Journal of Applied Analysis and Computation
Volume 6, Number 4, November 2016, 1105–1113
Website:http://jaac-online.com/
DOI:10.11948/2016072

DEGREE SEQUENCES BEYOND POWER LAWS IN COMPLEX NETWORKS*

Zhanying Zhang[1], Wenjun Xiao[2,†] and Guanrong Chen[3]

Abstract Many complex networks possess vertex-degree distributions in a power-law form of $ck^{-\gamma}$, where k is the degree variable and c and γ are constants. To better understand the mechanism of power-law formation in real-world networks, it is effective to analyze their degree variable sequences. We had shown before that, for a scale-free network of size N ,if its vertex-degree sequence is $k_1 < k_2 < \ldots < k_l$, where $\{k_1, k_2, \ldots, k_l\}$ is the set of all unequal vertex degrees in the network, and if its power exponent satisfies $\gamma > 1$, then the length l of the vertex-degree sequence is of order $logN$. In the present paper, we further study complex networks with more general distributions and prove that the same conclusion holds even for non-network type of complex systems. In addition, we support the conclusion by verifying many real-world network and system examples. We finally discuss some potential applications of the new finding in various fields of science, technology and society.

Keywords Network, degree variable sequence, power-law distribution, general distribution.

MSC(2010) 05C82, 05C07.

1. Introduction

Surrounding us are various complex networks, such as the Internet, wireless communication networks, software networks, social networks, and biological networks etc. Recently, studies on mathematical theory and modelling of complex networks have received a renewal of interest, with many generic and realistic graph-theoretic models developed, especially the small-world network model [16] and scale-free network model [2]. Scale-free networks, in particular, have a heterogeneous connectivity, where a small fraction of vertices are highly connected while the spare vertices have small numbers of connections. The scale-free network model of Barabsi and Albert [2] revealed an essential power-law distribution of vertex degrees in many complex networks, in the form of $P(k) \propto k^{-\gamma}$ where k is the degree variable and γ is a constant determined by the network. This power-law distribution is a direct consequence of two general mechanisms that generate the network topology: (i) network expansion over time through addition of new vertices; (ii) preferential attachment of new vertex to those existing ones in the network.

[†]the corresponding author. Email address:wjxiao@scut.edu.cn(W. Xiao)
[1]School of Computer Science and Engineering South China University of Technology, Guangzhou 510006, China
[2]School of Software Engineering South China University of Technology, Guangzhou 510006, China
[3]Department of Electronic Engineering City University of Hong Kong, Kowloon, Hong Kong SAR, China
*The authors were supported by National Natural Science Foundation of China (61170313, 61103037, 61370003).

In many real-world scale-free networks, the power-law exponent satisfies $\gamma \geq 2$ [1,11,14,15,18,19], but there are also many others with $\gamma < 2$ [3,6,8,9,12,17]. In our previous work [17–19], we presented some necessary conditions for the scale-free property of networks, based on the assumption of $\gamma > 1$. In this paper, we further show that the same conclusion holds also for complex networks with a more general type of vertex-degree distributions. We furthermore verify that the same results hold also for complex systems that may not be networks, and finally demonstrate that the new finding has many potential applications in science and technology, including economy, finance, society, and so on.

2. Scale-free networks and existing results

A complex network can be represented by an undirected or a directed graph, $G(V, E)$, where V is the set of vertices and E the set of edges. A graph has a number of local and global parameters that characterize its structure (e.g., regularity, modularity), connectivity (e.g., density, diffusion) and robustness (e.g., resilience to random attacks or malicious faults). The following list summarizes the main parameters to be used in this paper.

- M: Number of edges; $M = |E|$
- N: Number of vertices; $N = |V|$
- k_i:Degree of vertex,$i \in V$
- $\overline{d}$:Average vertex degree of the network
- n_{k_i}: Number of degree-k_i vertices;$\sum_{i=1}^{l} n_{k_i} = N$ and $\sum_{i=1}^{l} n_{k_i} k_i = 2M$
- l: Length of an unequal vertex-degree sequence $\{k_1, k_2, \ldots, k_l\}$,$1 \leq k_1 < k_2 < \ldots < k_l$
- K: Set of all unequal vertex degrees in a network
- $P(k)$: Degree distribution, or fraction of vertices of degree k;$P(k) = n_k/N$

For scale-free networks, one has

$$P(k) = ck^{-\gamma}, \ \gamma > 1.$$

Here, the requirement of $\gamma > 1$ ensures that $P(k)$ can be normalized and, in this case, the constant c can be used for normalization, $c = (\sum_{k \in K} k^{-\gamma})^{-1}$.

