Connections Between Several Distributions of Scale-free Networks

Xiaomin Wang, Bing Yao and Jin Xu

1 School of Electronics Engineering and Computer Science, Peking University, Beijing, 100871, China
2 College of Mathematics and Statistics, Northwest Normal University, Lanzhou, 730070, China
3 School of Electronics and Information Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China

PACS 89.75.Da – 89.75.Da
PACS 02.10.Ox – 02.10.Ox
PACS 89.75.Fb – 89.75.Fb

Abstract – In our empirical works, we find that there exists the equivalency of the cumulative degree distribution and edge-cumulative distribution. Furthermore, we employ three network models of the recursive graphs, Sierpinski networks and Apollonian networks to verify our conjecture: Both the cumulative degree distribution and the edge-cumulative distribution are equivalent to each other in deterministic network models.

Keywords: scale-free network; recursive graph; Sierpinski network; Apollonian network

Introduction. – The study of complex networks is getting intensive. There are many scholars focusing on the formulation of complex network models. Complex systems consist of many constituents such as individuals, substrates, and companies in social, biological, and economic systems, respectively, showing cooperative phenomena between constituents through diverse interactions and adaptations to the pattern they create [1]. Watts and Strogatz [2] analyzed a model for disordered networks and showed that they behave like a small-world model even for relatively small values of a disorder parameter. The authors in literature [3] pointed out that algorithm performance is analyzed theoretically for parallelism and errors, and numerical simulation results are presented to assess the efficiency and effectiveness of distributed multiplicative Schwarz routing in achieving load balancing.

In [4], the authors discussed an algorithm for finding the spanning trees with maximal leaves. The literature [5] introduced a network evolution process motivated by the network of citations in the scientific literature. In each iteration of the process a vertex is born and directed links are created from new vertex to a set of target vertices already in the network. To better understand the mechanism of power law formation in real world networks, explore and analyze the underlying mechanism based on vertex degree sequences of such networks [6]. In [7], the authors proposed a simple maximum entropy process which provides the best representation of what are typical properties of scale-free networks, and furnishes a standard against which real and algorithmically generated networks can be compared.

Power law distributions occur in many situations of scientific interest and have significant consequences for our understanding of natural and man-made phenomena. Unfortunately, the detection and characterization of power laws is complicated by the large fluctuations that occur in the tail of the distribution the part of the distribution representing large but rare events and by the difficulty of identifying the range over which power law behavior retains. Commonly used methods for analyzing power law data, such as least-squares fitting, can produce substantially inaccurate estimates of parameters for power law distributions, and even in cases where such methods return accurate answers they are still unsatisfactory because they give no indication of whether the data obey a power law at all [8]. Specifically, for one such site, the news aggregator, we show how a stochastic model distinguishes the effect of the increased visibility due to the network from how interested users are in the content. We find a wide range of interest, distinguishing stories primarily of interest to users in the network (nice interests) from those of more general interest to the user community. This distinction is useful for predicting a story’s eventual popularity from users early reactions to the story [9].
In this paper, we discover that there exists the equivalence of the cumulative degree distribution and edge-cumulative distribution. Furthermore, we list three classical and famous network models of the recursive graphs, Sierpinski networks and Apollonian networks to verify our conjecture: Both the cumulative degree distribution and the edge-cumulative distribution are equivalent to each other in deterministic network models. Finally, we propose some directions for further research in order to provide guidance in our daily life.

Cumulative distributions. – The relationship between cumulative degree distribution and edge-cumulative distribution is indistinct and there is no way to know the connection. Currently, it exists two approaches to calculate cumulative degree distribution and edge-cumulative distribution. The cumulative degree distribution is defined as \( \Delta_N^c(k) = \sum_{k'=k}^{\infty} \frac{N(k',t)}{N(t)} \sim k^{1-\gamma} \) in [10], where \( N(t) \) denotes the total number of vertices at time step \( t \), \( N(k',t) \) represents the number of vertices which the degree greater than \( k \) at time step \( t \). Some literatures are drew support from this statistical approach [11–13]. The edge-cumulative distribution is defined as \( P_{ecum}(k) = \sum_{k'=k}^{\infty} \frac{E(k',t)}{E(t)} \sim k^{1-\gamma} \) in [14], where \( E(t) \) denotes the total number of edges at time step \( t \), \( E(k',t) \) stands for the number of edges which the degree greater than \( k \) at time step \( t \). The distinction between cumulative distribution and edge-cumulative distribution is that the denominator of recursive graphs are drew support from this statistical approach [11–13].

