Available online at www.sciencedirect.com
SCIENCE DIRECT®
Physica A 318 (2003) 72–79
PHYSICA A
www.elsevier.com/locate/physa

Packet transport and load distribution in scale-free network models

K.-I. Goh, B. Kahng, D. Kim*

School of Physics, Seoul National University, Seoul 151-747, South Korea

Abstract

In scale-free networks, the degree distribution follows a power law with the exponent γ. Many model networks exist which reproduce the scale-free nature of the real-world networks. In most of these models, the value of γ is continuously tunable, thus is not universal. We study a problem of data packet transport in scale-free networks and define *load* at each vertex as the accumulated total number of data packets passing through that vertex when every pair of vertices send and receive a data packet along the shortest paths. We find that the load distribution follows a power law with an exponent δ for scale-free networks. Moreover, the load exponent δ is insensitive to the details of the networks in the range $2 < \gamma \leqslant 3$. For the class of networks considered in this work, $\delta \approx 2.2(1)$. We conjecture that the load exponent is a universal quantity to characterize and classify scale-free networks.

PACS: 89.75.Hc; 05.10. − a; 89.70.+c; 89.75.Da

Keywords: Scale-free networks; Load distributions

In the network approach of complex systems, vertices of a graph represent their constituents such as individuals, substrates, and companies in social, biological, and economic systems, respectively, and edges the interactions between the two constituents connected [1–3]. The number of edges incident on a vertex is the degree (or connectivity) of the vertex and one is interested in the probability distribution of the degree, $P_D(k)$, which is measured by the fraction of the vertices whose degree is k, averaged over an appropriate ensemble. The random network of Erdös and Rényi (ER) [4] and

* Corresponding author.
E-mail address: vrooom@phya.snu.ac.kr (K.-I. Goh).

0378-4371/03/$ - see front matter
PII: S0378-4371(02)01407-3

the small world network of Watts and Strogatz (WS) [5] are exponential networks in that the degree distribution is Poissonian. However, many real-world networks show power-law behavior in the degree distributions, and are termed as scale-free (SF) [6]. Examples of SF networks include the world-wide web [7,8], the Internet [9–11], the citation network [12] and the author collaboration network [13,14] of scientific papers, the protein–protein interaction network [15,16] and the metabolic networks in biological organisms [17].

There are many models which reproduce such scale-free features. For example, Barabási and Albert (BA) [18] introduced an evolving network where the number of vertices N increases linearly with time rather than fixed, and a newly introduced vertex is connected to m already existing vertices, following the so-called preferential attachment rule. In the original BA model, the probability for the new vertex to connect to an already existing vertex is proportional to the degree k of the selected vertex. But the generalized versions [19] assign the probability proportional to $k + m(a-1)$, a (>0) being a tunable parameter. Then the degree distribution $P_D(k)$ follows a power law $P_D(k) \sim k^{-\gamma}$ with $\gamma = 2 + a$. Many other models also possess parameters with which the degree exponent can be tuned continuously. Real-world networks also show varying values for γ [3], mostly in the range $2 < \gamma \leqslant 3$.

The SF networks show the small world behavior in that the diameter d of the network, defined as the average of shortest distances between every pair of vertices, scales with the size N, the number of vertices, as $d \sim \log N$ [20]. The small-world property in SF networks results from the presence of a few vertices with high degree. In particular, the hub, the vertex whose degree is the largest, plays a dominant role in reducing the diameter of the system. The transport from one position to another is mainly carried along the shortest path(s) between them. When a data packet is sent from one vertex to another through SF networks such as Internet, it is efficient to take a road along the shortest path between the two. Then vertices with higher degrees should be heavily loaded and jammed by lots of data packets passing along the shortest paths. To prevent such Internet traffic congestions, and allow data packets to travel in a free-flow state, one has to enhance the capacity, the rate of data transmission, of each vertex to the extent that the capacity of each vertex is large enough to handle appropriately defined "load".

