LOG-PERIODIC OSCILLATIONS
IN DEGREE DISTRIBUTIONS
OF HIERARCHICAL SCALE-FREE NETWORKS[*]

Krzysztof Suchecki and Janusz A. Hołyst[†]

Faculty of Physics
and Center of Excellence for Complex Systems Research
Warsaw University of Technology
Koszykowa 75, 00-662 Warsaw, Poland

Hierarchical models of scale free networks are introduced where numbers of nodes in clusters of a given hierarchy are stochastic variables. Our models show periodic oscillations of degree distribution $P(k)$ in the log–log scale. Periods and amplitudes of such oscillations depend on network parameters. Numerical simulations are in a good agreement to analytical calculations.

PACS numbers: 89.75.Hc, 89.75.–k, 89.75.Da, 89.75.Fb

1. Introduction

Recently there is a large interest in scale-free networks that seem to be good approximations for such systems as the Internet, World Wide Web, social or biological networks; for a review see [1–4]. A simple model that exhibits the power law for degree distributions $P(k)$ observed in real complex networks is the Barabási–Albert model of preferential attachment [1]. The model, however, suffers from very low values of the clustering coefficient C [5] for large networks as compared to observations of real systems [1–4]. To overcome this discrepancy a model of hierarchical networks has been introduced by Ravasz and Barabási (RB) where the clustering coefficient is much larger [6]. The RB network consists of hierarchically connected clusters, where numbers of nodes in every cluster of a given hierarchy, are the

---

same. The degree distribution $P(k)$ in this approach also exhibits power-law. However, it is only a general trend. In fact, the degree distribution consists of delta-peaks for only a few degree values, instead of continuous distribution observed in real networks. In this paper, we introduce a class of more general models, where number of nodes in every cluster is a stochastic variable, what seems to be more justified for real network models. As a result the peaks of $P(k)$ are blurred, creating a network with wide range of possible k values, but the log-periodic behavior of $P(k)$ is still clearly visible.

Let us remind that log-periodic oscillations are characteristic features of systems where a discrete self-similarity is present [7] and the effect can occur even without a preexisting hierarchy [7] in such various systems as earthquakes [8, 9] or financial markets where log-periodic oscillations were observed as possible precursors for financial crashes [10–12]. Such oscillations were also found for mean residence times at chaotic crisis, where a collision of a fractal attractor with a fractal or a non-fractal basin of another attractor, takes place [13, 14] and for the stochastic resonance in chaotic systems near a crisis point [15].

2. The model

Our model possesses two parameters, a distribution $P_M(m)$, where $m = 1, 2, 3, \ldots$ and a number $p \in (0, 1]$. We start out from a single cluster (a cluster of hierarchy 0) of $m + 1$ fully connected nodes (Fig.1), where m is a random number from a distribution $P_M(m)$. One node in the cluster is its *central* node. The central node of the cluster is a center of hierarchy 0. Next, we call our cluster the central one and create a random number m of similar clusters. Each is created in the same way as the central cluster, but we pick a random number m for each one independently, therefore, they may include different numbers of nodes. Next, we connect a part p of all nodes in non-central clusters to the central node in the central cluster. This node becomes the central node for the whole cluster of hierarchy 1 we have obtained so far. Similarly the central node of our cluster is a center of hierarchy 1. We repeat the process, until we get a network of a desired hierarchy. This model is referred to as P1 model. The model is generalization of the stochastic model proposed by Barabási and Ravasz [6]. If we take $P_M(m) = \delta(m, m_0)$, where m_0 is constant, our model simplifies to BR model, with number of nodes and degree distribution determined strictly by p and m_0 values.

