randomly generated, then the

average local clustering coefficient of the nodes is more likely to be independent of node degree (like the truly random graph in Figs. 1, 4) and the Randomness Index of the networks is likely to be close to zero or positive.

### 3.1 Comparison with the Erdos–Renyi random network graphs

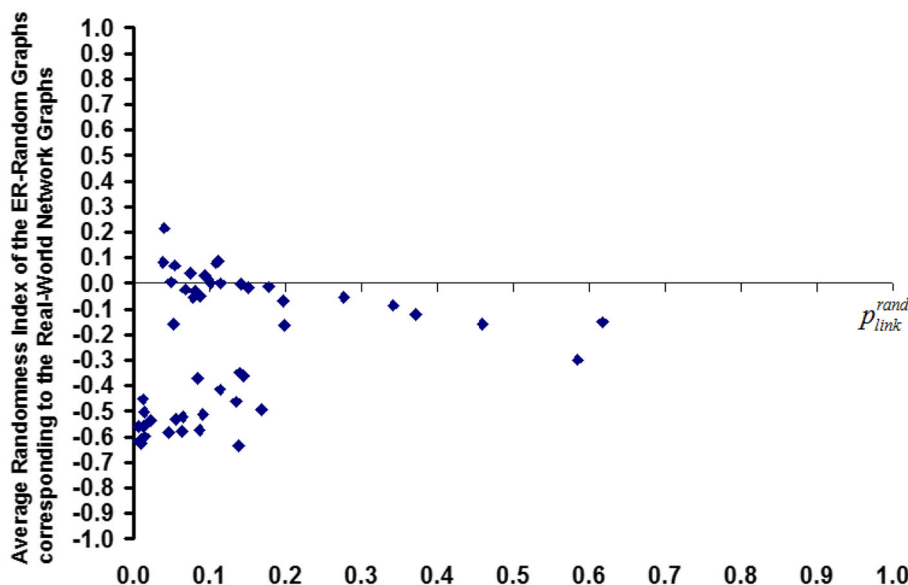
In this subsection, we compare the Randomness Index values obtained for the 48 real-world network graphs with those obtained for their corresponding random network graphs generated according to the well-known Erdos–Renyi (ER) model. For each real-world network graph, we generate 1000 instances of the corresponding ER-random network graphs following the procedure described in Sect. 2 and determine the average of the Randomness Index values obtained for these 1000 random network graphs. In Fig. 6, we plot the Randomness Index values observed for a real-world network graph versus the average of the Randomness Index values observed for its corresponding random network graphs generated according to the ER model. For 43 of the 48 real-world networks (the data points lie below the diagonal line), the Randomness Index values observed for the real-world networks are lower (i.e., relatively more closer to  $-1$ ) compared to the Randomness Index values observed for their corresponding ER-random graphs.

The median of the average Randomness Index values of the ER-random graphs is  $-0.12$  (see Fig. 7 for a distribution of the Randomness Index values vs. the probability for a link in the ER-random graphs corresponding to the real-world network graphs). For 13 of the 48 real-world network graphs, the average Randomness Index values of the



**Fig. 6** Randomness Index of the Real-World Network Graphs versus Average of the Randomness Index Values for the Corresponding Instances of the ER-Random Graphs

**Fig. 7** Probability for a link in the ER-Random Graphs corresponding to the Real-World Network Graphs versus Average of the Randomness Index Values for the ER-Random Graphs



**Table 3** Estimation of the power-law exponent ( $\gamma$ ) and the  $R^2$  values of these estimates for the real-world network graphs

#	Net.	Estimate of the power-law exponent ( $\gamma$ )	$R^2$ of the estimate	#	Net.	Estimate of the power-law exponent ( $\gamma$ )	$R^2$ of the estimate
1	ADJ	0.7768	0.6457	25	LMN	0.8565	0.6691
2	AKN	0.9355	0.7854	26	MDN	0.3699	0.0288
3	JBN	0.2661	0.1533	27	MTB	0.7191	0.5464
4	CEN	0.7987	0.4866	28	MCE	0.6822	0.1543
5	CLN	0.7727	0.7579	29	MSJ	2.0581	0.8941
6	CGD	1.3321	0.7973	30	AFB	0.7188	0.5206
7	CFN	0.7039	0.6502	31	MPN	0.3855	0.1131
8	DON	0.6233	0.4780	32	MMN	0.9689	0.5083
9	DRN	1.8458	0.8766	33	PBN	0.6614	0.2477
10	DLN	0.6602	0.6080	34	PSN	0.0330	0.0009
11	ERD	1.4945	0.8649	35	PFN	0.6967	0.2582
12	FMH	1.8304	0.8517	36	SJN	0.5759	0.1279
13	FHT	0.4441	0.2610	37	SDI	1.7105	0.6382
14	FTC	0.0022	0.0000	38	SPR	0.4538	0.3380
15	FON	6.4181	0.7148	39	SWC	0.3702	0.2024
16	CDF	0.2510	0.0907	40	SSM	0.3730	0.0635
17	GD96	1.2039	0.6458	41	TEN	3.1458	0.5016
18	MUN	1.2671	0.9395	42	TWF	0.8008	0.2530
19	GLN	1.2735	0.8422	43	UKF	0.2649	0.0759
20	HTN	0.0496	0.0045	44	APN	0.9456	0.7690
21	HCN	0.6304	0.4503	45	USS	0.0954	0.0043
22	ISP	0.7787	0.4158	46	RHF	0.2243	0.0308
23	KCN	0.5512	0.2914	47	WSB	0.3085	0.0683
24	KFP	0.2149	0.1541	48	WTN	0.4158	0.2005

