Topological properties of scale-free networks driven by a graph Hamiltonian

2011 EPL 93 38001

(http://iopscience.iop.org/0295-5075/93/3/38001)

A Letters Journal Exploring the Frontiers of Physics

February 2011

EPL, **93** (2011) 38001 www.epljournal.org
doi: 10.1209/0295-5075/93/38001

Topological properties of scale-free networks driven by a graph Hamiltonian

Soon-Hyung Yook[1(a)] and Juyong Park[2(b)]

[1] *Department of Physics and Research Institute for Basic Sciences, Kyung Hee University - Seoul 130-701, Korea*
[2] *Department of Physics, Kyung Hee University - Seoul 130-701, Korea*

received 27 July 2010; accepted in final form 18 January 2011
published online 14 February 2011

PACS 89.75.Da – Systems obeying scaling laws
PACS 89.75.Hc – Networks and genealogical trees
PACS 89.75.Fb – Structures and organization in complex systems

Abstract – We numerically study an equilibrium network model, generated via a graph Hamiltonian, that exhibits a power-law degree distribution $P(k) \sim k^{-\gamma}$ with $\gamma \approx 1.5$ in the thermodynamic limit. While the degree exponent, γ, is equal to the one found in the Merging-Creation (MC) model, we find that the detailed topological properties such as degree-degree correlation and the modular structure differ significantly. We discuss the possible origin of these differences.

Recently, the ubiquity of complex networks has garnered much attention. A complex network is one where nodes (vertices) are connected by links (edges) in a nontrivial fashion. Examples are technological networks such as the Internet and the World-Wide Web (WWW), biological networks such as the protein-protein interaction network and metabolic networks, and social networks such as the e-mail network [1–6], etc. Many such real networks have completely different topological properties from traditional models of networks, most notably the canonical random network of Erdös and Rényi [1,6,7]. One of the most striking differences is that these real networks exhibit a scale-free (SF) or power-law degree distribution $P(k) \sim k^{-\gamma}$, unlike the ER random network [8]. This property indicates a strong heterogeneity of connectivity in which only a few nodes have many connections to other nodes while the majority have a few. Empirical studies have revealed that γ varies from network to network, but is usually in the range $1.0 < \gamma \leqslant 3$ [1–6,9].

So far, much effort has been made in finding the underlying mechanisms that give rise to SF networks with $2 < \gamma \leqslant 3$ and understanding their topology [1,2]. In comparison, as reported in ref. [9], networks with $\gamma < 2$ have received less attention despite there being plenty of examples such as the Gnutella [10], e-mail [11], coauthorship [12], ecological [13], and software networks [14]. Some authors have recently proposed models of network generation based on the Merging-Creation (MC) process. For instance, in one such model SF networks with $\gamma < 2$ are generated by mimicking the effect of the reusability of software components [9,15–17]. In the MC model, the evolution of degree k of each node can be mapped to the evolution of mass m on each site in the mass aggregation model without desorption which is known to generate a power-law mass distribution $P(m) \sim m^{-3/2}$ [18]. Therefore, in the simple MC process with ignorable edge removal rate, the steady-state degree distribution is expected to follow the same power law $P(k) \sim k^{-3/2}$ [15].

It is important to note that the MC model as proposed by Seyed-Allaei *et al.* [9] is a non-equilibrium model. The model is composed of two processes at each step, *i.e.* merging of two randomly selected nodes and creation of a new node. In the merging process, self-connections and multiple connections are not allowed: if the two nodes to be merged are connected the link is removed, and when they have common neighbors the multiple links are replaced by a single link. In the creation process the new node is connected to a pre-existing, randomly chosen node. The creation rate is controlled so that it balances the merging process, and thus it is of interest to compare it with an equilibrium model where edges are rewired, keeping the numbers of nodes and edges constant. Since the most prominent feature of the MC model is the degree distribution $P(k) \sim k^{-3/2}$, we introduce a graph Hamiltonian that leads to the same power-law degree