A variation of the model has been also studied. In each hierarchy d we connect not a fraction p of nodes but a fraction p^d. This model is referred to as the PD model.
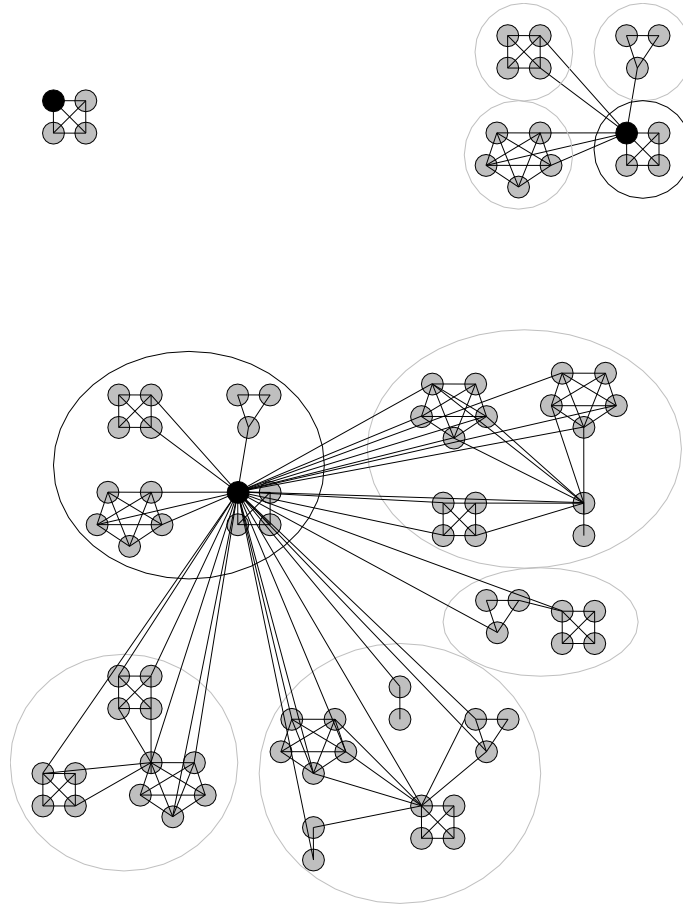
Fig. 1. The first three steps of network creation in the P1 model.

3. Degree distribution of P1 and PD models

As previously noted, for $P_{\mathrm{M}}(m) = \delta(m, m_0)$ we get a degree distribution identical to that of BR model. It consists of separate peaks, corresponding to degrees of central nodes of the following hierarchies. Central nodes of given hierarchy have a fixed degree, dependent only on the network parameter p. At the logarithmic scale the distance between neighboring peaks is approximately constant and equals $\log(m_0+1)$. The peaks follow laws of discrete scaling [7]. The heights of peaks with degrees k_i decrease as $k_i^{-\gamma}$, and distances between consecutive peaks fulfill the relation $k_{i+1}/k_i = \lambda$. The probability $P(k)$ between peaks equals zero what means that only nodes with peculiar degrees are possible. But what happens when the number m is not a fixed value ?

Numerical simulations show that each peak blurs, depending on the $P_M(m)$ distribution. If the blur is small, the distribution consists of separate peaks, but they are not delta-shaped. If the blur is large enough, the peaks overlap and a continuous degree distribution is obtained. Fig. 2 shows the degree distributions for both cases in P1 model. Both display a discrete scaling, and have the same scaling exponent (up to statistical fluctuations), independent of network parameters.

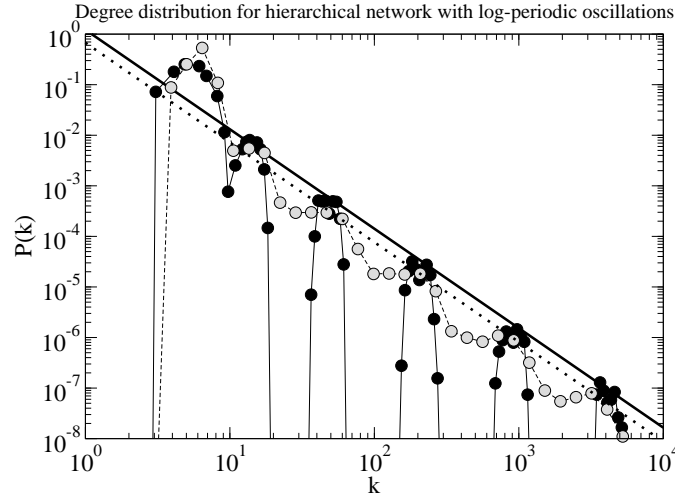Similar behavior has been observed in the PD model, although scaling exponent is parameters dependent.