corresponding ER-random graphs are positive (i.e., above 0.0) and for more than half of the real-world networks, the average Randomness Index values of the corresponding ER-random graphs are greater than  $-0.20$ . Such observations vindicate our earlier claim that if the links in a network came into existence due to just random association, the Randomness Index value for that network is more likely to be closer to 0 (rather than being closer to  $-1$ ). Only for 3 of the 48 real-world network graphs, their corresponding ER-random graphs had average Randomness Index values that are less than  $-0.60$ , whereas, as mentioned earlier (see Fig. 5), the Randomness Index values for 35 of the 48 real-world network graphs are less than  $-0.60$ . All of the above observations demonstrate the potential of the Randomness Index measure to categorically differentiate real-world network graphs from random network graphs.

### 3.2 Comparison with the power-law model scale-free network graphs

Scale-free network graphs are modeled using the power-law model (Barabasi and Albert 1999) according to

which the probability for observing a node with degree  $K$  is given by  $P(K) = CK^{-\gamma}$ , where  $C$  is the proportionality constant and  $\gamma$  is the power-law exponent. Unlike the Poisson model (Erdos and Renyi 1959; Ugarte et al. 2015) of degree distribution that is characteristic for random network graphs, the power-law model of degree distribution for scale-free network graphs is characteristic of a heavy tail (i.e., there exists a few, but appreciable number of nodes, with degree centrality values that are extremely larger compared to those for the rest of the nodes). The smaller the value for  $\gamma$ , the larger the chances for finding one or more node(s), typically called the hub nodes, with a degree that is much larger than the rest of the nodes in the network. On the other hand, the larger the value for  $\gamma$ , the lower the chances for finding the hub nodes; all the nodes are more likely to be of comparable degree (such a degree distribution is characteristic of the random network graphs). The convention in the literature (Barabasi 2016) is that networks with  $\gamma$  values less than 3.0 are considered to be in the scale-free regime and networks with  $\gamma$  values greater than 3.0 are considered to fall in the random network regime.



In this section, we fit the degree distributions of the real-world networks to follow a power-law model and estimate the  $\gamma$  values for the corresponding scale-free networks. For each real-world network, we count the number of nodes with a certain degree  $K$  and divide that count by the total number of nodes in the network to obtain the value of  $P(K)$ . Since  $P(K) = CK^{-\gamma}$ ,  $\ln P(K) = \ln C + (-\gamma) \ln(K)$ . We model the  $\ln(K)$  versus  $\ln P(K)$  distribution using linear regression and determine the slope of the line  $(-\gamma)$ . We observe the median of the estimated  $\gamma$  values to be 0.67, and only two of the 48 real-world networks (the # 15 FON and the # 41 TEN) have the estimated  $\gamma$  values greater than 3.0. However, we cannot very confidently claim that these real-world networks follow a power-law model, because the  $R^2$  values for the power-law model fits of these networks are not appreciably high; the median of the  $R^2$  values is only 0.43 (Table 3 displays the estimated  $\gamma$  values and the  $R^2$  values of these estimates for the 48 real-world networks considered, and Fig. 8 plots the same). It is interesting to note from Fig. 8 that lower the estimated  $\gamma$  values, the lower the  $R^2$  values of these estimates. Hence, even though the estimated  $\gamma$  values are low for the real-world networks, due to the lower  $R^2$  values for these estimates, the  $\gamma$  values cannot be alone construed as a measure of the scale-freeness (or indirectly, the lack of randomness) of a real-world network.

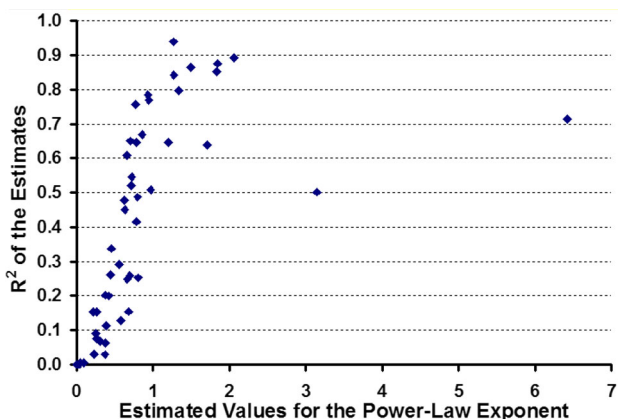
Our proposal for directly estimating the Randomness Index for a real-world network fills this gap. Rather than estimating the measure of scale-freeness of a real-world network and considering it as an indirect measure of the lack of randomness in the network, we propose to directly measure the extent of randomness of a real-world network using the Randomness Index measure. Figure 9 that plots the Randomness Index versus the estimated power-law exponent ( $\gamma$ ) indicates the same. Real-world networks with lower  $\gamma$  values have the Randomness Index values spread

over the entire range from  $[-1 \dots 0]$ . Hence, the power-law exponent  $\gamma$  would not be a sufficient measure to effectively quantify the extent of randomness of a real-world network.