### Algorithm 1 N-algorithm

**Initialization.** \( K(q,0) \) is a complete graph \( K_q \) of \( q \) vertices.

**Operation.** Add a new vertex for each of its existing subgraphs isomorphic to a \( q \)-clique and join it with all vertices of the subgraph. For example, at \( t = 1 \), add a new vertex and join it with all vertices of \( K(q,0) \), which forms a \((q+1)-\)clique.

**Iteration.** \( K(q,t) \) is obtained from \( K(q,t-1) \) by adding for each of its existing subgraphs isomorphic to a \( q \)-clique a new vertex in order to form a \((q+1)-\)clique.

At time step \( t \), the number of vertices of degree \( d = q, 2q, q^2 + 2q, \ldots, q + \sum_{j=1}^{t-1} q^j \) is \( n_d(t) = (q+1)^{t-1} - (q+1)^{t-2} - (q+1)^{t-3} - \ldots - q + 1 \). We can find that the degree spectrum of the recursive graph \( K(q,t) \) is discrete. Other values of degree are absent. For the large \( K(q,t) \), we let the vertices of degree \( k \) was the entered network at time step \( t \), then we obtain \( k = q \frac{\ln(q+1) + \ln(q-1)}{\ln(q)} + 1 \), Cornells et.al computed cumulative degree distribution \( P_{cum}^N(k) \) of the recursive graph \( K(q,t) \) as follows:

\[
P_{cum}^N(k) = \frac{1}{n_v(t)} \sum_{i=0}^{\tau} \Delta_v^N(i) = \frac{(q+1)^{-\tau} - 1 + q^2}{(q+1)^{t-1} - 1 + q^2}
\]

Plugging

\[
\tau - t = \frac{\ln q}{\ln(q+1)} \left[ \ln \left( \frac{k}{q} - 1 \right) (q-1) + 1 \right]
\]

into 2, we obtain

\[
P_{cum}^N(k) \approx k^{-\frac{\ln(q+1)}{\ln q}} = k^{1-\gamma_N}
\]

where \( \gamma_N \approx 1 + \frac{\ln(q+1)}{\ln q} \). Next, we compute edge-cumulative distribution \( P_{ecum}^N(k) \) of recursive graph.
\[ K(q, t) \]

\[ P_{\text{cum}}^N(k) = \frac{1}{n_v(t)} \sum_{i=0}^{\tau} \Delta_N^\gamma(i) = \frac{q(q+1)+q+1)^\gamma - 1}{(q+1)\gamma - 1} \]

By the above and (1), we then obtain \( P_{\text{cum}}^N(k) \propto k^{-\frac{\ln q + \ln t}{\ln(q+1)}} \) with \( \gamma_N \approx 1 + \frac{\ln(q+1)}{\ln q} \). Furthermore, it is not hard to see

\[
|P_{\text{cum}}^N(k) - P_{\text{cum}}^N(k)| \leq \frac{6}{(q+1)^{\gamma + 1}} < \varepsilon
\]

as \( t > \frac{\ln 6 - \ln \varepsilon}{\ln (q+1)} \), with respect to \( t > \tau > 0 \) and any real number \( \varepsilon > 0 \). Based on the above fact, both \( P_{\text{cum}}^N(k) \) and \( P_{\text{cum}}^N(k) \) are mutually equivalent.