We define and study such a quantity, which we simply call load, to characterize the transport dynamics in SF networks. To be specific, we suppose that a data packet sent from a vertex i to j is transmitted along the shortest path between them. If there exist more than one shortest paths, the data packet would encounter one or more branching points. In this case, we assume that the data packet is divided evenly by the number of branches at each branching point as it travels. Load contribution to a vertex k from a pair $(i \to j)$ is denoted as $\ell_k^{(i\to j)}$. A simple example for defining $\ell_k^{(i\to j)}$ is depicted in Fig. 1. Note that the contribution from the path $(i \to j)$ may be different from that of $(j \to i)$ even for undirected networks. Then we define the load ℓ_k of a vertex k as the total amount of data packets passing through that vertex when all pairs of vertices send and receive one unit of data packet between them; $\ell_k = \sum_{i,j} \ell_k^{(i\to j)}$.

Since the packets are conserved, total load contributed by one pair is simply related to the shortest path length d_{ij} between them, by $\sum_k \ell_k^{(i\to j)} = d_{ij} + 1$. Thus we have

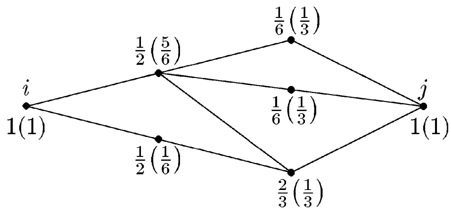

Fig. 1. The load at each vertex due to a unit packet transfer from the vertex i to the vertex j (from j to i). In this diagram, only the vertices along the shortest paths between (i, j) are shown.

the sum rule for ℓ_k:

$$\sum_k \ell_k = \sum_{i,j} (d_{ij} + 1) = N(N-1)(d+1) \sim N^2 d \; . \tag{1}$$

One can also define an equivalent quantity associated with bonds rather than vertices in analogy with the bond percolation problem. In this case, the bond-load contribution to a bond b from the pair $(i \to j)$, denoted as $\ell_b^{(i\to j)}$, is similarly defined and $\ell_b = \sum_{i,j} \ell_b^{(i\to j)}$ is the load of the bond b. The sum rule for the bond-load is $\sum_b \ell_b = N(N-1)d$. In social network theory, there are several kinds of *centrality* which measure the importance of a vertex in a network [21]. One is the degree centrality which simply associates the degree of a vertex as the measure of power. Another measure is the *betweenness centrality* [22]. It is defined as the number of times the shortest paths between every pair of vertices pass through that particular vertex. When there are more than one shortest path for a given pair, say g, each path contributes $1/g$. This quantity is exactly the same as the load for trees where there is only one shortest path for any pair of vertices, but is slightly different in general. However, we find the two quantities show very similar scaling behaviors for SF networks [23].

We now introduce the load distribution function $P_L(\ell)$, which can be measured by fraction of vertices or bonds whose load is ℓ, averaged over an appropriate ensemble. For SF networks, one may expect it also show the SF features and follow the power law. If one assumes that the vertex-load ℓ versus the rank s is of the power-law form $\ell \sim \ell_{\max} s^{-\beta}$, then one can easily see that

$$P_L(\ell) \sim \frac{\ell_{\min}^{\delta-1}}{\ell^{\delta}} \tag{2}$$

for $\ell_{\min} < \ell < \ell_{\max}$ with $\delta = 1 + 1/\beta$ and

$$\ell_{\min} \sim \ell_{\max}/N^{\beta} \sim \begin{cases} Nd & \text{if } \beta < 1 \; , \\ Nd/\ln N & \text{if } \beta = 1 \; , \\ N^{2-\beta} d & \text{if } \beta > 1 \; . \end{cases} \tag{3}$$

For the bond-load problem, assuming the total number of bonds scales with N as $\sim N^{1+\theta}$, Eq. (3) is modified to

$$\ell_{\min} \sim \ell_{\max}/N^{\beta(1+\theta)} \sim \begin{cases} N^{1-\theta}d & \text{if } \beta < 1\,, \\ N^{1-\theta}d/\ln N & \text{if } \beta = 1\,, \\ N^{2-\beta(1+\theta)}d & \text{if } \beta > 1\,. \end{cases} \tag{4}$$

When $\theta > 0$, we have an accelerated network [24].