In our earlier works [18,19], we proved that for a scale-free network of size N, having a power-law distribution with exponent $\gamma \geq 2$, the number of degree-1 vertices, if not zero, tends to be of order N; and we also proved that the average degree is of order lower than $\log N$. Our method provides an analytical tool that helps one to check if a given network is scale-free, which relies on static conditions that can be easily verified. Furthermore, we showed that the number of degree-1 vertices is divisible by the least common multiples of $k_1^{\gamma}, k_2^{\gamma}, \ldots, k_l^{\gamma}$, if they are all integers, where $k_1 < k_2 < \ldots < k_l$ is the vertex-degree sequence of the network. This remodeling method equips a scale-free network with some small-world features. Based on our earlier results [18,19], lately we further showed [17] that for scale-free networks with $\gamma > 1$, the length l of the vertex-degree sequence is of order $\log N$. Here, it should be emphasized that this result is very important, which demonstrates that the length of the degree sequence is an essence of a general scale-free network. In fact, all scale-free networks have very small numbers of degree sequences in comparison with the network sizes. Thus, by utilizing this characteristic, one can reconstruct a scale-free network with prominent small-world features [19] and can also improve some commonly-used maximal-degree search algorithms.

It must also be stressed that the above conclusion holds based on the precondition that the network obeys a precise power-law degree distribution. Actually, many real-world scale-free networks are not exactly so, but have approximate scale-free features; therefore, there exist subtle differences between such real networks and our theoretical results.

Furthermore, it must be pointed out that for many real-world networks, the lengths of their vertex-degree sequences are of order $(\log N)^{\varepsilon}$, at the most, namely $l \leq O((\log N)^{\varepsilon})$, where ε is a very small constant in the sense that l is very small in comparing with the network size value. Finally in this paper, the above conclusion will be verified by some real-world networks existing in different fields.

3. A characteristic of scale-free networks

In this section, we present a useful property of scale-free networks and its mathematical derivation.

Suppose that the vertex-degree sequence of network is $1 \leq k_1 < k_2 < \ldots < k_l$. For scale-free networks, one has [1]

$$P(k_i) = \frac{n_{k_i}}{N} = ck_i^{-\gamma}. \tag{3.1}$$

Here, n_{k_i} is the number of vertices with degree k_i , satisfying

$$N = \sum_{i=1}^{l} n_{k_i}, \tag{3.2}$$

and c is a normalizing constant: when $i = 1$,

$$P(k_1) = \frac{n_{k_1}}{N} = ck_1^{-\gamma}. \tag{3.3}$$

Thus,

$$c = \frac{n_{k_1} k_1^{-\gamma}}{N}. \tag{3.4}$$

By substituting (3.4) into (3.1), one obtains

$$n_{k_i} = n_{k_1} (\frac{k_1}{k_i})^{\gamma}. \tag{3.5}$$

4. General structure of vertex-degree sequnces in networks with general topologies

4.1. Vertex-Degree Sequences in General Networks

In this subsection, we study some conditions for the vertex degrees of a complex network to have a general distribution. Assume that the networks are connected, but similar arguments also apply to disconnected networks.

Recall that $P(K)$ is the probability distribution of the number of degree-k vertices, $\overline{d}$ represents the average vertex degree, and n_k denotes the number of vertices of degree k. Thus, $M = \frac{N\overline{d}}{2}$ and $n_k = NP(k)$, where N and M are the numbers of vertices and edges, respectively.