By observing two distributions \( P_{\text{cum}}^N(k) \) and \( P_{\text{cum}}^N(k) \), we can find that the accurate values \( P_{\text{cum}}^N(k) \) and \( P_{\text{cum}}^N(k) \) possess prodigious distribution, while both of them approximate to \( k^{-1-\gamma N} \), with \( \gamma_N = 1 + \frac{\ln(q+1)}{\ln q} \). Also, notice that when \( t \) gets in large, the maximal degree of a vertex is roughly \( q^t \sim (n_v(t))^\frac{\ln q}{\ln(q+1)} \). Besides, \( P_{\text{cum}}^N(k) \) and \( P_{\text{cum}}^N(k) \) obey the same power law distribution. It indicates that \( P_{\text{cum}}^N(k) \) and \( P_{\text{cum}}^N(k) \) are the reason which declares the recursive graph \( K(q, t) \) is scale-free.

**Sierpinski network model.**

We use an algorithm to build up the Sierpinski network model, it was denoted as \( S(t) \) and the algorithm is named as S-algorithm. \( S(t) \) was exactly and particularly introduced in [12].

Let \( \Delta_S^\gamma(t) \) be the numbers of vertices and edges of \( S(t) \) created at time step \( t \). According to construction of Sierpinski network, it is easily to calculate the values of \( \Delta_S^\gamma(t) \), \( \Delta_S^\gamma(t) \), in other words, \( \Delta_S^\gamma(t) = 3 \cdot 6^{t-1}, \Delta_S^\gamma(t) = 9 \cdot 6^t \). Let \( s_v(t), s_e(t) \) be the total numbers of vertices and edges of the network \( S(t) \), respectively. It is easily to calculate the size and order of the Sierpinski network model \( S(t) \)

\[ s_v(t) = \sum_{j=0}^{t} \Delta_S^\gamma(j) = \frac{3 \cdot 6^t + 12}{5} \]

\[ s_e(t) = \sum_{j=0}^{t} \Delta_S^\gamma(j) = \frac{9 \cdot 6^t + 6}{5} \]

At time \( t \), the degree spectrum of the Sierpinski network model \( S(t) \) is as: \( n_d(t) = 3 \cdot 6^{t-1}, 3 \cdot 6^{t-2}, 3 \cdot 6^{t-3}, \ldots, 6 \) when \( d = 4, 3^2 + 1, 3^3 + 1, \ldots, 3^t + 1 \). We can observe that the degree spectrum of the network model is discrete. Other values of degree are absent. It is easily to obtain cumulative degree distribution of Sierpinski network model is given by

\[ P_{\text{cum}}^S(k) = \frac{1}{s_v(t)} \sum_{i=0}^{r} \Delta_S^\gamma(i) = \frac{3 \cdot 6^t + 12}{3 \cdot 6^t + 12} = 6^{t-1} \]

Plugging for \( t = t + 1 - \ln(k - 1)/\ln 3 \), we can get that

\[ P_{\text{cum}}^S(k) \propto 6 \cdot (k - 1)^{-[1+\ln2/\ln3]} = 6 \cdot k^{1-\gamma_S} \]

with \( \gamma_S \approx 2 + (\ln 2/\ln 3) \).

Now we calculate edge-cumulative distribution is given by

\[ P_{\text{cum}}^S(k) = \frac{1}{s_e(t)} \sum_{i=0}^{r} \Delta_S^\gamma(i) = \frac{9 \cdot 6^t + 6}{9 \cdot 6^t + 6} = 6^{t-1} \]

Similarly, we can obtain that

\[ P_{\text{cum}}^S(k) \propto 6 \cdot (k - 1)^{-[1+\ln2/\ln3]} = 6 \cdot k^{1-\gamma_S} \]

