SCALING OF INTERNODE DISTANCES IN WEIGHTED COMPLEX NETWORKS

JULIAN SIENKIEWICZ

Faculty of Physics
Center of Execellence for Complex Systems Research
Warsaw University of Technology, Koszykowa 75
Warsaw, 00-662, Poland
julas@if.pw.edu.pl

JANUSZ A. HOŁYST

Faculty of Physics
Center of Execellence for Complex Systems Research
Warsaw University of Technology, Koszykowa 75
Warsaw, 00-662, Poland
jholyst@if.pw.edu.pl

We extend the previously observed scaling equation connecting the internode distances and nodes' degrees onto the case of weighted networks. We show that the scaling takes a similar form in the empirical data obtained from networks characterized by different relations between node's strength and its degree. In the case of explicit equation for $s(k)$ (e.g. linear or scale-free), the new coefficients of scaling equation can be easily obtained. We support our analysis with numerical simulations for Erdős-Rényi random graphs with different weight distributions.

Keywords: weighted networks; scaling; empirical data

PACS Nos.: 89.75.Hc, 89.75.Da

1. Introduction

The are different ways to represent the structure of real-world networks using the notions of graph theory. In particular, depending on the fact if we are interested only in the question of existence of connections between nodes or we would also like to consider the "intensity" of those links, the described network can be presented as *unweighted* or *weighted*. In most cases the initial network is of weighted character but in simple analytical methods the weights are usually removed.

The modeling of weighted networks has begun almost parallel with the unweighted case [1], however it is the higher availability of the empirical data that made possible to perform extensive and well set research on weighted transportation networks, networks of scientific collaboration [2] or in general the networks of

social contacts [3]. In the same time the works on connections between topology and weight's dynamics [4,5], transport on weighted networks [6,7] and optimal paths [8] were being developed. Only recently the weighted networks approach has been used to create a unified statistics combing Bose-Einstein and Fermi-Dirac distributions [9].

The considerations included in this paper are an attempt for extending the scaling relations presented in [10,11] onto the case of weighted networks. In the cited works it is shown that the mean distance (or internode distance) between nodes with degrees k_i and k_j is given by the following equation

$$\langle l_{ij} \rangle = a - b \log k_i k_j, \tag{1}$$

where a simple model of unweighted network with a negligible clustering coefficient leads to

$$a = 1 + \frac{\log(N \langle k \rangle)}{\log \kappa} \quad \text{and} \quad b = \frac{1}{\log \kappa}. \tag{2}$$

with being N the size of the network, $\langle k \rangle$ its average degree and κ the average branching degree. The scaling relation (1) has been spotted in many different networks ranging from public transport networks to biological ones[10,11] although the observed values of the coefficients a and b are different from the results given by Eq. (2).

2. Empirical data

In the case of weighted network instead of node's degree k_i one usually considers node's strength s_i i.e.

$$s_i = \sum_{j=1}^{j=k_i} w_{ij} \tag{3}$$

where w_{ij} is the weight of the link between nodes i and j. In order to find the relation between the internode distances $\langle l_{ij} \rangle$ and the product of nodes' strengths we examined the following systems: the network of flights between world-wide airports[a], public transport network in Warsaw[b] and two networks made from connections between the web-pages of web portals[c]. In the case of the airport network the weight w_{ij} stands for the number of different carriers making the connections between the airports i and j while the weights in Warsaw public transport network are