For constants $\gamma \geq 0, q > 1$, consider a network with a general vertex-degree distribution:

$$P(k_i) = \frac{n_{k_i}}{N} = ck_i^{-\gamma} q^{1-k_i}. \tag{4.1}$$

It follows that

$$n_{k_i} = n_{k_1} (\frac{k_i}{k_1})^{-\gamma} q^{k_1 - k_i}. \tag{4.2}$$

Therefore,

$$\log N \geq \log n_{k_1} \geq \log n_{k_l} + \gamma \log(\frac{k_l}{k_1}) + (k_l - k_1) \log q \geq (l-1) \log q.$$

That is, l is of order $\log N$ for $\gamma \geq 0$, $q > 1$.

Example 4.1. Let $P(k) = ck^{-\gamma} e^{-\frac{k}{\kappa}}$, where c and κ are constants. Many real small-world networks are of this type, including the World Wide Web and collaboration graphs of scientists as well as Fortune 1000 company directors [13].

Based on the preceding definitions, one has

$$\sum_{k=1}^{N-1} n_k = N, \tag{4.3a}$$

$$\sum_{k=1}^{N-1} kn_k = 2M. \tag{4.3b}$$

Assuming that $n_1 \neq 0$, a condition that is satisfied by most networks with a large size N, one has $n_k = n_1 P(k)/P(1)$ and

$$\sum_{k=1}^{N-1} [P(k)/P(1)] = N/n_1, \tag{4.4a}$$

$$\sum_{k=1}^{N-1} [kP(k)/P(1)] = 2M/n_1. \tag{4.4b}$$

For networks with a general distribution, one has $P(k) = P(1)k^{-\gamma} q^{1-k}$, which leads to

$$\sum_{k=1}^{N-1} k^{-\gamma} q^{1-k} = \frac{N}{n_1}. \tag{4.5}$$

Therefore, based on the assumption of $q > 1$, a condition which holds for many real networks [13], one obtains

$$\frac{N}{n_1} \leq \sum_{k=1}^{\infty} k^{-\gamma} q^{1-k} \leq \frac{q^2}{q-1}. \tag{4.6}$$

This leads to the result of $N \approx n_1$ for some constant $q > 1$. Generally, for any network, one has

$$N = \sum_i n_{k_i} \le n_{k_1} q^{k_1} \sum_i q^{-k_i} \le \frac{n_{k_1} q^{k_1+1}}{(q-1)}, \tag{4.7}$$

$$2M = \sum_i k_i n_{k_i} \le n_{k_1} q^{k_1} \sum_i k_i q^{-k_i} \le \frac{n_{k_1} q^{k_1+1}}{(q-1)^2}. \tag{4.8}$$

All these together implies

$$\overline{d} \le \frac{n_{k_1} q^{k_1+1}}{N(q-1)^2} \le \frac{q^{k_1+1}}{(q-1)^2}. \tag{4.9}$$

It implies that the average vertex degree $\overline{d}$ is related to q and k_1. Thus, $\overline{d}$ is small because q and k_1 are generally small, in real networks (see Table 1).

Table 1. LENGTHS OF VARIABLE(VERTEX-DEGREE) SEQUENCES IN SOME REAL-WORLD SYSTEMS (NETWORKS)

System(Network)	Type	N	M	$\overline{d}$	l	$\log N$	ε
TG City	TP[a]	18263	23797	2.606	7	14.157	1
OL City	TP	6105	7029	2.303	5	12.576	1
US Air	TP	332	2126	12.807	58	8.375	2
Linux	SW[b]	5285	11352	4.296	51	12.368	2
Mysql2	SW	1480	4190	5.662	43	10.531	2
AbiWord	SW	1035	1719	3.322	29	10.015	2
Helico	Bio[c]	710	1396	3.932	31	9.472	2
Elegans	Bio	314	363	2.312	17	8.295	2
Wiki-Vote	SC[d]	7066	100736	28.519	308	12.787	2.3
Twitter	SC[e]	81306	2420766	59.547	1245	16.311	2.6
2004 Internet	TN[f]	18408	33963	3.690	121	14.168	2
2009 Internet	TN	31164	63226	4.058	172	14.928	2
Web-Google	Web[g]	855802	5066842	11.841	748	19.707	2.2
Web-Stanford	Web[h]	255265	2234572	17.508	730	17.962	2.3
Ncstrlwg2	CN[i]	6396	15872	4.963	42	12.643	2
CA-HepPh	CN[j]	11204	117634	20.999	288	13.452	2.2

a. For data on transportation networks (TP), see www.cs.fsu.edu/ lifeifei/SpatialDataset.htm.