with \( \gamma_S \approx 2 + (\ln 2/\ln 3) \). Notice that

\[
|P_{\text{cum}}^S(k) - P_{\text{cum}}^S(k)| \\
= \frac{1}{s_v(t)} \sum_{i=0}^{r} \Delta_S^\gamma(i) - \frac{1}{s_e(t)} \sum_{i=0}^{r} \Delta_S^\gamma(i) \\
= \frac{3 \cdot 6^t + 12}{3 \cdot 6^t + 12} - \frac{9 \cdot 6^t + 6}{9 \cdot 6^t + 6} \\
= \frac{90(6^t - 6^2)}{(3 \cdot 6^t + 12)(9 \cdot 6^t + 6)} \\
\leq \frac{180 \cdot 6^t}{27 \cdot 6^2} = \frac{20}{3} \cdot 6^t < \frac{8}{6^t} < \varepsilon
\]
as \( t \) is large enough namely, \( t > \frac{\ln s - \ln \varepsilon}{\ln \sigma} \), with respect to \( t > \tau > 0 \) and any real number \( \varepsilon > 0 \). So, we declare that \( P_{\text{cum}}^S(k) \) and \( P_{\text{cum}}^S(k) \) are equivalent to each other.

Also, notice that when \( t \) gets enough large, \( P_{\text{cum}}^S(k) \) and \( P_{\text{cum}}^S(k) \) decrease to a power law distribution. By this means, we can discover that cumulative degree distribution and edge-cumulative distribution are both about to \( 6 \cdot (k - 1)^{1 + \left[(\ln 2/\ln 3)\right]} \). It indicates that \( P_{\text{cum}}^S(k) \) and \( P_{\text{cum}}^S(k) \) import the same coefficient and power law exponent \( \gamma_S \), wtih \( \gamma_S \approx 2 + (\ln 2/\ln 3) \). The edge-cumulative distribution illustrate the Sierpinski network model is scale-free from another aspect. It indicates that \( S(t) \) is scale-free network.

**Apollonian network.** Finally, we use an algorithm to rewrite the process of construction of high dimensional Apollonian networks. It have been discussed in literature [13] in detailed. In the iterative process for the construction of high-dimensional Apollonian network at each iteration, for each new hypersphere added, \( d + 1 \) new interstices are created in the associated Apollonian packing which will be filled in the next iteration. When building networks, we can say in equivalent words that for each new vertex added, \( d + 1 \) new d-simplices are generated in the network, into which vertices will be inserted in the next iteration. We denote the network model as \( A(t) \). The A-algorithm can be listed as follows

**Algorithm 3 A-algorithm**

**Initialization.** \( A(d,0) \) is a complete graph \( K_{d+1} \) of \( d + 1 \) vertices.

**Operation.** Add a vertex for each of its existing subgraphs isomorphic to a \((d + 1)\)-clique and join it to all vertices of this subgraphs. Such as, at \( t = 1 \), add one new vertex and \( d + 1 \) new vertices of this subgraphs. Such as, at \( t = t \), a \( d \)-simplex in order to form a \((d + 1)\)-clique in the network model as \( A(t) \).

**Iteration.** \( A(d, t) \) is obtained from \( A(d−1, t) \) by adding for each of its existing subgraphs isomorphic to \((d + 1)\)-clique in order to form a \((d + 2)\)-clique.

Let \( \Delta^A(t) \), \( \Delta^A(t) \) be the numbers of vertices and edges created at time step \( t \). According to the A-algorithm, it is easily to calculate the values of \( \Delta^A(t), \Delta^A(t) \), namely, \( \Delta^A(t) = (d + 1)^{t−1} \), \( \Delta^A(t) = (d + 1)^t \). The notation \( a_v(t), a_e(t) \) represent the total number of vertices and edges of Apollonian network model \( A(t) \). Then we can easily see that at time step \( t \), the total number of vertices and edges of the Apollonian network \( A(t) \) can be computed as

\[
a_v(t) = \sum_{j=0}^{t} \Delta^A_v(j) = d + 1 + \sum_{j=1}^{t} (d + 1)^{j−1} = d + 1 + \frac{(d + 1)^t - 1}{d} - 1
\]

\[
a_e(t) = \sum_{j=0}^{t} \Delta^A_e(j) = \frac{d(d + 1)}{2} + \sum_{j=1}^{t} (d + 1)^j = \frac{d(d + 1)}{2} + \frac{(d + 1)^{t+1} - d - 1}{d}
\]