---

proportional to the total number of buses and trams traveling between stops i and j. In the web portals data the weight was defined as the total number of users who, being first at page i surfed to page j. For simplicity reasons in all datasets the weights were symmetrized by calculating the average of the original values i.e. $w = (w_{ij} + w_{ji})/2$. Due to high reciprocity between w_{ij} and w_{ji} this does not affect further analysis.
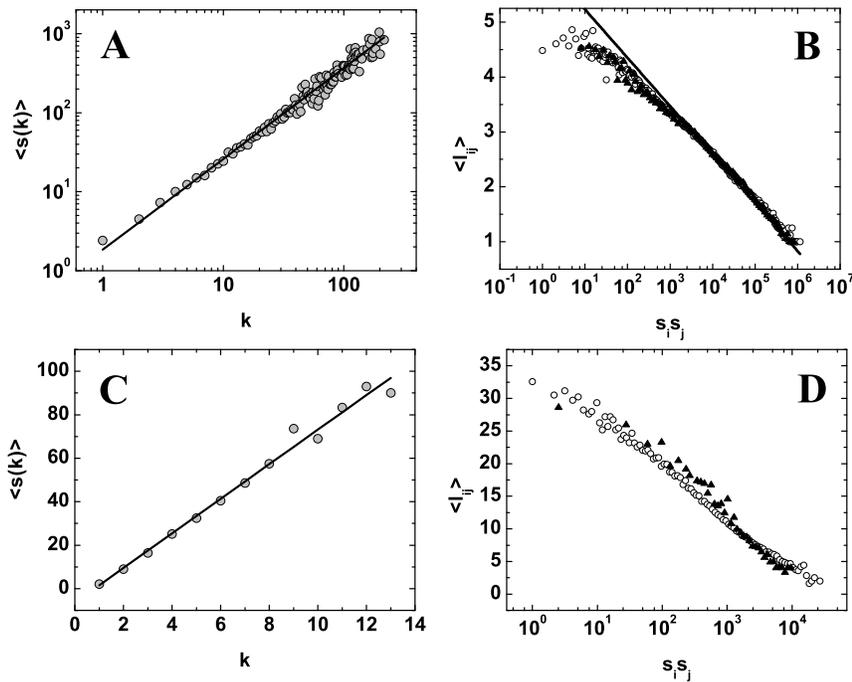

Fig. 1.   The average node's strength versus node's degree $\langle s(k) \rangle$ for world-wide airport network (A) and public transport network in Warsaw (C). The solid line is a fit to the empirical data $\langle s(k) \rangle \sim k^{1.15}$ (A) and $\langle s(k) \rangle = 8k - 6$ (C). The dependence of the internode distances $\langle l_{ij} \rangle$ on the product of nodes' strengths $s_i s_j$ (circles) and on the product of nodes' degrees $k_i k_j$ rescaled by the relation $\langle s(k) \rangle$ (triangles) for the airport connections (B) and public transport in Warsaw (D). The solid line in (B) is defined by the rescaled coefficients $\widetilde{a}$ and $\widetilde{b}$ obtained from Eqs (14).

The analysis of the results obtained from the above described datasets is presented in Figures 1 and 2. For each network we plot the dependence of average node's strength and its degree $\langle s(k) \rangle$ (Figures 1A, 1C, 2A and 2C) as well as the internode distances $\langle l_{ij} \rangle$ separately as a function of the product of degrees $k_i k_j$ and as a function of the product of strengths $s_i s_j$ (Figures 1B, 1D, 2B and 2D).

The plots presenting $\langle l_{ij} \rangle$ as a function of $k_i k_j$ were rescaled with respect to node degree values, substituting it with an approximated relation between the node strength and its degree, marked as solid curve in $\langle s(k) \rangle$ plots. The $\langle s(k) \rangle$ equation can take different forms: starting from lack of correlations in case of public

transport network in Warsaw ($\langle s(k) \sim k \rangle$, through a scale-free relation in airport network ($\langle s(k) \rangle \sim k^\alpha$), ending at a complicated dependence seen in web portals data ($\langle s(k) \rangle \sim k^\alpha \exp(\beta k)$). Nevertheless in each example the coincidence between the rescaled data and the empirically observed relation $\langle l_{ij}(s_i s_j) \rangle$is clearly visible. The oscillations observed on the logarithmic trend seen in Figures 2B and 2D are an effect of the high average degree present in both portals data. The origin of this phenomenon is extensively described in [14].
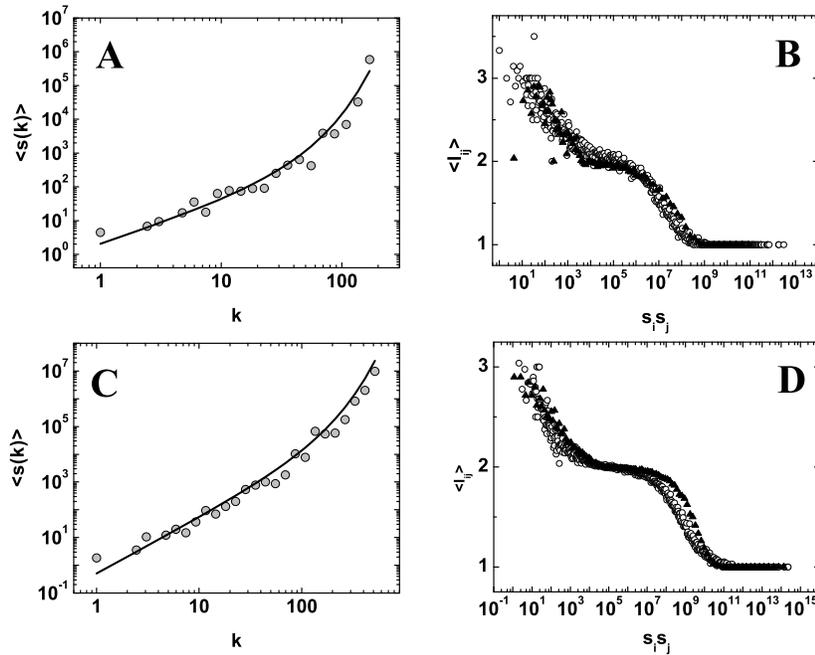