(a) E-mail: syook@khu.ac.kr
(b) E-mail: perturbation@gmail.com

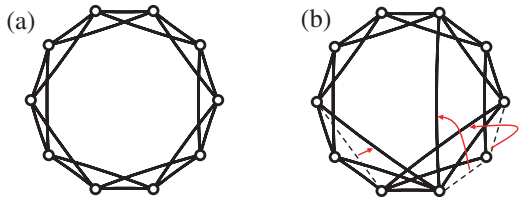

Fig. 1: (Colour on-line) Example of the (a) initial topology and (b) rewiring process with $N = 10$ and $\langle k \rangle = 4$. The dashed links in (b) are rewired. The new topology is accepted only if $\Delta H < 0$ at $T = 0$.

distribution at ground state, and compare the topological properties such as clustering coefficients and the degree-degree correlations of the two models.

Our equilibrium network is as follows. It is composed of N nodes connected by undirected edges. The corresponding adjacency matrix $\mathbf{a}$, where its element a_{ij} is equal to 1 if nodes i and j are connected, and 0 otherwise, is symmetric. The connectivity or degree k of node i is defined as $k_i = \sum_j a_{ij}$, and the total number of links is $L = \sum_{i<j} a_{ij} = \mathrm{Tr}\ \mathbf{a}^2/2$. The average degree is therefore $\langle k \rangle = 2L/N$. Then we can define the partition function as

$$Z(T, N, L) = \sum_{\Omega(\mathbf{a})} \delta \left(L - \frac{\mathrm{Tr}\ \mathbf{a}^2}{2} \right) \exp\left[-\beta \mathcal{H}(\mathbf{a}) \right], \qquad (1)$$

for any network Hamiltonian $\mathcal{H}(\mathbf{a})$ over all possible permutations of adjacency matrix $\Omega(\mathbf{a})$. T represents the temperature and $\beta = 1/k_B T$ is the inverse temperature with Boltzmann constant k_B. This follows the ensemble approach that is often used as a methodology of generating networks [19–23]. In this report, we study specifically the network ensemble defined via a graph Hamiltonian

$$\mathcal{H}(\mathbf{a}) = -J \sum_{(ij)} k_i k_j = -\frac{J}{2} \sum_{i,j} a_{ij} k_i k_j, \qquad (2)$$

where (ij) means connected pairs.

Imagine that we are tracking the changes in $\mathcal{H}(\mathbf{a})$ as we follow an MC process. One can easily verify that $\mathcal{H}(\mathbf{a})$ always decreases in the MC process when L stays constant. Therefore, it is natural to ask if the MC model is equivalent to, or share common topological properties with, the equilibrium network defined by eq. (2) obtained via rewiring.

In the following numerical simulations, for convenience we focus only on networks generated using the Hamiltonian at $T = 0$, with parameters fixed at $J = 1$ and $k_B = 1$. The simulation proceeds as follows: We start from a regular ring of N nodes, each connected to its $\langle k \rangle$ nearest neighbors (see fig. 1(a)), and at each step we select a link (*i.e.* a connected node pair) and an unconnected node pair, and decide whether to rewire the link. The link is rewired when the energy is lower at the new potential configuration, *i.e.* if $\Delta H = H_{rewired} - H_{current} < 0$ (see fig. 1). Otherwise, we keep the current topology. Note

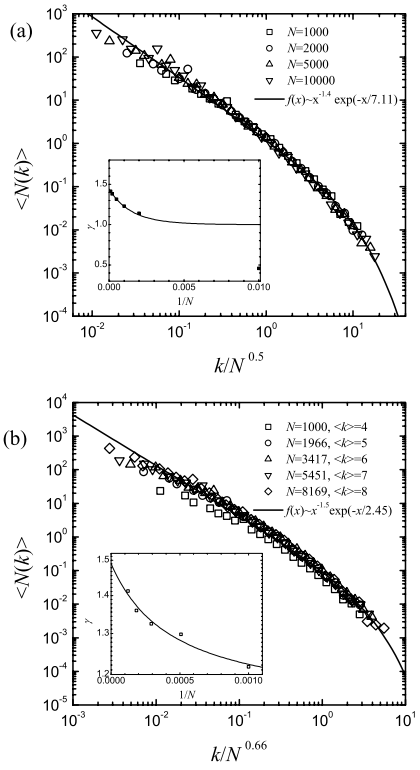

Fig. 2: Collapsed plot of the degree distributions for various network sizes, (a) $N = 1000$, 2000, 5000, and 10000, with fixed $\langle k \rangle = 4$, and (b) N increases according to the relation $\langle k \rangle \propto N^{\xi}$ with $\xi = 1/3$. From the insets we obtain (a) $\gamma = 1.4 \pm 0.3$ and (b) $\gamma = 1.5 \pm 0.3$.