The degree spectrum of the network model is discrete since \( n_d(t) = 1, d + 1, (d + 1)^2, \ldots, (d + 1)^{t−1} \) with \( d = (d+1)(\sum_{j=0}^{t-2} d^2+1), (d+1)(\sum_{j=0}^{t-3} d^3+1), (d+1)(\sum_{j=0}^{t-3} d^3+1), \ldots, d + 1 \). Other values of degree are absent. According to the degree spectrum of the Apollonian network model, we can compute the cumulative degree distribution of Apollonian network model is given by

\[
P^A_{\text{cum}}(k) = \frac{1}{a_v(t)} \sum_{i=0}^{\tau} \Delta^A_v(i) = \frac{d(d + 1) + (d + 1)^\tau - 1}{d(d + 1) + (d + 1)^\tau - 1} = \frac{(d + 1)^\tau - t}{d + 1}
\]

for large \( t \), plugging \( \tau = t + 1 - \ln(d + 1)/\ln d \) into the above form

\[
P^A_{\text{cum}}(k) \propto k^{-\frac{\ln(d+1)}{\ln d}} = k^{1-\gamma_A}
\]

It means that the apollonian network is scale-free. Because the degree distribution follow the scaling law, with \( \gamma_A \approx 1 - \frac{\ln(d+1)}{\ln d} \).

Now we introduce another statistic method, we calculate the edge cumulative degree distribution is given by

\[
P^A_{\text{cum}}(k) = \frac{1}{a_e(t)} \sum_{i=0}^{\tau} \Delta^A_e(i) = \frac{(d + 1)d^2 + 2[(d + 1)^{t+1} - d - 1]}{(d + 1)d^2 + 2[(d + 1)^{t+1} - d - 1]} = \frac{(d + 1)^{t+1} - d - 1}{(d + 1)^{t+1} - d - 1}
\]

As \( t \) becomes large, plugging \( \tau = t + 1 - \ln(d + 1)/\ln d \) into the above form, we can obtain

\[
P^A_{\text{cum}}(k) \propto k^{-\frac{\ln(d+1)}{\ln d}} = k^{1-\gamma_A}
\]

where \( \gamma_A \approx 1 + \frac{\ln(d+1)}{\ln d} \). On the basis of above analysis and sharply computation, we verify that both \( P^A_{\text{cum}}(k) \) and
Connections Between Several Distributions of Scale-free Networks

$P_{\text{cum}}^A(k)$ are mutually equivalent in the following way

\[
\left| P_{\text{cum}}^A(k) - P_{\text{cum}}^A \right| = \frac{1}{a(v(t))} \sum_{i=0}^{\tau} \Delta^i - \frac{1}{a(v(t))} \sum_{i=0}^{\tau} \Delta^i \leq \frac{d(d+1)+(d+1)^{\tau}-1}{d(d+1)+(d+1)^{\tau}-1} - \frac{(d+1)^2 + 2[(d+1)^{\tau+1} - d-1]}{(d+1)^2 + 2[(d+1)^{\tau+1} - d-1]} \leq \frac{4d(d+1)^{\tau+2} + 2 + (2d^2 + 4d)(d+1)^{\tau+1}}{(d+1)^{2\tau+1}} \leq \frac{8d}{(d+1)^{\tau+1} - 1} < \varepsilon
\]

when $t$ gets large in other words, $t > 1 + \ln \frac{8d - \ln \varepsilon}{\ln (d+1)}$, with respect to $t > \tau > 0$ and arbitrary real number $\varepsilon > 0$.

The accurate value of the cumulative degree distribution have relation with the number of vertices, while the edge-cumulative distribution is connected with the number of edges. Even though we know the number of vertices and the number of edges always different. But the ratio of the number of vertices to the number of vertices have relation with the number of vertices, while the edge-cumulative distribution approximate to the same power law distribution when analyse on different network models to verify our conjecture. We analyse the results above for several classical and famous network models to verify our conjecture. We employ several classical and famous network models to verify our conjecture. We analyse connections between $P_{\text{cum}}(k)$ and $P_{\text{cum}}(k)$, the former represents the ratio of the number of vertices, the latter stands for the ratio the number of edges, while both of them obey the same power law distribution when analyse the same deterministic network model. We didn’t know whether exists other approaches for computing $P_{\text{cum}}(k)$ and $P_{\text{cum}}(k)$, it implies that the high dimensional Apollonian network is scale-free.