Fig. 2.   The average node's strength versus node's degree $\langle s(k) \rangle$ for Portal A (A) and portal B (C). The solid line is a fit to the empirical data $\langle s(k) \rangle = 2k^{1.2} \exp(k/30)$ (A) and $\langle s(k) \rangle = 0.5k^2 \exp(k/100)$ (C). The dependence of the internode distances $\langle l_{ij} \rangle$ on the product of nodes' strengths $s_i s_j$ (circles) and on the product of nodes' degrees $k_i k_j$ rescaled by the relation $\langle s(k) \rangle$ (triangles) for the Portal A (B) and Portal B (D).

3. Scaling in the absence of $s(k)$ correlations

The scaling relation observed in the absence of strength-degree correlations can be easily explained using the approach previously applied for internode distances in [10]. In fact, considering a branching process one obtains the following equation connecting the degrees of nodes i and j, the distance x between them, the size of

the network N, its average degree $\langle k \rangle$ and the average branching degree κ:

$$k_i \kappa^{x-1} = \frac{N\langle k \rangle}{k_j}. \tag{4}$$

Multiplying the nominator and denominator at both sizes of the equation by the average value of the weight in the network $\langle w \rangle$ and assuming that in the absence of the $s(k)$ correlations the following relation is satisfied

$$s_i = \langle w \rangle k_i, \tag{5}$$

one gets

$$s_i \kappa^{x-1} = \frac{N\langle k \rangle \langle w \rangle^2}{s_j}. \tag{6}$$

This equation immediately results in the form of the internode distances scaling relation as a function of nodes' strengths

$$\langle l_{ij} \rangle = \widetilde{a} - \widetilde{b} \log s_i s_j, \tag{7}$$

where coefficients $\widetilde{a}$ and $\widetilde{b}$ are given by

$$\widetilde{a} = 1 + \frac{\log(N\langle k \rangle \langle w \rangle^2)}{\log \kappa} \quad \text{oraz} \quad \widetilde{b} = \frac{1}{\log \kappa}. \tag{8}$$

The above considerations lead to the following conclusions: (a) the slope $\widetilde{b}$ of the scaling function is the same as in the unweighted case, (b) the relation between coefficients a and $\widetilde{a}$ takes the form

$$\widetilde{a} = a + \frac{\log \langle w \rangle^2}{\log \kappa}. \tag{9}$$

In order to compare the consistency of the Eqs (7) and (8) with the numerical simulations we performed the trails (see Figure 3) for the Erdős-Rényi random network (ER network) of $N = 1000$ nodes and $\langle k \rangle = 5$. The weights were randomly drawn from the Gauss distribution

$$p_w(w) = \frac{1}{\sqrt{2\pi}\sigma_w} \mathrm{e}^{-\frac{(w - \langle w \rangle)^2}{2\sigma_w^2}}, \tag{10}$$

characterized by the mean value $\langle w \rangle = 10$ and three different values of the standard deviation: $\sigma_w = 0.1$ (Figures 3A and 3D), $\sigma_w = 1$ (Figures 3B and 3E) and $\sigma_w = 3$ (Figures 3C and 3F). A comparison of those weight distributions is presented in Figure 3H. As one can see the width of the distribution plays the key role with respect to the shape of the strength distribution $P(s)$; in case when $\sigma_w = 0$ one gets of course the Poisson distribution

$$P(s) = \mathrm{e}^{-\langle k \rangle} \frac{\langle k \rangle^{\frac{s}{\langle w \rangle}}}{\left(\frac{s}{\langle w \rangle}\right)!}, \tag{11}$$