b. For software networks (SW), see www.tc.cornell.edu/ myers/Data/SoftwareGraphs/index.html.

c. For biological networks (Bio), see www.cosin.org/extra/data.

d. For Wiki-Vote(SC), see http://snap.stanford.edu/data/Wiki-Vote.html.

e. For twitter(SC), see http://snap.stanford.edu/data/egonets-Twitter.html.

f. For Internet (TN), see www.caida.org.

g. For Web-Google (Web), see http://snap.stanford.edu/data/web-google.html.

h. For Web-Stanford(Web), see http://snap.stanford.edu/data/web-Stanford.html.

i. The data of Ncstrlwg2 (CN) are provided by M E J Newman [10].

j. For CA-HepPh(CN), see http://snap.stanford.edu/data/ca-HepPh.html.

In general, let $f(x)$ be a monotonously decreasing function, satisfying $n_{k_i} = n_{k_1} f(k_i)$. One can obtain networks of various vertex-degree distributions. For example, when $f(k_i) = (\frac{k_1}{k_i})^{\gamma}$, one has a scale-free network, and when $f(k_i) = q^{k_1 - k_i}$, one gets a network with an exponential distribution. On the other hand, one has

$$N = \sum_i n_{k_i} = n_{k_l} \sum_i \frac{n_{k_i}}{n_{k_l}} = n_{k_l} \sum_i \frac{f(k_i)}{f(k_l)}.$$

If l is of order $\log(\sum_i \frac{f(k_i)}{f(k_l)})$, then l is of order $\log N$. For instance, letting $f(k_i) = (\frac{k_i}{k_1})^{-\gamma} q^{k_1 - k_i}$ gives

$$N = n_{k_l} \sum_i \frac{f(k_i)}{f(k_l)} = n_{k_l} k_l^{\gamma} q^{k_l - k_1} \sum_i \frac{q^{k_1 - k_i}}{k_i^{\gamma}} \geq q^{k_l - k_1} \geq q^{l-1},$$

which shows that l is of order $\log N$ for constant $q > 1$.

It can be seen from Table 1 that $l < (\log N)^3$ for many real-world networks. It must be stressed that the above conclusion holds based on the precondition that the network obeys precise vertex-degree distributions. However, many real-world networks are not exactly so, which only have some approximate vertex-degree distributions, so there are small differences between such real networks and our theoretical results.

4.2. General System Structure and Potemtial Applications

In this subsection, the above results are generalized to general systems, not necessarily networks. Some typical examples are given to demonstrate the universality of the theoretical results.

Here, the network G is replaced by the system G, the vertex is replaced by some system quantity, the degree of vertex is replaced by the corresponding discrete variable, the number of degree-k vertices is replaced by the number of variable-k quantities, and so on. Although there may not have edge in such a general system, one can still define a quantity or index by

$$2M = \sum_i k_i n_{k_i}.$$

Thus, the results in Subsection 4.1 above hold for such systems.

Example 4.2. It follows from (3.5) that $n_{k_i} = n_{k_1}(\frac{k_1}{k_i})^{\gamma}$. There are many systems of this type. Let, for instance, $i = 1, 2, \ldots, l$, and assume that $\gamma = \log_q p$, $k_i = q^i$, where p is a positive constant. Then, one has $n_{k_i} = n_{k_l} p^{l-i}$. The familiar recursive clique trees belong to this type, with $p = 3$, $q = 2$ ([5]). It is known that the vertex-degree sequence of the network is $2, 2^2, \ldots, 2^l, 2^{l+1}$. The corresponding vertex numbers are $3^l, 3^{l-1}, \ldots, 3, 3$, respectively. It is not an exact scale-free network, however, if the final number 2^{l+1} of the sequence is ignored, then the new sequence is $2, 2^2, \ldots, 2^l$, with the corresponding vertex numbers $3^l, 3^{l-1}, \ldots, 3$, respectively. Hence, in this case, $k_i = 2^i, n_{k_i} = 3^{l+1-i}, i = 1, 2, \ldots, l$. Let $p = 3, q = 2$. Thus, $\gamma = \log_2 3$ and so (3.5) holds.