In the following, we review the results reported in [25]. We generate several large SF networks with tunable parameters and measure the load of all vertices. We find indeed that the load distribution $P_L(\ell)$ follows a power law $P_L(\ell) \sim \ell^{-\delta}$. Moreover, the exponent $\delta \approx 2.2$ we obtained is surprisingly robust, insensitive to the detail of the SF network structure as long as the degree exponent is in the range, $2 < \gamma \leqslant 3$. When $\gamma > 3$, δ increases as γ increases, however. The universal behavior is also valid for directed networks, when $2 < \{\gamma_{\rm in}, \gamma_{\rm out}\} \leqslant 3$. Since the degree exponents in most of real-world SF networks satisfy $2 < \gamma \leqslant 3$, the universal behavior is interesting.

First, we discuss the *static* model of SF network which is constructed as follows. There are N vertices in the system from the beginning, which are indexed by an integer i $(i = 1, \ldots, N)$. We assign the weight $p_i = i^{-\alpha}$ to each vertex, where α is a control parameter in $[0, 1)$. Next, we select two different vertices (i, j) with probabilities equal to the normalized weights, $p_i/\sum_k p_k$ and $p_j/\sum_k p_k$, respectively, and add an edge between them unless one exists already. This process is repeated until mN edges are made in the system. Then the mean degree is $2m$. Since edges are connected to a vertex with frequency proportional to the weight of that vertex, the degree at each vertex is given as

$$k_i \sim \left(\frac{N}{i}\right)^{\alpha} \tag{5}$$

and $\sum_j k_j = 2mN$. Then it follows that the degree distribution follows the power law, $P_D(k) \sim k^{-\gamma}$, where γ is given by $\gamma = 1 + 1/\alpha$. Thus, adjusting the parameter α in [0,1), we can obtain various values of the exponent γ in the range, $2 < \gamma < \infty$.

Once a SF network is constructed, we select an ordered pair of vertices (i, j) on the network, and identify the shortest path(s) between them and measure the load on each vertex along the shortest path using the modified version of the breath-first search algorithm introduced by Newman [26]. We have measured the load ℓ_i for the networks with various γ. It is found numerically that the load ℓ_i follows the formula,

$$\frac{\ell_i}{\sum_j \ell_j} \sim \frac{1}{N^{1-\beta} i^{\beta}}\,, \tag{6}$$

with $\beta = 0.80(5)$. The value of β is insensitive to different values of the exponent γ in the range, $2 < \gamma \leqslant 3$ as shown in Fig. 2. The total load, $\sum_j \ell_j$ scales as $\sim N^2 \log N$ confirming the small-world property that $d \sim \ln N$. From Eq. (6), it follows that the load exponent is $\delta = 1 + 1/\beta \approx 2.2(1)$, independent of γ in the range, $2 < \gamma \leqslant 3$. Direct measure of $P_L(\ell)$ also gives $\delta \approx 2.2(1)$ as shown in Fig. 3. We also check δ for different mean degrees $m = 2$, 4 and 6, but obtain the same value, $\delta \approx 2.2(1)$ as shown