Fig. 2. Degree distribution for two networks with different $P_M(m)$. Filled circles are for $m = 3$ or 4 with equal probability, gray circles are for uniform m distribution between 1 and 5. The straight lines show scaling of peak heights and correspond to $\gamma = 0.967$ for the first case (continuous line) and $\gamma = 0.973$ for the second (dotted line).

4. Mean m value approach

Numerical simulations have shown that when m is not a constant but a random number from a given distribution, peaks blur, eventually overlapping and creating a continuous distribution. However, regardless of the actual shape, the distribution still consists of peaks. Each peak has an average degree k and a mass n representing number of nodes that belong to this peak. All nodes in a peak are centers of the same hierarchy. Using a *mean value* of m, the distance between peaks and their relative heights can be easily found. From these two values we get directly the discrete scaling ratio λ and the scaling exponent γ. In the following calculations we neglect the

degree increase of nodes due to their connections to the central cluster, as this effect increases the node degree at most by d, what is insignificant for higher hierarchy centers.

4.1. P1 model

Let us denote an average degree of peak of hierarchy d by k_d, an average number of nodes in a cluster of hierarchy d by N_d, and an average number of centers of hierarchy d in a network of hierarchy h by n_d^h. The network size N_d increases exponentially with hierarchy d as $N_d = \langle m+1 \rangle^{d+1}$. Centers of hierarchy 0 have a degree k_0 equal to $\langle m \rangle$ and it increases by $p \langle m \rangle N_{d-1}$ in each next hierarchy d. We obtain

$$k_d = \langle m \rangle + p \langle m+1 \rangle \left(\langle m+1 \rangle^d - 1 \right). \tag{1}$$

If $\langle m+1 \rangle > 1$ and $d \gg 1$, the above expression can be simplified to

$$k_d \approx p \langle m+1 \rangle^{d+1}. \tag{2}$$

If the condition is not satisfied, distances between peaks are not constant at logarithmic scale and the network is not scale-free.

Since the discrete scaling ratio λ simply equals k_{d+1}/k_d thus we get $\lambda \approx \langle m+1 \rangle$.

The scaling exponent γ can be found using the cumulative degree distribution. Starting from

$$P(k) = \frac{\Delta P_{\text{cum}}}{\Delta k} = \frac{\Delta P_{\text{cum}}}{\Delta d} \frac{\Delta d}{\Delta k}, \tag{3}$$

where d are consecutive hierarchies, and using calculations presented in Appendix A we get $P(k) \sim k^{-2}$ so the scaling exponent γ equals 2, regardless of p and $P_{\text{M}}(m)$. Note that this scaling is valid for peak masses n_d^h only.

4.2. PD model

The case of PD model is very similar to the P1 model. However, since instead of a fraction p we connect a fraction p^d of nodes from non-central clusters, the degree k_d is

$$k_d = \langle m \rangle \frac{1 - (p \langle m+1 \rangle)^{d+1}}{1 - p \langle m+1 \rangle}. \tag{4}$$

When we assume that $p \langle m+1 \rangle > 1$ we can omit one in the numerator and get the discrete scaling ratio $\lambda = k_{d+1}/k_d \approx p \langle m+1 \rangle$. Similarly to the P1 model, if it is not true, the network is not scale-free.

To find the scaling exponent, again we use cumulative degree distribution and Eq. (3). For the PD model we get $\gamma = 1 + (\ln \langle m + 1 \rangle)/(\ln p \langle m + 1 \rangle)$. Note that since $p \in (0, 1]$ and $p \langle m + 1 \rangle > 1$ for scale-free networks, the scaling exponent is always greater than 2.