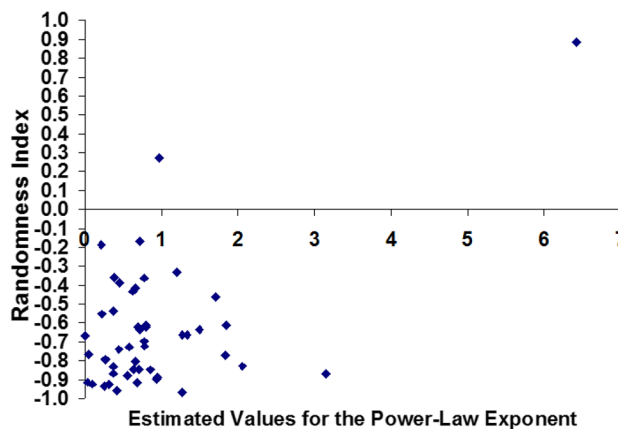
### 3.3 Comparison with the small-worldness index

A network is considered to be a small-world network (Newman 2010) if the average path length between any two nodes is proportional to the logarithm of the number of nodes in the network as well as the average local clustering coefficient of the nodes is high. The clustering coefficient for a network (Newman 2010) is the average of the local clustering coefficient of the individual nodes in the network. Structurally, a small-world network is considered to be in between a regular network (wherein all nodes have the same degree) and random network (Newman 2010). Regular networks have a larger clustering coefficient, but a larger average path length too. On the other hand, random networks have a lower average path length, but a lower clustering coefficient too. The well-known Watts–Strogatz model (Watts and Strogatz 1998) is typically used to transform a regular network to a small-world network by arbitrarily rewiring the links of the regular network with a probability  $\beta$ . If  $\beta = 1$ , the network is truly random. Typically,  $\beta$  values of 0.10 or lower have been observed (Meghanathan 2015) to be appropriate to transform a regular network to a random network without significantly decreasing the clustering coefficient, but at the same time appreciably decreasing the average path length.

In Humphries and Gurney (2008), the authors proposed a metric called “Small-worldness Index” that could be used to quantify the extent of small-worldness in a real-world network. The Small-worldness Index of a real-world network is computed as follows: We generate 1000 instances of the ER-random graphs for the real-world



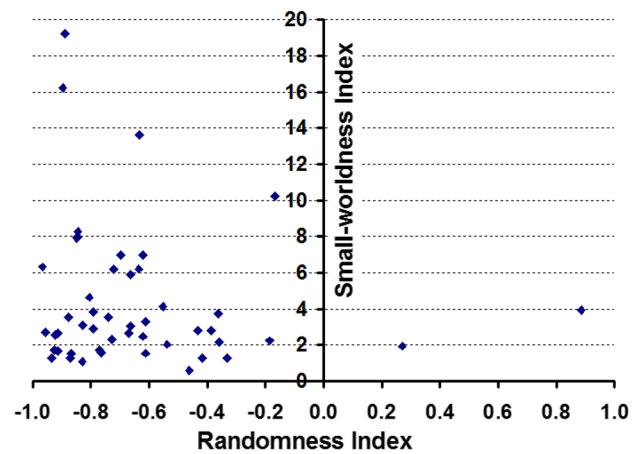
**Fig. 8** Distribution of the Estimated Values for the Power-Law Exponent ( $\gamma$ ) and the  $R^2$  Values for these Estimates for the Real-World Network Graphs



**Fig. 9** Distribution of the Estimated Values for the Power-Law Exponent ( $\gamma$ ) versus the Randomness Index Values for the Real-World Network Graphs

network (using the procedure described in Sect. 2) and determine the overall average clustering coefficient of these random graphs as well as the overall average path lengths between any two nodes in these graphs. We determine the clustering coefficient ratio as the ratio of the average clustering coefficient of the real-world network and the overall average clustering coefficient for all the corresponding ER-random graphs. Likewise, we determine the path length ratio as the ratio of the average path length between any two nodes in the real-world network and the overall average path length between any two nodes in the corresponding ER-random graphs. The average path length for a real-world network is expected to be larger than the overall average path incurred for its corresponding ER-random graphs. Hence, the average path length ratio for a real-world network is expected to be larger than 1.0, but closer is the ratio to 1.0 or even lower than 1.0, the larger the extent of small-worldness. On the other hand, the average clustering coefficient for a real-world network is expected to be larger than that incurred for its corresponding ER-random graphs. Hence, the Small-worldness Index of a real-world network is computed as the clustering coefficient ratio divided by the average path length ratio. The larger the magnitude of the Small-worldness Index for a real-world network, the larger the extent of small-worldness in the network (i.e., smaller path length and larger clustering coefficient of the nodes).