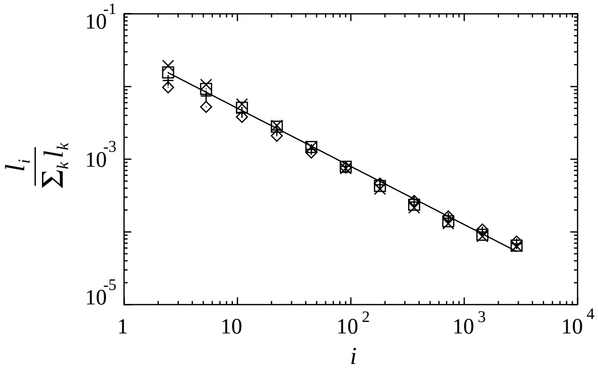

Fig. 2. Plot of the normalized load $\ell_i/\sum_k \ell_k$ versus vertex index i in double logarithmic scales for the scale-free networks with different degree exponents $\gamma = 2.25$ ($\times$), 2.5 ($\square$), 2.75 ($+$), and 3.0 ($\diamond$). The solid line is the linear fit and has a slope -0.80. Simulations are performed for $N = 10,000$ and $m = 2$ and all data points are averaged over 10 configurations.

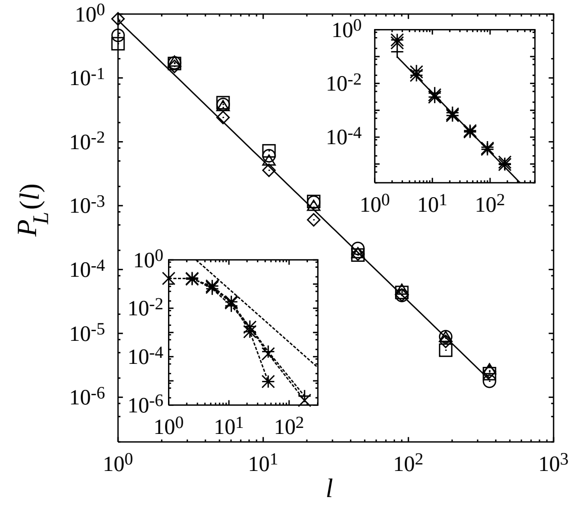

Fig. 3. Plot of the load distribution $P_L(\ell)$ versus ℓ for various $\gamma = 2.25$ ($\diamond$), 2.5 ($\triangle$), 2.75 ($\bigcirc$) and 3.0 ($\square$) in double logarithmic scales. The linear fit (solid line) has a slope -2.2. Simulations are performed for $N = 10,000$ and $m = 2$ and all data points are averaged over 10 configurations. Lower inset: Same plot for $\gamma = 4$ ($+$), 5 ($\times$), and ∞ ($*$). The line having a slope -2.2 is drawn to compare the data with the case for $2 < \gamma \leqslant 3$. Upper inset: Plot of $P_L(\ell)$ versus ℓ for different $m = 2$, 4 and 6, but for the same $\gamma = 2.5$.

in the upper inset of Fig. 3. Thus, we conclude that the exponent δ is a generic quantity for this network. However, for $\gamma > 3$, δ depends on γ in a way that it increases as γ increases. Eventually, the load distribution decays exponentially for $\gamma = \infty$ as shown in the lower inset of Fig. 3. Thus, the transport properties of the SF networks with

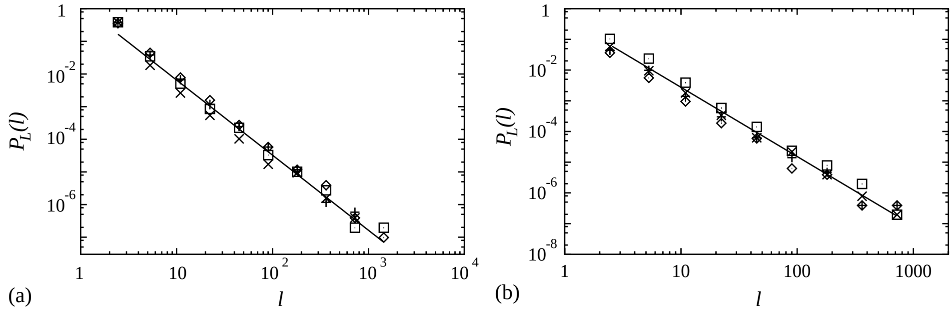

Fig. 4. (a) Plot of the load distribution $P_L(\ell)$ versus ℓ for the evolving model. The data are obtained for γ=2.25 (×), 2.5 (□), 2.75 (+) and 3.0 (◇). The fitted line has a slope −2.2. (b) Plot of the load distribution $P_L(\ell)$ versus ℓ for the directed case. The data are obtained for $(\gamma_{in}, \gamma_{out}) = (2.1, 2.3)$ (◇), (2.1, 2.7) (+), (2.5, 2.7) (□) and (2.5, 2.2) (×). The fitted line has a slope −2.3.