Example 4.3. Let $P(k) = ce^{-\lambda k}$, where λ is an adjustable parameter of the underlying system, which may not be a network. Two mechanisms, growing and random adjacent attachment, introduce a new type of systems. The simulation results [4] indicate that this type of systems indeed have exponential variable (may not be degree) distributions, which matches the empirical data very well, showing that the length of discrete variable sequences in the system is of order $\log N$ as described by (4.2).

Example 4.4. Let $P(k) = ck^{-\gamma}$, where $1 \leq \gamma \leq 3$ for large N. Systems with variables obeying this type of distributions are common in economy, finance and society [4]. It possesses a hierarchical structure and the length of its discrete variable sequences is small, also described by (4.2).

Example 4.5. Let

$$P(k_i) = \frac{n_{k_i}}{N} = ck_i^{-\gamma} q^{1-k_i}, \quad i = 1, 2, \ldots, l,$$

where $0 \leq \gamma \leq 3$, $1 \leq q \leq 3$ for large N. Systems with this type of variable distributions are also common in economy, science, technology and society [20]. The length l of their discrete variable sequences, denoted also by $\{k_1, k_2, \ldots, k_l\}$, which may not be degrees but also described by (4.2), is small when $q > 1$. Hence, by (4.2), one can write

$$n_{k_l} = n_{k_1} (\frac{k_l}{k_1})^{-\gamma} q^{k_1 - k_l}. \tag{4.10}$$

Finally, one can let $a = \frac{n_{k_1}}{n_{k_l}}$, $b = \frac{k_l}{k_1}$, so as to obtain

$$a = b^{\gamma} q^{(b-1)k_1}. \tag{4.11}$$

Example 4.6. Let the variable k_i denote income in finance. Then, n_{k_i} is the number of some quantity associated with the income k_i. Let also $a =$(number of quantity of minimum income)/(number of quantity of maximum income) and $b =$(maximum income)/(minimum income). If $a = 10^8$, $b = 10^4$, which are reasonable assumptions on the total income of the world population, then by (4.11) one has

$$10^8 = 10^{4\gamma} q^{(10^4-1)k_1}.$$

One may assume without loss of generality that $k_1 = 1$, so q is close to 1. If $q = 1$, then the distribution is a power law, with $\gamma = 2$. If $\gamma = 0$, then the distribution is exponential, $q > 1$ and is close to 1.For the lower class of income [20], one may assume that $a = 10^4$, $b = 10$, so that by (4.11) one has

$$10^4 = 10^{\gamma} q^{(10-1)k_1}.$$

Example 4.7. Let the variable k_i denote the size of a city, i, and rank all cities according to their sizes. Then, n_{k_i} is the rank of the city size k_i in the ranked sequence. Clearly, n_{k_1} is the rank of the minimum size k_1,n_{k_l} is the rank of the maximum size k_l. Furthermore, let a =(rank of minimum size)/(rank of maximum size), and b =(maximum size)/(minimum size). If $a = 10^3$, $b = 10^3$, which are reasonable assumptions on the size of the city population [4], then it follows from (4.11) that

$$10^3 = 10^{3\gamma} q^{(10^3-1)k_1}.$$

Thus, q is close to 1. If $q = 1$, then the distribution is a power law with $\gamma = 1$.

The above examples, to some extent, have demonstrated the universality of the general formula (4.2).

5. Verification by real-world system examples

We have tested and verified a large number of real datasets available on the Internet, and observed that the lengths of the variable (e.g., vertex-degree) sequences in the studied systems (networks) are all less than $(\log N)^3$, that is, $\varepsilon < 3$ in formula (4.2). It means that l is very small as compared with the network size N. In fact it is easy to see that $(\log N)^\varepsilon/N \to 0$ for any $\varepsilon < 3$ as $N \to +\infty$. Table I shows some typical examples, which are scale-free (with power-laws), exponential, or more general types of systems (networks) in various areas, including scientific collaboration networks (CN), transportation networks (TP), software package networks (SW), biological networks (Bio), social networks (SC), and so on.