Table 4 displays the Small-worldness Index values computed for the 48 real-world networks, and Fig. 10 presents a plot showing the distribution of the Randomness



**Fig. 10** Distribution of the Randomness Index Values versus the Small-worldness Index Values for the Real-World Network Graphs

Index vs. Small-worldness Index values for the real-world network graphs. The median of the Small-worldness Index values is 2.85. For real-world networks with Small-worldness Index values much larger than the median (say, for example, networks with Small-worldness Index values of 6.0 or above), the Randomness Index values are less than  $-0.60$  (except for # 30: AFB network). Though we could see such an inverse relationship between the two indexes for highly small-world networks, we are not able to generalize it for all the real-world networks studied. In particular, for 25 of the 48 real-world networks that had Small-worldness Index values less than 3.0, the Randomness Index values range from  $-1.0$  to 0.20. Thus, the

**Table 4** Small-worldness Index of the real-world network graphs

#	Net.	Small-worldness Index	#	Net.	Small-worldness Index	#	Net.	Small-worldness Index
1	ADJ	3.7454	17	GD96	1.2951	33	PBN	4.6476
2	AKN	16.2410	18	MUN	6.3164	34	PSN	2.6542
3	JBN	3.8344	19	GLN	3.0292	35	PFN	2.4758
4	CEN	6.9976	20	HTN	1.5768	36	SJN	2.3040
5	CLN	6.9937	21	HCN	7.9463	37	SDI	0.5886
6	CGD	5.9216	22	ISP	6.2047	38	SPR	2.8095
7	CFN	8.2710	23	KCN	3.5292	39	SWC	2.0331
8	DON	2.8175	24	KFP	2.2426	40	SSM	1.5488
9	DRN	1.5486	25	LMN	7.9039	41	TEN	1.2619
10	DLN	1.2839	26	MDN	1.0799	42	TWF	3.2944
11	ERD	13.6027	27	MTB	6.2209	43	UKF	2.8887
12	FMH	1.7365	28	MCE	1.6601	44	APN	19.2159
13	FHT	3.5309	29	MSJ	3.0847	45	USS	2.5624
14	FTC	2.6611	30	AFB	10.2375	46	RHF	4.1281
15	FON	3.9315	31	MPN	2.1514	47	WSB	1.7135
16	CDF	1.2747	32	MMN	1.9368	48	WTN	2.7231

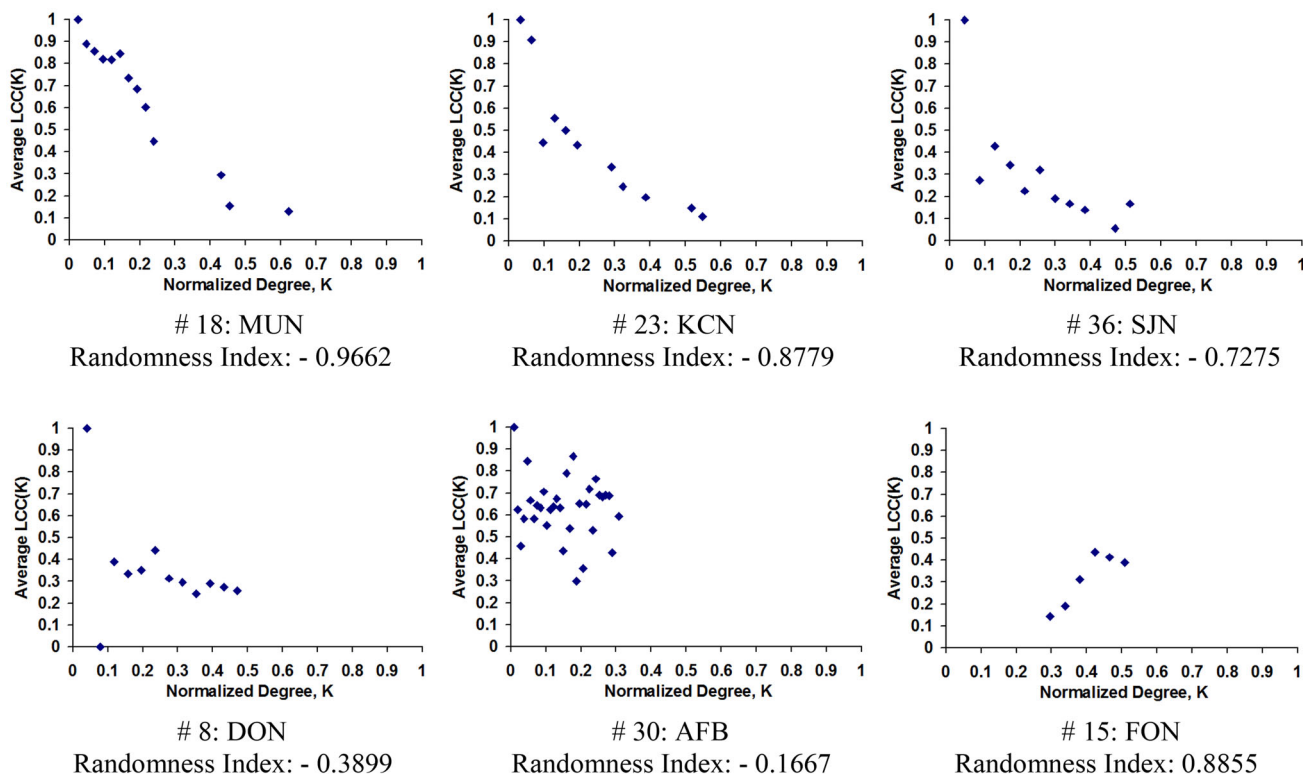


Fig. 11  $K$  vs.  $\overline{LCC}(K)$  Distribution for Selected Real-World Networks and Randomness Index Values

Small-worldness Index measure cannot be used as a direct measure to infer the extent of randomness of any arbitrary real-world network.