$\gamma > 3$ are fundamentally different from those with $2 < \gamma \leqslant 3$. This is probably due to the fact that for $\gamma > 3$, the second moment of $P_D(k)$ exists, while for $\gamma \leqslant 3$, it does not. Note that Eqs. (5) and (6) combined gives a scaling relation between the load and the degree for this network as

$$\ell \sim k^{(\gamma-1)/(\delta-1)} . \tag{7}$$

Thus, when and only when $\gamma = \delta$, the load at each vertex is directly proportional to its degree. Otherwise, it scales nonlinearly.

Next, we generate SF networks in an evolving way according to the method proposed by Kumar et al. [27], which is similar to the method proposed by Simon in 1955 in their idea [28]. The stochastic rule includes two ingredients, the duplication and the mutation. At each time step, a new vertex is introduced and it creates m edges connecting to existing vertices by the following rule: Select an existing vertex randomly. Associated with it are m vertices to which edges were added previously at its creation. Add an edge to the selected vertex or to any one of the associated vertices with probability p_K and $1 - p_K$, respectively. Repeat this m times. The network generated in this way exhibits a power law in its degree distribution, where the degree exponent is given by $\gamma = (2 - p_K)/(1 - p_K)$. The BA model is the case when $p_K = 0.5$. Through this model, we also obtain the load exponent $\delta \approx 2.2$ for different values of the degree exponent in $2 < \gamma \leqslant 3$ as shown in Fig. 4a, which confirms the previous result. The load-degree scaling, Eq. (7), is also satisfied.

Next, we consider the case of directed SF network. The directed SF networks are generated following the static rule. In this case, we assign two weights $p_i = i^{-\alpha_{out}}$ and $q_i = i^{-\alpha_{in}}$ $(i = 1, \ldots, N)$ to each vertex for outgoing and incoming edges, respectively. Both control parameters α_{out} and α_{in} are in the interval $[0, 1)$. Then two different vertices (i, j) are selected with probabilities, $p_i / \sum_k p_k$ and $q_j / \sum_k q_k$, respectively, and an edge from the vertex i to j is created with an arrow, $i \to j$. The SF networks generated in this way show the power law in both outgoing and incoming degree distributions with the exponents γ_{out} and γ_{in}, respectively. They are given as $\gamma_{out} = (1 + \alpha_{out})/\alpha_{out}$

and $\gamma_{in} = (1 + \alpha_{in})/\alpha_{in}$. Thus, choosing various values of α_{out} and α_{in}, we can determine different exponents γ_{out} and γ_{in}. Following the same steps as for the undirected case, we obtain the load distribution on the directed SF networks. The load exponent δ obtained is $\approx 2.3(1)$ as shown in Fig. 4b, consistent with the one for the undirected case, also being independent of γ_{out} and γ_{in} in $2 < \{\gamma_{out}, \gamma_{in}\} \leqslant 3$. Therefore, we conjecture that the load exponent is a universal value for both the undirected and directed cases.

In conclusion, we have introduced a physical quantity, load $\{\ell_i\}$ associated with each vertex i motivated by the problem of data packet transport on networks. For the SF networks generated in various ways, it is found that the load distribution follows a power law, $P_L(\ell) \sim \ell^{-\delta}$. The load exponent $\delta \approx 2.2(1)$ turns out to be insensitive to the details of the network structure, as long as the degree exponent γ is in the range $(2, 3]$. Moreover, it is also the same for both directed and undirected cases within our numerical uncertainties. Therefore, we conjecture that the load exponent is a generic quantity to characterize SF networks. Since the degree exponents for most of real-world SF networks are in the range $2 < \gamma \leqslant 3$, the universal behavior we found may have interesting implications to the interplay of their structure and dynamics, and could be taken as a generic nature of the SF networks by which we can classify them into universality classes. Work in this direction is in progress [23].