According to the City Population Statistics of the Sixth National Survey conducted by the Census of China [7], the city of largest is Chongqing, with population 28846200, which ranked 1; the city of minimum size is Ngari, with population 95500, which ranked 339. For this system, $a = 339$, $b = 302.05$, so that by (4.11) the distribution is a power law with $q \approx 1$ and $\gamma \approx 1.02$.

6. Conclusion

We have shown earlier that when the vertex-degree sequence of a scale-free network of large size N follows a power-law distribution with exponent $\gamma > 1$, the length l of the vertex-degree sequence $k_1 < k_2 < \ldots < k_l$ is of order $\log N$. In this paper, we furthermore show that for many real-world systems(networks), the lengths of their variable(vertex-degree) sequences are at most of order $(\log N)^\varepsilon$, where ε is a very small constant as compared to the system (network) size N. We also extend the results to more general distributions and verify that the same conclusion holds for many real-world systems(networks). Our results may find new and useful applications in various fields in the future.

Acknowledgements

This Project is supported by the National Nature Science Foundation of China (Nos. 61170313, 61103037, 61370003).

References

[1] R. Albert and A.-L.Barabsi, *Statistical mechanics of complex networks,* Rev. Mod. Phys., 74(2002)(1), 47–91.

[2] A.-L. Barabsi and R. Albert, *Emergence of scaling in random networks,* Science, 286(1999), 509–512.

[3] R. F. I. Cancho and R. V. Sole, *The small world of human language,* Proceedings: Biological Sciences, 268(2001)(1482), 2261–2265.

[4] A. Clauset, C. R. Shalizi and M. E. J. Newman, *Power-law distributions in empirical data,* SIAM Review, 51(2009)(4), 661C-703.

[5] F. Comellas, G. Fertin and A. Raspaud, *Recursive graphs with small-world scale-free properties,* Phys. Rev. E, 69(2004)(3), 037104.

[6] H. Ebel, L. I. Mielsch and S. Bornholdt, *Scale-free topology of e-mail networks,* Phys. Rev. E, 66(2002)(3), 1–4.

[7] http://www.stats.gov.cn/tjsj/tjgb/rkpcgb/dfrkpcgb/
Appendix: http://www.ee.cityu.edu.hk/ gchen/pdf/Appendix.pdf

[8] R. A. Horn and C. R. Johnson, *Matrix Analysis,* Cambridge University Press, 1985.

[9] M. A. Jovanovic, F. S. Annexstein and K. A. Berman, *Scalability Issues in Large Peer-to-peer Networks–A Case Study of Gnutella,* Tech. Rep., University of Cincinnati, 2001.

[10] M. Newman, *Scientific collaboration networks.I.Network construction and fundamental results,* Phys. Rev. E, 64(2001)(1), 2001, 016131.

[11] M. Newman, *The structure and function of complex networks,* SIAM Rev., 45(2003)(2), 167–256.

[12] M. Newman, *Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality,* Phys. Rev. E, 64(2001)(1), 016132.

[13] M. Newman, S. Strogatz and D. Watts, *Random graphs with arbitrary degree distributions and their applications,* Phys. Rev. E, 64(2001)(2), 026118.

[14] R. Pastor-Satorras and A. Vespignani, *Epidemic spreading in scale-free networks,* Phys. Rev. Lett., 86(2001)(14), 3200–3203.

[15] S. H. Strogatz, *Exploring complex networks,* Nature, 410(2001), 268–276.

[16] D. J. Watts and S. H. Strogatz, *Collective dynamics of small-world networks,* Nature, 393(1998), 440–442.

[17] W. J. Xiao, Y. Liu and G. R. Chen, *Characterizingvertex-degree sequences in scale-free networks,* Physica A, 404(2014), 291–295.

[18] W. J. Xiao, W. D. Chen and B. Parhami, *On necessary conditions for scale-freedom in complex networks,with applications to computer communication systems,* Int. J. of Systems Science, 42(2011)(6), 951–958.

[19] W. J. Xiao, S. Z. Jiang and G. R. Chen, *A small-world model of scale-free networks: Features and verifications,* Applied Mechanics and Materials, 50-51(2011), 166–170.

[20] V. M. Yakovenko and J. B. Rosser, *Colloquium: Statistical mechanics of money, wealth, and income,* Reviews of Modern Physics, 2009, arXiv:0905.1518.