6   *Julian Sienkiewicz, Janusz A. Hołyst*

for $s = \langle w \rangle, 2\langle w \rangle, 3\langle w \rangle, ...$ In general, the strength distribution takes the form

$$P(s) = \sum_{k=1}^{\infty} p(k) \text{Prob} \left(\sum_{i=1}^{k} w_i = s \right), \tag{12}$$

assuming that weights w_i are i.i.d. random variables drawn from the distribution $p_w(w)$. Taking into account the considered distributions $p(k)$ and $p_w(w)$ one arrives at the relation

$$P(s) = \sum_{k=1}^{\infty} \frac{1}{\sqrt{2\pi k}\sigma_w} e^{-\frac{(s-\langle w \rangle k)^2}{2k\sigma_w^2}} e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!}. \tag{13}$$

The strength distributions taken from the numerical simulations for three different values of σ_w as well as the values $P(s)$ numerically summed using Eq. (13) are shown in Figures 3A, 3B and 3C. The characteristic oscillations seen in $P(s)$ distribution around the multiply of the average weight ($\langle w \rangle, 2\langle w \rangle, 3\langle w \rangle, ...$) vanish along with the increasing value of σ_w which in result gives the distribution a smooth shape. The form of the strength distribution changes the relation between the internode distances and the product of strengths (see Figures 3D–F). The relation bewteen the average node's strength and node's degree follows Eq. (5) and is shown in Figure 3G.

For $\sigma_w = 0.1$ (Figure 3D) the points obtained from the numerical simulations overlap those predicted by Eq. (7) with the coefficients $\widetilde{a}$ and $\widetilde{b}$ defined by the parameters of the ER network and the average weight $\langle w \rangle$. A comparison between relations $\langle l_{ij} \rangle (s_i s_j)$ and $\langle l_{ij} \rangle (k_i k_j)$ is presented in Figure 3I showing that indeed the slopes in both cases are identical as predicted by Eqs (2) and (8). The increase in the σ_w value leads to emergence of the oscillations on the $log(s_i s_j)$ trend line (Figure 3E) and eventually for large values of σ_w the scaling relations breaks down. This effect is illustrated in Figure 3F.

4. Scaling in the presence of $s(k)$ correlation

Section 2 contains examples of the scaling of internode distances on the product of nodes' strengths in the presence of non-linear $s(k)$ correlations. Previously proposed branching-tree model has no application in this case, as one should sum up all the weights on the tree's surface, which is impossible for any general form of $s(k)$. Certainly, if $s(k)$ has an explicit form, there is a possibility of reversing the relation, putting it into Eq. (4) and obtaining the coefficients $\widetilde{a}$ and $\widetilde{b}$ defining the $\langle l_{ij} \rangle$ relation as a function of $s_i s_j$. For instance, if $s(k) = Ak^\alpha$, reversing gives $k = (s/A)^{1/\alpha}$ and as a result the rescaled coefficients $\widetilde{a}$ and $\widetilde{b}$ will take the form

$$\widetilde{b} = \frac{b}{\alpha} \quad \text{and} \quad \widetilde{a} = a + 2\log A\widetilde{b}. \tag{14}$$