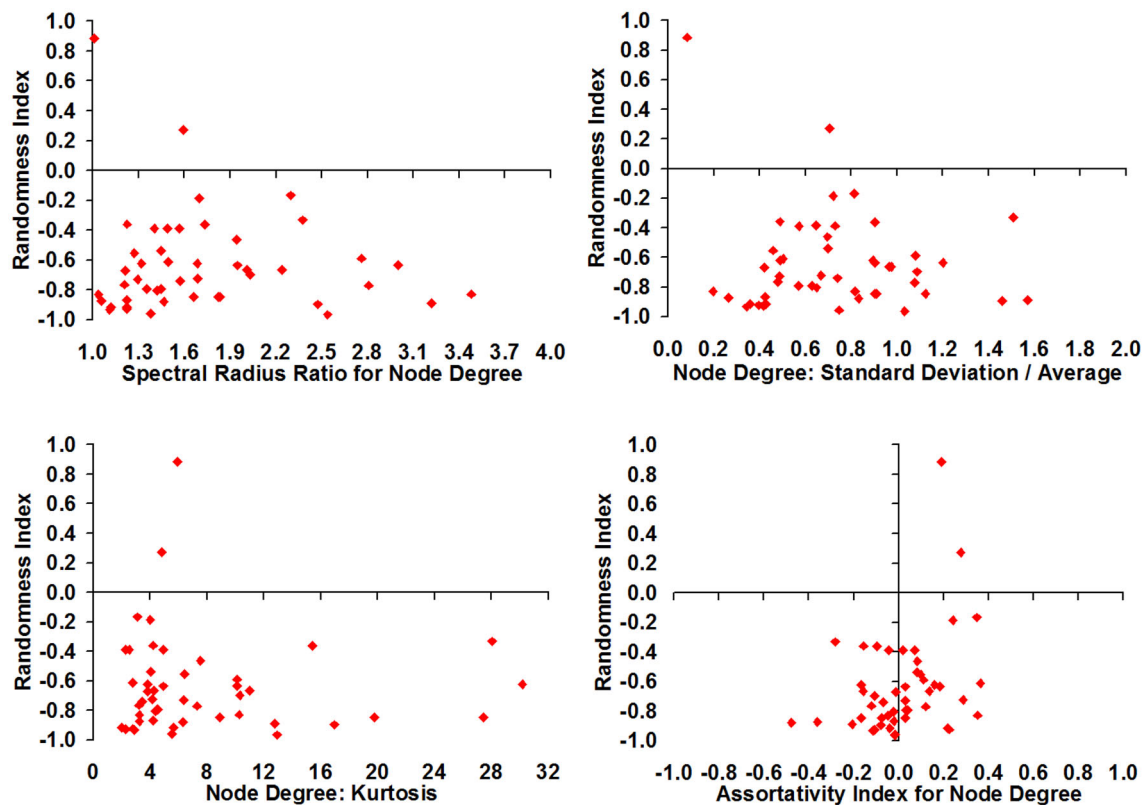
### 3.4 Transitions in the pattern of the $K$ versus $\overline{LCC}(K)$ distribution

Figure 11 presents the  $K$  versus  $\overline{LCC}(K)$  distribution (where the degree centrality values are normalized in a scale of 0 to 1 for uniformity) for selected networks such that we capture the transitions in the pattern of the distribution as the Randomness Index values change from  $-1$  to  $1$ . The  $K$  versus  $\overline{LCC}(K)$  distributions for networks MUN (# 18), KCN (# 23), SJN (# 36), DON (# 8), with Randomness Index values smaller than  $-0.35$  clearly show a negative correlation. The Author Facebook Network (# 30: AFB) shows a random pattern for the  $K$  versus  $\overline{LCC}(K)$  distribution, vindicating its Randomness Index value of  $-0.1667$ . The Football Network (# 15: FON) shows an opposite trend, wherein the  $\overline{LCC}(K)$  values increase with increase in  $K$  until certain values of  $K$  and then decrease; such a behavior is unexpected for real-world networks and can happen only when the LCC for a node is independent of the degree of the node (a characteristic of truly random networks and is also vindicated by the lowest spectral

radius ratio for node degree value of 1.01 for the Football Network).