that, at a low temperature, when $J > 0$ low-degree nodes would prefer to be connected to high-degree nodes for lower energy. Thus, as we will see later, this leads to a disassortative mixing of degrees [24]. On the other hand, if $J < 0$, depending on the initial condition, the initial network would not be rewired at $T = 0$ because the rewiring always increases the energy, unlike the model suggested in ref. [25].

In growing SF networks with $\gamma < 2$, the average degree $\langle k \rangle$ is related to N in a non-trivial way, *i.e.* $\langle k \rangle \sim N^{\xi}$. Then $\xi > 0$ implies that the total number of links grows faster than the number of nodes. This also leads to the structural cut-off k_c that diverges with N as [9]

$$k_c \sim N^{(1+\xi)/2} \sim N^{\sigma}, \qquad (3)$$

where ξ is known to satisfy

$$\xi = \frac{(2-\gamma)}{\gamma} \quad \text{and} \quad \sigma = \frac{1+\xi}{2}. \qquad (4)$$

In our equilibrium model, by contrast, $\langle k \rangle$ is in the initial condition, and thus k_c can be adjusted to satisfy any ξ in $\langle k \rangle \sim N^{\xi}$. Therefore the first relation in eq. (4) is not always valid in our model. If we fix $\langle k \rangle$ to be a constant for all sizes of networks, then $\xi = 0$ and $k_c \sim N^{1/2}$ but still $\gamma \approx 1.5$ [26], as we see in fig. 2(a). Figure 2(a) shows $N(k)$

for various N's with fixed $\langle k\rangle = 4$. Using the best fit of the data to the function

$$N(k) \sim k^{-\gamma} \exp(-k/k_c), \tag{5}$$

we obtain $k_c(N) \sim N^{\sigma}$ with $\sigma = 1/2$ (not shown). With this exponent $\sigma = 1/2$, we find that the *N(k)*'s for various N collapse well into a single curve as shown in fig. 2(a). Then we find the thermodynamic limit ($N \to \infty$) of γ via plotting $\gamma(N)$ against $1/N$ (inset), yielding $\gamma(N \to \infty) = 1.4 \pm 0.3$. This confirms our claim that the first relation in eq. (4) is not generally valid, and only a specific value $\xi = 1/3$ can satisfy the relation when $\gamma = 1.5$. The latter point can be checked via simulation as well: in a network in which $\langle k\rangle \sim N^{1/3}$, from eqs. (3) and (4) we expect that $k_c \sim N^{2/3}$ with $\sigma = 2/3$. As expected, with this value *N(k)*'s for various N collapse well into a single curve, shown in fig. 2(b). We also find $\gamma(N \to \infty) = 1.5 \pm 0.3$ (inset). Therefore, the degree distribution $P(k)$ (or equivalently $N(k)$) of our model satisfies the same scaling ansatz given in ref. [9], regardless of the applicability of the first relation in eq. (4). For other values of $\xi(> 0)$, we obtain the same results (which is not shown).

To investigate the detailed structural properties of our equilibrium model, we have also measured the clustering coefficient and the degree-degree correlation. The clustering coefficient C is known to give us a very important clue regarding the hierarchical structure of network modules. In a network with no hierarchical modular structure C shows no dependence on k, *i.e.* $C(k) = \text{const}$. On the other hand, if the networks exhibits a well-defined hierarchical modules, the clustering coefficient is a decreasing function of k, *e.g.* $C(k) \sim k^{-\beta}$, implying that the nodes in smaller module tend to be more tightly connected to one another [27]. The degree-degree correlation is another important measure in characterizing the network structure. Depending on degree-degree correlation, networks can be classified into two groups: networks with positive (negative) correlations are said to exhibit assortative (disassortative) mixing [24]. Generally, social networks obtained via one-mode projection from affiliation relations are assortative, while technological networks are disassortative.