4.3. Numerical data

Numerical simulations have been performed for networks of hierarchy 6, with $p = 0.5$ and various uniform distributions of m, to find out whether analytic predictions are correct.

Tables I and II contain obtained data. Fig. 3 shows the comparison between prediction and results.

As it can be seen the numerical data are in a good agreement with our analytic predictions. The largest deviation is for low $\langle m \rangle$ and for low p, where our approximations were poor.

TABLE I

Distribution of m numbers, analytic and numerical scaling exponents γ and the logarithm of discrete scaling ratio λ for the model P1. Data obtained from averaging over 30 networks.

m	$\langle m + 1 \rangle$	γ_{analyt}	γ_{numer}	$\log \lambda_{\text{analyt}}$	$\log \lambda_{\text{numer}}$
1 to 2	2.5	2	1.981	0.398	0.397
1 to 3	3	2	1.978	0.477	0.461
1 to 4	3.5	2	1.931	0.544	0.556
1 to 5	4	2	1.973	0.602	0.606

TABLE II

Distribution of m numbers, analytic and numerical scaling exponents γ and the logarithm of discrete scaling ratio λ for the model PD. Data obtained from averaging over 2000 networks. In the first row, the exact λ value was impossible to obtain, due to a very weak periodic behavior.

m	$\langle m + 1 \rangle$	γ_{analyt}	γ_{numer}	$\log \lambda_{\text{analyt}}$	$\log \lambda_{\text{numer}}$
1 to 2	2.5	5.106	3.858	0.097	~ 0.16
1 to 3	3	3.710	3.067	0.176	0.208
1 to 4	3.5	3.239	3.038	0.243	0.271
1 to 5	4	3.000	2.846	0.301	0.32

Discrete scaling ratio and scaling exponent for P1 and PD models
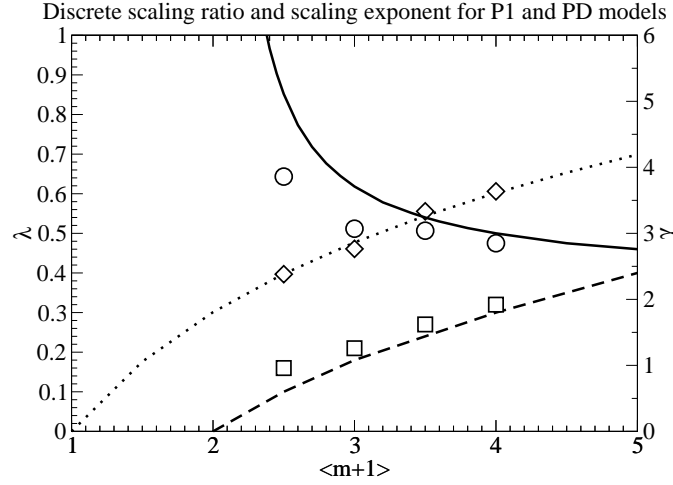

Fig. 3. The discrete scaling ratio λ and the scaling exponent γ. Symbols are the values obtained from numeric simulations, lines are analytic predictions. Diamonds and dotted line are for λ in P1 model. Squares and dashed line are for λ in PD model. Circles and solid line are for γ in PD model. In all cases, networks are of hierarchy 6, with parameter $p = 0.5$.

5. Exact degree distribution

Up to now, all calculations have been performed using only the average m value, treating the degree distribution as series of peaks. We have been concentrating on relations between peak's masses and distances, while ignoring their shape. Here we find a shape of the degree distribution for the P1 model.

Let $P_{\mathrm{M}}(m)$ be a distribution of m, where m is a number of non-central clusters in each hierarchy. Let $\widetilde{P}_d(N)$ be a distribution of the network sizes N for hierarchy d. $P_d(k)$ is a degree distribution for a network of hierarchy d, $P_d^{\mathrm{c}}(k)$ is a degree distribution for the central node of hierarchy d.