### 3.5 Randomness Index versus metrics to quantify the extent of variation in node degree

Figure 12 presents a distribution of the Randomness Index values versus metrics that have been hitherto used to quantify the extent of variation in node degree (spectral radius ratio for node degree (Meghanathan 2014), ratio of the standard deviation to the average of node degree) or the heavy-tailed nature of a degree distribution (Kurtosis) or the degree-degree correlation of the end vertices of the edges (assortativity index) (Newman 2002). We do not observe any particular trend of increase or decrease in the Randomness Index values for each of these four metrics. We observe networks with lower variation in node degree (lower values for the spectral radius ratio for node degree or the ratio of the standard deviation to the average of node degree) need not have lower values for the Randomness Index, implying that the lack of variation in node degree cannot be perceived as the sole reason to expect a real-world network to exhibit randomness in the distribution of its links. For example, the Taro Exchange Network (# 41: TEN) and the Football Network (# 15: FON) have spectral radius ratio for node degree and standard deviation/average



**Fig. 12** Randomness Index versus {Spectral Radius Ratio, Standard Deviation/Average Ratio, Kurtosis, Assortativity Index} for node degree

degree ratio values of (1.06, 0.265) and (1.01, 0.083), respectively; however, the Randomness Index of the TEN is  $-0.8712$  (lack of randomness in the distribution of links among the nodes) and that of the FON is  $0.8855$  (significant randomness in the distribution of links among the nodes). Likewise, we also observe that two networks with vastly different Kurtosis values could have very close Randomness Index values. For example, the Anna Karenina Network (# 2: AKN) and the Fraternity College Dormitory Network (# 16: CDF) have Kurtosis values of 16.97 and 2.87, respectively, whereas their Randomness Index values ( $-0.8962$  and  $-0.9336$ ) are very close to each other. With respect to assortativity of the end vertices of the edges on the basis of node degree, we observe a majority of the real-world networks to exhibit assortativity index (calculated as the Pearson's correlation coefficient of the degrees of the end vertices of the edges) values closer to 0, irrespective of the Randomness Index.

#### 4 Related work

Several works in the literature [e.g., (Foudalis et al. 2011) for social networks, (Pizzuti and Rombo 2014) for protein-protein interaction networks] have reported the negative correlation between the degree of a vertex and the local

clustering coefficient of vertices with a particular degree. In (Bloznelis 2013), the authors attempted to study this phenomenon [to explore the hierarchical organization of real-world networks (Ravasz and Barabasi 2003)] using two random graph models (Godehardt and Jaworski 2001) that admit power-law degree distribution. However, there is no attempt to quantify the strength of this correlation and relate the value of the correlation coefficient with that of the extent of randomness in the network. To the best of our knowledge, ours is the first attempt to quantify the extent of randomness in the distribution of the links on the basis of the negative correlation of the  $K$  versus  $\overline{LCC}(K)$  distribution.

We use the classical Erdos–Renyi model (Erdos and Renyi 1959) as the basis for the degree-independent local clustering coefficient for the nodes (the local clustering coefficient for any node is close to the probability for a link between any two nodes) in truly random graphs. We do not advocate the use of other random network models [such as the configuration model (Britton et al. 2006)] because these models restrict the distribution of the links among the nodes in the network subject to certain constraints [such as a particular degree sequence (Meghanathan 2016) or an LCC sequence (Li and O’Riordan 2013)]. For example, in (Li and O’Riordan 2013), the authors exploited the negative  $K$  versus  $\overline{LCC}(K)$  correlation to evaluate the

cooperation among the nodes as well as the robustness of the network with regard to link removals and proposed algorithms that would generate graphs with certain LCC values for the nodes for a given degree distribution. Such models cannot be used to evaluate the extent of randomness in real-world networks.

In Soffer and Vazquez (2005), the authors proposed a novel approach to determine the maximum possible number of links between the neighbors of a node as part of the formulation to compute the local clustering coefficient for a node. If the degree constraints of the neighbors of a node are taken into consideration, the maximum possible number of links among the neighbors of a node is more likely to be lower than the maximum possible number of links used in the standard LCC formulation (that is also used in this paper). For example, if a node  $i$  has four neighbors (say, A, B, C, D) and all these neighbors have degree 2, as per (Soffer and Vazquez 2005), the maximum possible number of links between these neighbors is 2 (say, A–B and C–D) and not  $4(4 - 1)/2 = 6$ . The extent of randomness of a real-world network could not be evaluated with such a formulation for the LCC as it forces all the possible links incident on the neighbors of a node to be those connecting the neighbors of the nodes themselves (and the chances of randomness in the distribution of the links are suppressed). To corroborate our observation with regard to the computation of LCC, the  $K$  versus  $\overline{LCC}(K)$  distribution for several real-world networks based on the formulation of (Soffer and Vazquez 2005) has been observed to be a flat distribution, and with such a degree-independent distribution of LCC, the extent of randomness in a real-world network could not be evaluated.