Such an example is presented in Figure 1B (solid line) where, after getting coefficients a and b from fitting to relation (1) they were put into Eqs 14) thus obtaining the values $\widetilde{a}$ and $\widetilde{b}$.
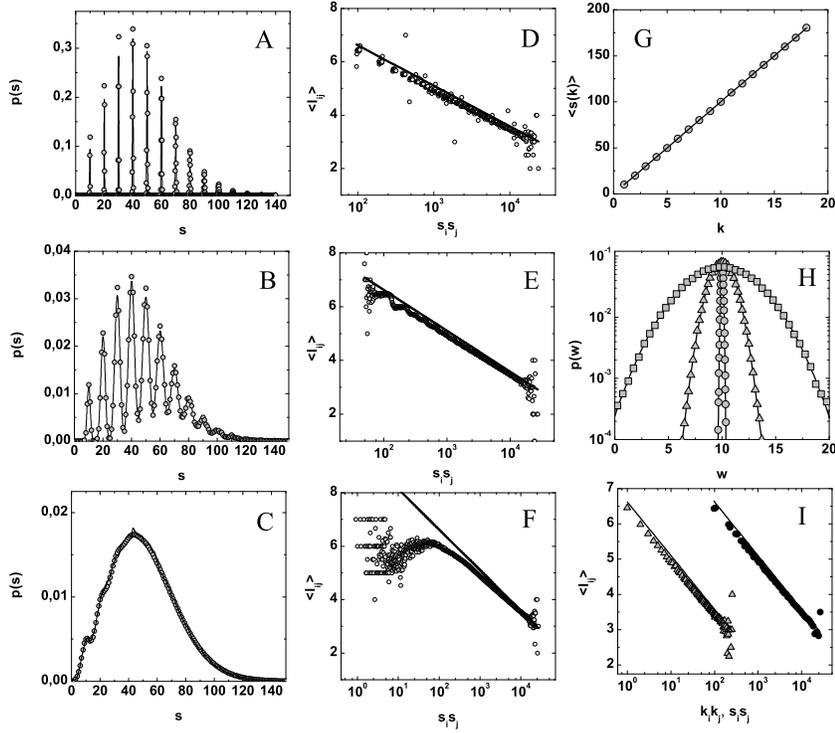
Fig. 3.   Numerical simulations of the ER network with $N = 1000$ and $\langle k \rangle = 5$ and the weights coming from Gauss distribution (10) with $\langle w \rangle = 10$ and different values of standard deviation σ_w. (A–C) Strength distributions for $\sigma_w = 0.1$ (A), $\sigma_w = 1$ (B) and $\sigma_w = 3$ (C); solid lines correspond to numerically summed Eq. (13). (D–F) Internode distances versus the product of strengths for $\sigma_w = 0.1$ (D), $\sigma_w = 1$ (E) and $\sigma_w = 3$ (F); solid lines correspond to Eq. (7). (G) The node's strength versus its degree; solid line corresponds to equation $s(k) = 10k$. (H) Weight distributions for $\sigma_w = 0.1$ (circles), $\sigma_w = 1$ (triangles) and $\sigma_w = 3$ (squares); solid lines come from Eq. (10). (I) A comparison between relations $\langle l_{ij} \rangle (s_i s_j)$ (circles) and $\langle l_{ij} \rangle (k_i k_j)$ (triangles) for $\sigma_w = 0.1$; solid lines come from Eqs (1) and (7).

For more systematical study we performed numerical simulations for ER network with $N = 1000$ nodes and $\langle k \rangle = 5$, with specific $s(k)$ relations. To generate the correlations between node's strength and its degree we used the observations from work [2]: if in the given network there are no degree-degree correlations (the asortativity coefficient r is equal to 0) and the edge between nodes i and j has weight $\langle w_{ij} \rangle$, proportional to the expression $(k_i k_j)^\alpha$ then the strength of node i is

$$s_i \sim k_i \langle w_{ij} \rangle \sim k_i^{1+\alpha}. \tag{15}$$

Using the above described procedure, we imposed weights onto ER network according to relations $s(k) = 130k^3$ (Figure 4A) and $s(k) = 80k \sin(\pi k/18)$ (Figure 4C). As one can observe in Figures 4B and 4D the expected overlap of relation $\langle l_{ij} \rangle$ as a function of $s_i s_j$ (circles) with the rescaled data of $\langle l_{ij} \rangle (k_i k_j)$ (triangles) was ob-