The number of nodes in the network can be found as follows. Network of hierarchy $d = 0$ has $m + 1$ nodes what means $\widetilde{P}_0(N) = P_{\mathrm{M}}(N - 1)$. The size of each next hierarchy $d+1$ is a sum of $m+1$ independent values, which are the sizes of networks of hierarchy d

$$\widetilde{P}_{d+1}(N) = \sum_m P_{\mathrm{M}}(m) \sum_{n_1, n_2, \ldots, n_m} \widetilde{P}_d(n_1)\widetilde{P}_d(n_2)\ldots\widetilde{P}_d(N - n_m - \ldots - n_1). \quad (5)$$

This recursive formula describes the probability distribution for the network size N.

A network of hierarchy $d = 0$ has the degree distribution $P_0(k) = P_M(k)$. This distribution describes both regular nodes and a center of hierarchy 0, which have the same degree values. In each next hierarchy $d + 1$ the degree distribution for all nodes of hierarchy d or less is the same, since we omit the degree increase due to connections to the central node of higher hierarchy. Now, we multiply the distribution by $(m + 1)$ and add the degree distribution $P_{d+1}^c(k)$ for the central node of the network. This way we obtain an unnormalized degree distribution for the whole network of hierarchy $d + 1$

$$P_{d+1}(k) = \sum_m \left[(m + 1)P_M(m)P_d(k) + P_{d+1}^c(k) \right] . \tag{6}$$

The center is roughly connected to fraction p of all nodes in the network, what means it possesses the degree $p\,N$. This yields the distribution of its degree equal to $P_1^c(k) = \widetilde{P}_1(k/p)$. As the result we obtain

$$P_{d+1}(k) = \sum_m \left[(m + 1)P_M(m)P_d(k) + \widetilde{P}_{d+1}\left(\frac{k}{p}\right) \right] . \tag{7}$$

This recursive formula describes the unnormalized degree distributions for networks of consecutive hierarchies d, with the exception of $d = 1$. Since $P_0(k)$ describes not only centers of hierarchy 0 but both regular nodes and centers, we must account for that. We do so by multiplying $P_0(k)$ in the formula by the average basic cluster size $\langle m + 1 \rangle$.

$$P_1(k) = \sum_m \left[\langle m + 1 \rangle \, (m + 1)P_M(m)P_0(k) + \widetilde{P}_1\left(\frac{k}{p}\right) \right] . \tag{8}$$

In the above calculations, as in the calculations using average m, we omitted the degree increase due to connections to the central node of higher hierarchy. This is insignificant for higher hierarchy centers, as the increase is at most d, while the center degree increases exponentially with d. We have used the formula $P_d^c(k) = \widetilde{P}_d(k/p)$ in the above calculations. In reality, the degree distributions are discrete, with natural k values. Depending on how we round the number of connections to the central node, we should interpret the above formula accordingly.

We rounded the number of connections $p\,N$ down, what gives the following formula for interpreting the probability with fractional argument

$$P_d^c(k) = \sum_{l \geq k/p, l < (k+1)/p} \widetilde{P}_d(l) . \tag{9}$$

It means that the probability of getting the center of degree k equal to the sum of probabilities for N that lead to this k. Along with

$$\widetilde{P}_0(N) = P_M(N - 1) \tag{10}$$

and

$$P_0(k) = P_M(k)\,. \tag{11}$$

Eqs. (5) and (7)–(9) allow one to find numerically an exact but unnormalized degree distribution for the P1 model.

Comparing these formulas with numerical data one can see that our calculations are correct for higher degrees, where approximations we used are accurate (Figs. 4, 5).
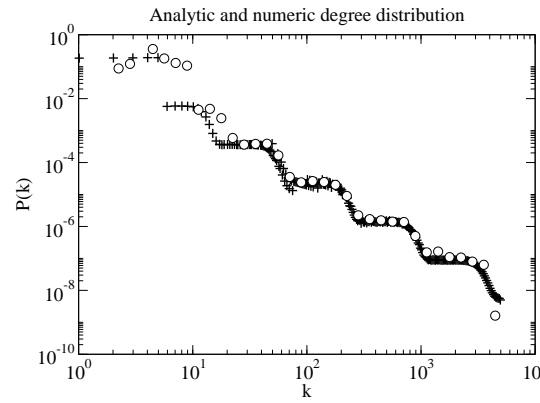