This work was supported by the Korean Research Foundation (Grant No. 01-041-D00061).

References

[1] S.H. Strogatz, Nature (London) 410 (2001) 268.
[2] R. Albert, A.-L. Barabási, Rev. Mod. Phys. 74 (2002) 47.
[3] S.N. Dorogovtsev, J.F.F. Mendes, Adv. Phys. 51 (2002) 1079.
[4] P. Erdös, A. Rényi, Publ. Math. Inst. Hung. Acad. Sci. Ser. A 5 (1960) 17.
[5] D.J. Watts, S.H. Strogatz, Nature (London) 393 (1998) 440.
[6] A.-L. Barabási, R. Albert, H. Jeong, Physica A 272 (1999) 173.
[7] R. Albert, H. Jeong, A.-L. Barabási, Nature (London) 401 (1999) 130.
[8] A. Broder, et al., Computer Networks 33 (2000) 309.
[9] M. Faloutsos, P. Faloutsos, C. Faloutsos, Comput. Commun. Rev. 29 (1999) 251.
[10] R. Pastor-Satorras, A. Vázquez, A. Vespignani, Phys. Rev. Lett. 87 (2001) 258701.
[11] K.-I. Goh, B. Kahng, D. Kim, Phys. Rev. Lett. 88 (2002) 108701.
[12] S. Redner, Eur. Phys. J. B 4 (1998) 131.
[13] M.E.J. Newman, Proc. Natl. Acad. Sci. USA 98 (2001) 404.
[14] A.-L. Barabási, H. Jeong, Z. Neda, R. Ravasz, A. Schubert, T. Vicsek, Physica A 311 (2002) 590.
[15] H. Jeong, S.P. Mason, A.-L. Barabási, Z.N. Oltvai, Nature (London) 411 (2001) 41.
[16] T. Ito, et al., Proc. Natl. Acad. Sci. USA 97 (2000) 1143;
T. Ito, et al., Proc. Natl. Acad. Sci. USA 98 (2001) 4569.
[17] H. Jeong, B. Tombor, R. Albert, Z.N. Oltvani, A.-L. Barabási, Nature (London) 407 (2000) 651.
[18] A.-L. Barabási, R. Albert, Science 286 (1999) 509.
[19] P.L. Krapivsky, S. Redner, F. Leyvraz, Phys. Rev. Lett. 85 (2000) 4629;
S.N. Dorogovtsev, J.F.F. Mendes, A.N. Samukhin, Phys. Rev. Lett. 85 (2000) 4633.
[20] B. Bollobás, Random Graphs, Academic Press, New York, 1985.
[21] S. Wasserman, K. Faust, Social Network Analysis, Cambridge University, Cambridge, 1994.
[22] L.C. Freeman, Sociometry 40 (1977) 35.
[23] K.-I. Goh, E.S. Oh, H. Jeong, B. Kahng, D. Kim, Proc. Natl. Acad. Sci. USA 99 (2002) 12583.

[24] S.N. Dorogovtsev, J.F.F. Mendes, Phys. Rev. E 63 (2001) 025101(R).
[25] K.-I. Goh, B. Kahng, D. Kim, Phys. Rev. Lett. 87 (2001) 278701.
[26] M.E.J. Newman, Phys. Rev. E 64 (2001) 016131;
M.E.J. Newman, Phys. Rev. E 64 (2001) 016132.
[27] R. Kumar et al., Proceedings of the 41st IEEE Symposium on Foundations of Computer Science, 2000.
[28] H. Simon, Biometrika 42 (1955) 425.