tained. Moreover, in Figure 4B there can be seen a good fitting of data with linear function determined by the coefficients taken from the Eqs (14).
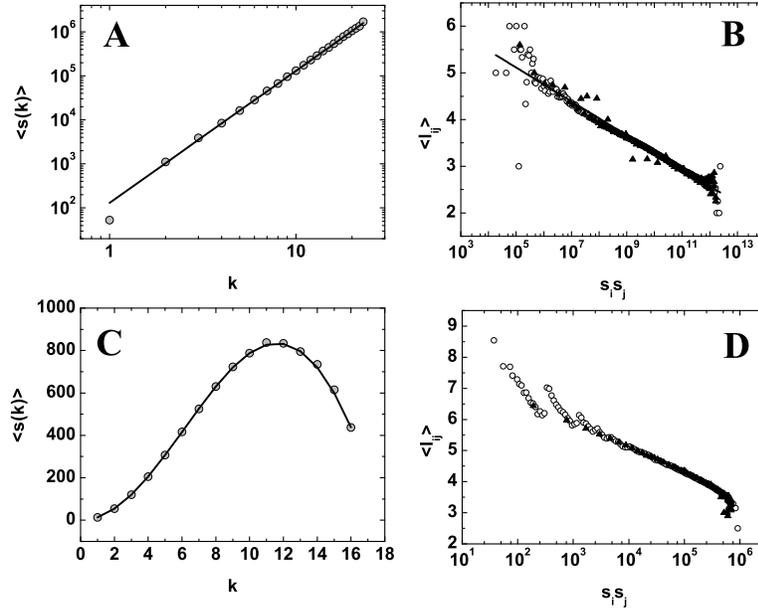

Fig. 4. Plots A and C: the average node's strength versus its degree k. Plots B and D: the internode distance versus the product of nodes' degrees $k_i k_j$ rescaled by the relation between node's strength and its degree (triangles); the internode distance versus the product of nodes' strengths $s_i s_j$ (circles). Plots A and B were obtained for ER network with $N = 1000$, $\langle k \rangle = 5$ where weights fulfill the equation $w_{ij} \sim (k_i k_j)^2$, while in the case of plots C and D a corresponding relation was $w_{ij} \sim \sin(k_i)\sin(k_j)$.

5. Conclusions

The relation between internode distances and the product of the nodes' strengths is an extension of the previously examined dependence linking $\langle l_{ij} \rangle$ and $k_i k_j$. As expected the scaling takes a similar form while the coefficients a and b are altered by the relation between average node's strength and its degree. In the case of explicit equation for $s(k)$ (e.g. linear or scale-free), the new coefficients $\widetilde{a}$ and $\widetilde{b}$ can be easily obtained. A breakdown in scaling relation while increasing the width of the weight distribution has also been observed.

Acknowledgments

References

1. S. H. Yook, H. Jeong, A.-L. Barabási, Phys. Rev. Lett. **86**, 5835 (2001).
2. A. Barrat, M. Barthélemy, R. Pastor-Satorras, A. Vespignani, Proc. Natl. Acad. Sci. U.S.A. **101**, 3747 (2004).
3. J. M. Kumpula, J.-P. Onnela, J. Saramäki, K. Kaski, J. Kertész, Phys. Rev. Lett. **99**, 228701 (2007).
4. Z. Wu, L. A. Braunstein, S. Havlin, H. E. Stanley, Phys. Rev. Lett. **96**, 148702 (2006).
5. J. J. Ramasco, B. Gonçalves, Phys. Rev. E **76**, 066106 (2007).
6. A. Barrat, M. Barthélemy, A. Vespignani, Phys. Rev. Lett. **92**, 228701 (2004).
7. W.-X. Wang, B.-H. Wag, B. Hu, G. Yan, Q. Ou, Phys. Rev. Lett. **94**, 188702 (2005).
8. Y. Chen, E. Lopez, S. Havlin, H. E. Stanley, Phys. Rev. Lett. **96**, 068702 (2006).
9. D. Garlaschelli, M. I. Loffredo, Phys. Rev. Lett. **102**, 038701 (2009).
10. J. A. Hołyst, J. Sienkiewicz, A. Fronczak, P. Fronczak, K. Suchecki, Phys. Rev. E **72**, 026108 (2005).
11. J. A. Hołyst, J. Sienkiewicz, A. Fronczak, P. Fronczak, K. Suchecki, Physica A **351**, 167 (2005).
12. J. Sienkiewicz, J. A. Hołyst, Phys. Rev. E **72**, 046127 (2005).
13. A. Chmiel, K. Kowalska, J. A. Hołyst, Phys. Rev. E **80**, 066122 (2009).
14. J. Sienkiewicz, P. Fronczak, J. A. Hołyst, Phys. Rev. E **75**, 066102 (2007).