Fig. 4. The degree distribution for the network with uniform m distribution from 1 to 5, and the hierarchy $d = 5$. The graph shows analytic (crosses) and numeric data (circles).
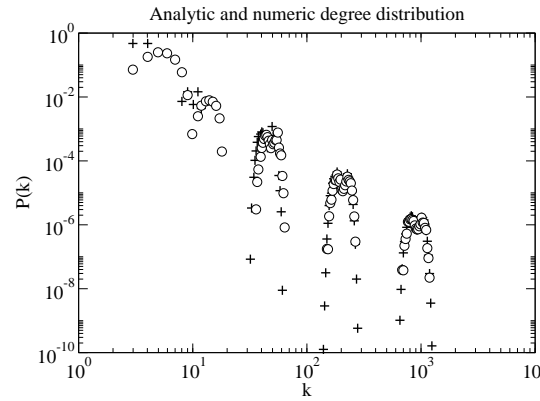

Fig. 5. The degree distribution for the network with uniform m distribution (3 or 4), and the hierarchy $d = 4$. The graph shows analytic (crosses) and numeric data (circles).

Using the degree distributions obtained with our formulas (Eq. (5), Eqs. (7)–(11)), a relation between the distribution of m and a peak shape has been found. We have studied various uniform distributions of m and have found linear relation between the standard deviation of the distribution $P(\ln m)$ and the standard deviation of peaks in the $P(\ln k)$ distribution (Fig. 6).
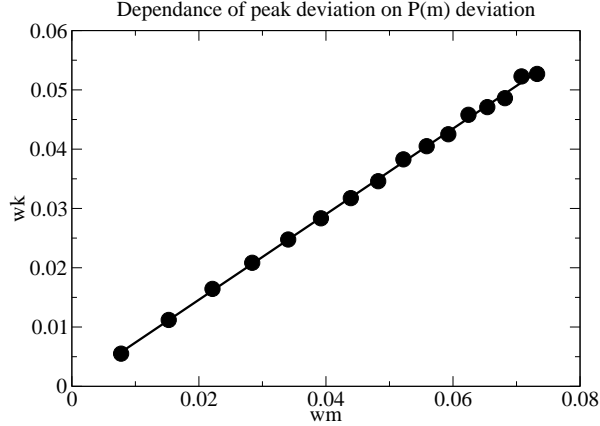

Fig. 6. The relation between the peak width wk (a standard deviation of peaks in $P(\ln k)$) and the distribution width wm (a standard deviations of $P(\ln m)$). The slope of the line is $\alpha = 0.722$.

The peak deviations in the distribution $P(k)$ are calculated at the logarithmic scale of k

$$\sigma = \frac{\sum (\ln k - \langle \ln k \rangle)^2 \, P(\ln k)}{\sum P(\ln k)} \; . \tag{12}$$

Similar formula has been used to calculate the deviation of $P(m)$. Peak deviations have been calculated for the peak of the highest hierarchy. In the case of overlapping peaks the minimums of $P(k)$ have been considered the borders of the peak. The approximation is quite accurate, as $P(k)$ decays fast when we go away from the peak average k value.

6. Discussion

The question occurs, whether our model corresponds to real network systems. It is obvious that many real networks posess a hierarchical structure but, of course, a detailed mechanism responsible for its emergence is unknown. According to our knowledge, log-periodic oscillations around the power law in degree distributions were never directly reported in the studies of real networks or corresponding models. One can suspect, however, that

in many cases such oscillations were visible and could be overlooked if the binning or data averaging had been performed. Small amplitude oscillations can be also easily confused with random fluctuations. The situation resembles oscillations around the scaling law in chaotic crises, where the periodic part is also often omitted as fluctuations [13, 14].