## 5 Conclusions

The high-level contribution of this paper is the proposal for a quantitative metric to assess the extent of randomness in the distribution of the links in real-world networks. We make use of the fact that if a real-world network is not a truly random network, the chances for any two neighbors of a node to be connected (quantified in terms of the local clustering coefficient, LCC) decrease with increase in the degree for the node. We define the Randomness Index of a real-world network as the Pearson's correlation coefficient of the degree and average LCC of the nodes with the particular degree. We propose that stronger the negative correlation (i.e., closer the values of the correlation coefficient to  $-1$ ), the lower the extent of randomness in the distribution of the links in the network. For networks with lower values of the Randomness Index (close to  $-1$ ), the chances for any two neighbors of a node to be connected decrease with increase in degree, and hence, the links are

more likely to have not formed due to random association. For networks with Randomness Index closer to 0 or above 0, the chances for any two neighbors of a node to be connected do not depend on the degree of the node, and hence, the links between nodes are more likely to have formed due to random association.

We evaluate a suite of 48 real-world networks of diverse degree distribution and observe that the median of the Randomness Index values is as low as  $-0.72$  (indicating that the extent of randomness in a majority of the real-world networks is limited). We also show that the Randomness Index values for a real-world network and its corresponding ER-random graphs are indeed different and the former is more likely to be lower. We advocate that the Randomness Index metric is a dimensionless metric (the values always range from  $-1$  to  $1$  and do not depend on the actual degree values of the nodes or the number of nodes and edges in the network) that could be unequivocally used (unlike the currently used approaches of evaluating the degree distribution, standard deviation/average ratio, connectedness, assortativity index, etc.) to evaluate the extent of randomness in any real-world network. Further, we show that the Randomness Index metric is a unique metric for complex network analysis and is different from some of the well-known modeling parameters like the power-law exponent for scale-free networks and Small-worldness Index for small-world networks. We also show that it does not have similarity to any of the existing metrics to quantify the variation or correlation of node degree.

## References

- Al-Shiridah G, Mahdi K, Safar M (2013) Feedback capacity of different complex network models. In: Proceedings of the 27th international conference on advanced information networking and applications workshops. Barcelona, pp 996–1003
- Barabasi AL (2016) Network science, 1st edn. Cambridge University Press, Cambridge
- Barabasi AL, Albert R (1999) Emergence of scaling in random networks. *Science* 286(5439):509–512
- Batagelj V, Mrvar A (2006) Pajek datasets. <http://vlado.fmf.uni-lj.si/pub/networks/data/>
- Bernard HR, Killworth PD, Sailer L (1980) Informant accuracy in social network data IV: a comparison of clique-level structure in behavioral and cognitive network data. *Soc Netw* 2(3):191–218
- Biedl T, Franz BJ (2001) Graph-drawing contest report. In: Proceedings of the 9th international symposium on graph drawing. pp 513–521
- Bloznelis M (2013) Degree and clustering coefficient in sparse random intersection graphs. *Ann Appl Probab* 23(3):1254–1289
- Bonacich P (1987) Power and centrality: a family of measures. *Am J Sociol* 92(5):1170–1182
- Britton T, Deijfen M, Martin-Lof A (2006) Generating Simple random graphs with prescribed degree distribution. *J Stat Phys* 124(6):1377–1397