Conclusion. – We employ several classical and famous network models to verify our conjecture. We analyse connections between $P_{\text{cum}}(k)$ and $P_{\text{cum}}(k)$, the former represents the ratio of the number of vertices, the latter stands for the ratio the number of edges, while both of them obey the same power law distribution when analyse the same deterministic network model. We didn’t know whether exists other approaches for computing $P_{\text{cum}}(k)$ and $P_{\text{cum}}(k)$, it implies that the high dimensional Apollonian network is scale-free.

The author, Jin Xu, thanks the National Natural Science Foundation of China Grant 61300915. The author, Bing Yao, was supported by the National Natural Science Foundation of China under grants No. 61163054 and No. 61363060.

REFERENCES

[1] Albert-László Barabási and Reka Albert (Editor), Emergence of scaling in random networks, Vol. 286 (Science) 1999, p. 509-512.
[2] Watts DJ, Strogatz SH (Editor), Collective dynamic of small-world networks, Vol. 393 (Nature) 1998, p. 440-442.
[3] Ren-Song Ko (Editor), A load-balancing routing algorithm for wireless sensor networks based on domain decomposition, Vol. 9 (Ad Hoc Networks) 2015, p. 3063-3083.
[4] Xiaomin Wang, Xiyang Zhao, Bing Yao, Ming Yao, Xiang-en Chen (Editor), Algorithms for maximum leaf spanning trees of Edge-growing network models in information system. Engineering solutions for industrial production., Vol. 775, 2015, p. 431-435.
[5] E.R. Colman, G.J. Rodgers (Editor), Complex scale-free networks with tunable power-law exponent and clustering, Vol. 392 (Physical A) 2013, p. 5501-5510.
[6] Wenyun Xiao, Yanxia Liu, Guanrong Chen (Editor), Characterizing vertex-degree sequences in scale-free networks, Vol. 404 (Physical A) 2014, p. 291-295.
[7] Linjun Zhang, Michael Small, Kevin Judd (Editor), Exactly scale-free scale-free networks, Vol. 433 (Physical A) 2015, p. 182-197.
[8] Aaron Clauset, Cosma Rohilla Shalizi M.E.J Newman (Editor), Power-law distribution in empirical data, Vol. 51 (Society for Industrial and Applied Mathematics) 2009, p. 661-703.
[9] Kristina Lerman Tad Hogg (Editor), Using Stochastic Models to Describe and Predict Social Dynamics of Web Users, Vol. 34 (ACM Transactions on Intelligent Systems and Technology) 2012, p. 62.
[10] S.N. Dorogovtsev, A.V. Goltsev, J.F.F. Mendes (Editor), Pseudo-fractal scale-free web, Vol. 62 (Physical review) 2002, p. 066122-066125.
[11] Francesc Comellas (Editor), Recursive graphs with small-world scale-free properties, Vol. 69 (Physical Review Letters) 2004, p. 037101-037104.
[12] Zhongzhi Zhang, Shuigeng Zhou, Lijun Fan, Jihong Guan, Yichao Zhang (Editor), Maximal planar scale-free Sierpinski networks with small-world effect and power-law strength-degree correlation, Vol. 380 (Journal of Physics. A, Mathematical and Theoretical) 2007, p. 621-626.
[13] Zhongzhi Zhang, Francesc Comellas, Guillaume Fertin, Lili Rong (Editor), High dimensional apollonian networks, Vol. 398 (Journal of Physics. A, Mathematical and Theoretical) 2006, p. 063120-063120.
[14] Xia Liu, Bing Yao, Wanjia Zhang, Xiang-en Chen, Xinsheng Liu, Ming Yao (Editor), Uniformly bound-growing network models and their spanning trees. The 2014 International Conference on Information and Communica-
tion Technologies (ICT2014) will be held 16th-18th May 2014 in anjing, China. 2014, p. 35-38.