A clear example of log-periodic oscillations for real networks can be seen in the study of liability connections between Austrian banks [16]. As the authors stress [16] a significant part of the studied banking sector posesses a strong hierarchical structure, what can be easily detected looking at a corresponding connection graph. Two periods of oscillations can be identified at out-degree distribution describing the number of liabilities to other Austrian banks (regardless of liability size) [16]. The period of the oscillations is approximately $\lambda = k_{i+1}/k_i \approx 3$. According to our theory, they are a result of the network's hierarchical structure.

We have found also far less visible oscillations in the studies of computer directory trees [17] and World Trade Web [18], where the hierarchical structure can be identified at a corresponding connection graph [17] or in dependence of a clustering coefficient on a node degree [18]. Sizes of such oscillations are, however, at a fluctuations level. The other possible example can be found in the paper [19], which presents a non-monotenous behavior of degree distribution of $P(k)$ for a shareholding network in Japan. Here, a single wave around the power law can be observed, where $\lambda = k_{i+1}/k_i \approx 10$.

7. Conclusions

In conclusion, we have shown that hierarchical networks models display log-periodic oscillations in the degree distribution when the number of clusters forming the self-similar hierarchy is a stochastic variable. The period and the amplitude of these oscillations reflect the hierarchical structure of the network. We also point out examples of real networks that display such features. It follows that observations of log-periodic oscillations in degree distributions of real networks can give hints towards the existence of hidden hierarchical structures in such systems.

Appendix A

Mean m value calculations

This Appendix contains exact calculations regarding the discrete scaling ratio λ and the scaling exponent γ for the mean m approach. The number k_d is an average degree of peak of hierarchy d, N_d is an average number of nodes in network of hierarchy d, n_d^h is an average number of centers of hierarchy d in a network of hierarchy h. The number $\langle m + 1 \rangle$ is a mean number of clusters in each hierarchy.

The Eq. (2) can be obtained as

$$
\begin{aligned}
k_d &= k_{d-1} + \langle m \rangle \ N_{d-1} \cdot p = \langle m \rangle + p \, \langle m \rangle \sum_{d}^{i=1} \langle m + 1 \rangle^i \\
&= \langle m \rangle + p \, \langle m + 1 \rangle \left(\langle m + 1 \rangle^d - 1 \right) .
\end{aligned}
\tag{A.1}
$$

The scaling exponent γ for model P1 has been obtained using the cumulative degree distribution

$$
P(k) = \frac{\Delta P_{\text{cum}}}{\Delta k} = \frac{\Delta P_{\text{cum}}}{\Delta d} \frac{\Delta d}{\Delta k} .
\tag{A.2}
$$

First we find an expression for n_d^h. There is only one center of hierarchy $d = h$ in the network of such a hierarchy. Each time the network hierarchy h increases, the number of centers of hierarchy d increases $\langle m + 1 \rangle$ times. The exception is the first step, where one node becomes center of the higher hierarchy. Because of that n_d^h increases only by the factor of $\langle m \rangle$ for that step. We obtain $n_d^h = \langle m \rangle \, \langle m + 1 \rangle^{h-d-1}$ except for $n_h^h = 1$.

Now we calculate expressions in Eq. (A.2). Each next peak is smaller $\Delta P_{\text{cum}}/\Delta d = n_d^h \sim \langle m + 1 \rangle^{-d}$ while its average degree k_d increases $\Delta k/\Delta d \sim \langle m + 1 \rangle^d$. In such a way we obtain

$$
P(k) \sim \langle m + 1 \rangle^{-d} \, \langle m + 1 \rangle^{-d} = \langle m + 1 \rangle^{-2d} .
\tag{A.3}
$$