- Cross RL, Parker A, Cross R (2004) *The hidden power of social networks: understanding how work really gets done in organizations*, 1st edn. Harvard Business Review Press, Cambridge
- de Nooy W (1999) A literary playground: literary criticism and balance theory. *Poetics* 26(5–6):385–404
- Erdos P, Renyi A (1959) On random graphs I. *Publ Math* 6:290–297
- Evans JD (1995) *Straight forward Statistics for the behavioral sciences*, 1st edn. Brooks Cole Publishing Company, Pacific Grove
- Foudalis I, Jain K, Papadimitriou C, Sideri M (2011) Modeling social networks through user background and behavior. In: *Proceedings of the 8th international conference on algorithms and models for the web graph*. Atlanta, pp 85–102
- Freeman LC, Freeman SC, Michaelson AG (1989) How humans see social groups: a test of the Sailer–Gaulin models. *J Quant Anthropol* 1:229–238
- Freeman LC, Webster CM, Kirke DM (1998) Exploring social structure using dynamic three-dimensional color images. *Soc Netw* 20(2):109–118
- Geiser P, Danon L (2003) Community structure in jazz. *Adv Complex Syst* 6(4):563–573
- 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
- Gil-Mendieta J, Schmidt S (1996) The political network in Mexico. *Soc Netw* 18(4):355–381
- Girvan M, Newman MEJ (2002) Community structure in social and biological networks. *Proc Natl Acad Sci USA* 99(12):7821–7826
- Gleiser PM (2007) How to become a superhero. *J Stat Mech Theory Exp* 2007:P09020
- Godehardt E, Jaworski J (2001) Two models of random intersection graphs and their applications. *Electron Notes Discrete Math* 104:129–132
- Grimmer J (2010) A Bayesian hierarchical topic model for political texts: measuring expressed agendas in Senate press releases. *Political Anal* 18(1):1–35
- Hayes B (2006) Connecting the dots. *Am Sci* 94(5):400–404
- Hummon NP, Doreian P, Freeman LC (1990) Analyzing the structure of the centrality-productivity literature created between 1948 and 1979. *Sci Commun* 11(4):459–480. doi:10.1177/107554709001100405
- Humphries MD, Gurney K (2008) Network ‘small-world-ness’: a quantitative method for determining canonical network equivalence. *PLoS ONE* 3(4):e0002051
- Isella L, Stehle J, Barrat A, Cattuto C, Pinton JF, Van den Broeck W (2011) What’s in a crowd? Analysis of face-to-face behavioral networks. *J Theor Biol* 271(1):166–180. doi:10.1016/j.jtbi.2010.11.033
- Knuth DE (1993) *The Stanford GraphBase: a platform for combinatorial computing*, 1st edn. Addison-Wesley, Reading
- Krackhardt D (1999) The ties that torture: Simmelian tie analysis in organizations. *Res Sociol Organ* 16:183–210
- Krebs V (2003) Proxy networks: analyzing One network to reveal another. *Bull Methodol Sociol* 79:61–70
- Lee J-S (2004) Generating networks of illegal drug users using large samples of partial ego-network data. In: *Intelligence and security informatics (Lecture notes in computer science)*, vol 3073. pp 390–402
- Li M, O’Riordan C (2013) The Effect of clustering coefficient and node degree on the robustness of cooperation. In: *Proceedings of the 2013 IEEE congress on evolutionary computation*. Cancun, pp 2833–2839
- Loomis CP, Morales JO, Clifford RA, Leonard OE (1953) *Turrialba social systems and the introduction of change*. The Free Press, Glencoe, pp 45–78
- Lusseau D, Schneider K, Boisseau OJ, Haase P, Slooten E, Dawson SM (2003) The bottlenose dolphin community of Doubtful Sound features a large proportion of long-lasting associations. *Behav Ecol Sociobiol* 54(3):396–405
- MacRae D (1960) Direct factor analysis of sociometric data. *Sociometry* 23(4):360–371
- C. McCarty and L. Freeman, <http://moreno.ss.uci.edu/data.html>, 2008
- Meghanathan N (2014) 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*. Hainan, pp 30–33
- Meghanathan N (2015) Distribution of maximal clique size of the vertices for theoretical small-world networks and real-world networks. *Int J Comput Netw Commun* 7(4):21–41
- Meghanathan N (2016) On the sufficiency of using the degree sequence of the vertices to generate random networks corresponding to real-world networks. *Polibits Res J Comput Sci Comput Eng Appl* 53:5–21
- Meghanathan N (2017a) 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
- Meghanathan N (2017b) Complex network analysis of the contiguous United States graph. *Comput Inf Sci* 10(1):54–76
- Michael JH (1997) Labor dispute reconciliation in a forest products manufacturing facility. *For Prod J* 47(11–12):41–45
- Moreno JL (1960) *The sociometry reader*. The Free Press, Glencoe, pp 534–547
- 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
- Newman MEJ (2002) Assortative mixing in networks. *Phys Rev Lett* 89(20):208701
- Newman MEJ (2010) *Networks: an introduction*, 1st edn. Oxford University Press, Oxford
- Newman MEJ (2006) Finding community structure in networks using the eigenvectors of matrices. *Phys Rev E* 74(3):036104
- 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
- Pizzuti C, Rombo SE (2014) Algorithms and tools for protein-protein interaction networks clustering, with a special focus on population-based stochastic methods. *Bioinformatics* 30(10):1343–1352
- Ravasz L, Barabasi AL (2003) Hierarchical organization in complex networks. *Phys Rev E* 67:026112
- Resnick MJ, Bearman PS, Blum RW, Bauman KE, Harris KM, Jones J, Tabor J, Beuhring T, Sieving RE, Shew M, Ireland M, Bearinger LH, Udry JR (1997) Protecting adolescents from harm. Findings from the National Longitudinal Study on Adolescent Health. *J Am Med Assoc* 278(10):823–832
- Rogers EM, Kincaid DL (1980) *Communication networks: toward a new paradigm for research*. Free Press, New York
- 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., London
- Scott JP (1980) *The anatomy of Scottish capital: Scottish companies and Scottish capital*, 1st edn. Brooks Cole Publishing Company, Pacific Grove, pp 1900–1979
- Smith DA, White DR (1992) Structure and dynamics of the global economy: network Analysis of international trade 1965–1980. *Soc Forces* 70(4):857–893
- Soffer SN, Vazquez A (2005) Clustering coefficient without degree correlations biases. *Phys Rev* 7:057101
- Strang G (2006) *Linear algebra and its applications*, 4th edn. Brooks Cole, Pacific Grove
- Takahata Y (1991) Diachronic changes in the dominance relations of adult female Japanese monkeys of the Arashiyama B group. In:

- Fedigan LM, Asquith PJ (eds) The monkeys of Arashiyama. State University of New York Press, Albany, pp 124–139
- Ugarte MD, Militino AF, Arnholt AT (2015) Probability and statistics with R, 2nd edn. Chapman and Hall, Boca Raton
- Watts DJ, Strogatz SH (1998) Collective dynamics of ‘small-world’ networks. *Nature* 393:440–442
- White JG, Southgate E, Thomson JN, Brenner S (1986) The structure of the nervous system of the nematode *Caenorhabditis elegans*. *Philos Trans B* 314(1165):1–340
- Zachary WW (1977) An information flow model for conflict and fission in small groups. *J Anthropol Res* 33(4):452–473