In fig. 3(a) we display the largest cluster of a sample network of $N = 50$ nodes generated from our equilibrium model. This example captures interesting properties of the network topology generated by equilibrium model. As expected from the definition of the equilibrium model, the hubs show a strong tendency to be connected to low-degree nodes, the characteristic feature of disassortative networks. In the ground state, the biggest hub is connected to all the other nodes which causes a star-like topology similar to the model studied in ref. [19]. Since a true star (with one central hub connected to all other degree-one nodes) has a mean degree $\langle k\rangle = 2$, the remaining links form yet smaller star-like structures also connected to larger hubs to reduce the total energy. Then a hierarchy of sorts between modules of differences emerge. Therefore,

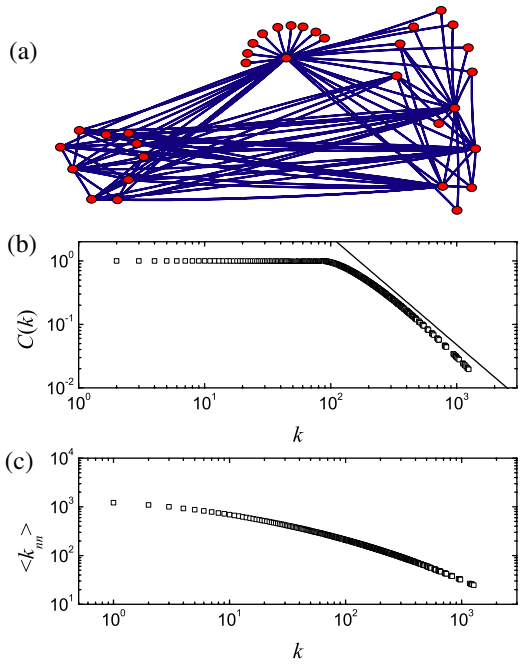

Fig. 3: (Colour on-line) (a) The largest cluster in the equilibrium network sample of $N = 50$. (b) Plot of $C(k)$ *vs.* k with $N = 10000$. For $k > 10^2$, we fit the data to $C(k) \sim k^{-\beta}$ with $\beta \approx 1.7$. (c) $\langle k_{nn}\rangle$ decreases as k increases, indicating disassortative degree mixing.

as outlined above, we can expect that $\langle k_{nn}(k)\rangle$ and $C(k)$ to decrease as k increases. First, $C(k)$ of the equilibrium model is shown in fig. 3(b). Here, we see a plateau for $k < 10^2$, indicating the absence of hierarchy in that range of k. However, for larger k ($> 10^2$), we find that $C(k)$ decreases and can be approximated by $C(k) \sim k^{-\beta}$ with $\beta \approx 1.7$ which indicates the existence of hierarchical modules. This agrees with our naïve expectation based on fig. 3(a). We also investigate the degree-degree correlation by measuring $\langle k_{nn}(k)\rangle$, the average degree of the nearest neighbors of the nodes having degree k, shown in see fig. 3(c). $\langle k_{nn}\rangle$ clearly decreases as k increases, indicating disassortative degree mixing or hub repulsion [24].

For comparison, we also measured the same quantities for the simplest MC model of ref. [9]. As shown in fig. 4(a), the network structure of MC model is visibly different from that of the equilibrium model shown in fig. 3(a), even though $P(k)$ has the same form (eq. (5)). From fig. 4(b) and (c), we find that both $C(k)$ and $\langle k_{nn}\rangle$ increase as k increases and then decrease for the largest k.