From Eq. (A.1) we get $d \approx \log_{\langle m+1 \rangle}(k/p) - 1 = (\ln k - \ln p)/(\ln \langle m+1 \rangle) - 1$. By putting so calculated d into Eq. (A.3) we get

$$
\begin{aligned}
P(k) \sim \langle m + 1 \rangle^{-2d} &= \exp\left(-2d \ln \langle m + 1 \rangle \right) \\
&= \exp\left(-2 \frac{\ln k - \ln p}{\ln \langle m + 1 \rangle} \ln \langle m + 1 \rangle + 2 \ln \langle m + 1 \rangle \right) \\
&\sim \exp\left(-2 \ln k \right) = k^{-2} .
\end{aligned}
\tag{A.4}
$$

For the PD model the Eq. (4) can be obtained in the following way

$$
\begin{aligned}
k_d &= k_{d-1} + \langle m \rangle \ N_{d-1} \ p^d = \langle m \rangle \sum_{i=0}^{d} p^i \ \langle m+1 \rangle^i \\
&= \langle m \rangle \ \frac{1 - (p \ \langle m+1 \rangle)^{d+1}}{1 - p \ \langle m+1 \rangle} \ .
\end{aligned}
\tag{A.5}
$$

The scaling exponent has been calculated in a similar way to the case of model P1. The slope $\Delta P_{\mathrm{cum}}/\Delta d$ is the same as in the previous case but $\Delta k/\Delta d \sim (p \ \langle m+1 \rangle)^d$ thus $P(k) \sim p^{-d} \ \langle m+1 \rangle^{-2d}$.

Expressing d by k we get $d = \log_{p\langle m+1 \rangle}((p \ \langle m+1 \rangle - 1) \ k/\langle m \rangle) - 1$ and putting that into Eq. (A.2) we get $P(k) \sim k^{-(1+(\ln\langle m+1 \rangle/\ln p\langle m+1 \rangle))}$ which yields exponent $\gamma = 1 + (\ln \langle m+1 \rangle/\ln p \ \langle m+1 \rangle)$ for PD model.

REFERENCES

[1] R. Albert, A.-L. Barabási, *Rev. Mod. Phys.* **74**, 47 (2002).

[2] M.E.J. Newman, *SIAM Rev.* **45**, 167 (2003).

[3] S.N. Dorogovtsev, J.F.F. Mendes, *Evolution of Networks: From Biological Nets to the Internet and WWW*, Oxford University Press, Oxford 2003.

[4] *Handbook of Graphs and Networks*, Ed. S. Bornhold, H.G. Schuster, Viley-VCH, Berlin 2002.

[5] A. Fronczak, P. Fronczak, J.A. Hołyst, *Phys. Rev.* **E68**, 046126 (2003).

[6] E. Ravasz, A.-L. Barabási, *Phys. Rev.* **E67**, 026112 (2003).

[7] D. Sornette, *Phys. Rep.* **297**, 239 (1998).

[8] D. Sornette, C.G. Sammis, *J. Phys. (France)* **5**, 607 (1995).

[9] H. Saleur, C.G. Sammis, D. Sornette, *Nonlinear Processes Geophys.* **3** 102 (1996).

[10] M. Ausloos, *Physica A* **285**, 48 (2000).

[11] D. Sornette, *Phys. Rep.* **378**, 1 (2003).

[12] W.X. Zhou, D. Sornette, *Int. J. Mod. Phys.* **C14**, 1107 (2003).

[13] K. Kacperski, J.A. Hołyst, *Phys. Rev.* **E60**, 403 (1999).

[14] K. Kacperski, J.A. Hołyst, *Phys. Lett.* **A254**, 53 (1999).

[15] S. Matyjaśkiewicz, A. Krawiecki, J.A. Hołyst, K. Kacperski, W. Ebeling, *Phys. Rev.* **E63**, 026215 (2001).

[16] M. Boss, H. Elsinger, M. Summer, S. Thurner, `cond-mat/0309582`.

[17] K. Klemm, V. Eguiluz, M. San Miguel, `cond-mat/0403239`.

[18] M.-Á. Serrano, M. Boguñá, *Phys. Rev.* **E68**, 015101(R) (2003).

[19] W. Souma, Y. Fujiwara, H. Aoyama, *Physica A* **344**, 73 (2004).