Qualitatively, this can be explained as the consequence of the evolutionary rule of the MC model. In MC model, at each time step two randomly chosen nodes are merged into a single node, and a new node created then connected to one of the existing nodes. Thus the evolutionary rule of the MC model effectively has an aging dynamics via which a node created at step t_0 is chosen to merge, on average, with another at time t with probability proportional to $\Delta t = t - t_0$. This merging process affects

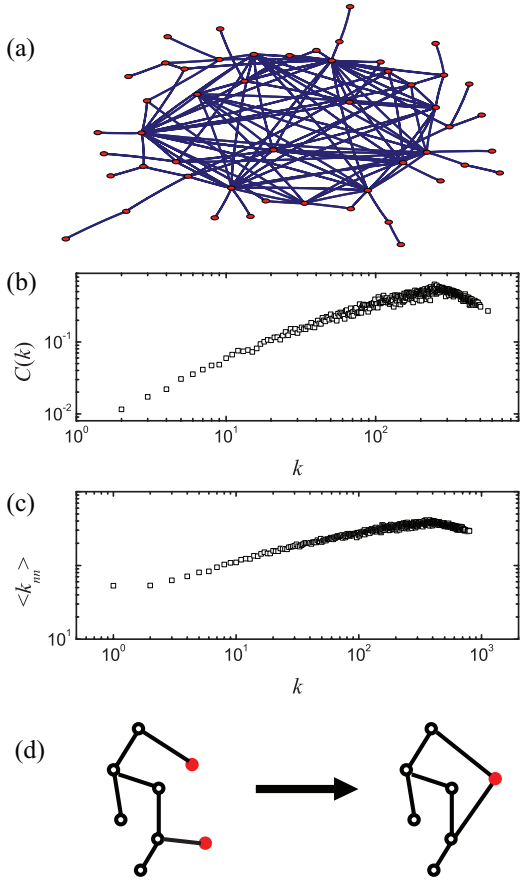

Fig. 4: (Colour on-line) (a) Sample network topology of the MC model. Plots of (b) $C(k)$ and (c) $\langle k_{nn}\rangle$ against k. In (d) we illustrate the loop creation by merging. Two filled circles (left) are selected to merge into a single node (right). This creates a loop.

the resulting topology of the networks in two crucial ways. First, older nodes naturally gain more links than younger ones, leading to the degree heterogeneity in the form of a power-law degree distribution. Moreover, since the merging process occurs between pre-existing nodes, encouraging connections between older nodes. This leads to an increasing $\langle k_{nn}(k)\rangle$ with increasing k. Second, the merging of two nodes can create loops in the network (fig. 4(d)). If we repeat the merging many times then the density of the loops can increase. As a result, the old nodes having more degrees can have larger clustering coefficient (concerning loops of length three) than the young nodes having smaller degrees. Therefore, similarly, $C(k)$ is also an increasing function of k in general.

The slight decreasing behavior $\langle k_{nn}(k)\rangle$ and $C(k)$ near k_{max}, however, indicates that for the emergence of hubs of $k \simeq k_{max}$ very young, low-degree nodes play a crucial role. Let τ be the average time it takes a newly born node with initial degree $k=1$ to be merged with another node. Since merging and creation are balanced, $\tau \propto N$. The accumulative probability that a node created at t_0 is selected to merge with a new node at time t is again proportional to Δt, which increases as k increases. As a result, old nodes having degree $k \approx k_{max}$ would have merged with $\tau\Delta t \propto N\Delta t$ newly born nodes of $k=1$ on average. Therefore, for large N, a sizable number of degrees of a hub may have originated from merging with $k=1$ nodes. The decreasing behavior of $C(k)$ and $\langle k_{nn}\rangle$ near $k \simeq k_{max}$ are the consequences of this dynamics. We have also verified the effect of young nodes with modified MC models incorporating preferential attachment [28].

We have introduced a network Hamiltonian model that generates a power-law degree distribution with $\gamma \approx 1.5$, identical to the MC model. However, we find that the detailed topological properties of the two models are very different, indicating that the degree exponent γ is merely one characteristic of a network (albeit an important one) and is woefully insufficient as a general classifier of networks[1]. We have argued that the differences in the topological properties stem from how the MC networks are generated, an evolutionary dynamics. Based on the measurements of degree-degree correlation and clustering coefficients, we expect that the modular structure found in many real-world networks such as software networks [14] can be better modeled via our equilibrium Hamiltonian model. This also raises a question regarding the possible origin of modularity in networks. The concept of modularity is central in engineering and biology where a system is often a composition of semi-independent, functionally distinct parts (modules) [29]. It is widely believed that the modularity in biological systems in particular is a result of evolution. The fact that one can create modular networks using a simple graph Hamiltonian rather than imagining a likely evolutionary process suggests the possibility of alternative methodologies for studying modularity in networks. We believe that such questions present future opportunities for further studies including deep analytical treatment.

* * *

We would like to thank Yup Kim, Sung-Chul Kwon, and Sooyeon Yoon for useful discussions. This work was supported by grants KHU-20100116 from Kyung Hee University, and KRF-20100004910 from the National Research Foundation of Korea.

REFERENCES

[1] Dorogovtsev S. and Mendes J. F. F., *Evolution of Networks: From Biological Nets to the Internet and WWW* (Oxford University Press, New York) 2003.
[2] Albert R. and Barabási A.-L., *Rev. Mod. Phys.*, **74** (2002) 47.
[3] Albert R., Jeong H. and Barabási A.-L., *Nature*, **401** (1999) 130.

[1] Furthermore, one should take caution not to use γ alone to define universality classes of networks as in studies of critical phenomena; networks in general are not critical, and thus γ cannot be viewed as a critical exponent.

[4] Yook S. H., Jeong H. and Barabási A.-L., *Proc. Natl. Acad. Sci. U.S.A.*, **99** (2002) 13382.
[5] Jeong H., Mason S., Barabási A.-L. and Oltvai Z., *Nature*, **411** (2001) 41; Yook S. H., Oltvai Z. and Barabási A.-L., *Proteomics*, **4** (2003) 928.
[6] Barabási A.-L. and Albert R., *Science*, **286** (1999) 509.
[7] Watts D. J. and Strogatz S. H., *Nature*, **393** (1998) 440.
[8] Erdös P. and Rényi A., *Publ. Math.*, **6** (1959) 290; *Publ. Math. Inst. Hung. Acad. Sci.*, **5** (1960) 17.
[9] Seyed-Allaei H., Bianconi G. and Marsili M., *Phys. Rev. E*, **73** (2006) 046113.
[10] Jovanovic M., Annexstein F. and Berman K., Technical Report, University of Cincinnati (2001), URL `http://www.ececs.uc.edu/~annexste/Papers/scalabilityissues.ps`.
[11] Ebel H., Mielsch L. I. and Bornholdt S., *Phys. Rev. E*, **66** (2002) 035103(R).
[12] Newman M. E. J., *Phys. Rev. E*, **64** (2001) 016131.
[13] Montoya J. M., Pimm S. L. and Solé R. V., *Nature*, **442** (2006) 259.
[14] Valverde S. and Solé R. V., *Europhys. Lett.*, **72** (2005) 858.
[15] Minnhagen P., Rosvall M., Sneppen K. and Trusina A., *Physica A*, **340** (2004) 725.
[16] Kim B. J., Trusina A., Minnhagen P. and Senppen K., *Eur. Phys. J. B*, **43** (2005) 369.
[17] Alava M. J. and Dorogovtsev S. N., *Phys. Rev. E*, **71** (2005) 036107.
[18] Takayasu H., *Phys. Rev. Lett.*, **63** (1989) 2563.
[19] Berg J. and Lässig M., *Phys. Rev. Lett.*, **89** (2002) 228701.
[20] Palla G., Derenyi I., Farkas I. and Vicsek T., *Phys. Rev. E*, **69** (2004) 046117.
[21] Park J. and Newman M. E. J., *Phys. Rev. E*, **70** (2004) 066117.
[22] Biely C. and Thurner S., *Phys. Rev. E*, **74** (2006) 066116.
[23] Minnhagen P., Bernhardson S. and Kim B. J., *EPL*, **78** (2007) 28004.
[24] Newman M. E. J., *Phys. Rev. E*, **67** (2003) 026126.
[25] Noh J. D., *Phys. Rev. E*, **76** (2007) 026116.
[26] Boguña M., Pastor-Satorras R. and Vespignani A., *Eur. Phys. J. B*, **38** (2004) 205.
[27] Ravasz E., Somera A. L., Mongru D. A., Oltvai Z. N. and Barabási A.-L., *Science*, **297** (2002) 1551.
[28] Yook S. H. *et al.*, in preparation.
[29] Dovrolis C. and Streelman J. T., *ACM SIGCOMM Comput. Commun. Rev.*, **40** (2010) 72.